

# Systems & Control *Transactions*

**Volume 3**

Proceedings of the 10th International Conference on  
Foundations of Computer-Aided Process Design

**FOCAPD 2024**



Editors

**Thomas A. Adams II**

**Matt Bassett**

**Selen Cremaschi and**

**Monica Zanfir**

---

# Systems & Control Transactions

**Vol 3**

Proceedings of the 10th International Conference on  
Foundations of Computer-Aided Process Design  
**FOCAPD 2024**

July 14-18, 2024  
Brekenridge, Colorado, USA

---



[PSEcommunity.org](https://PSEcommunity.org)

PSE Press • Hamilton • Notre Dame



Systems & Control Transactions, Volume 3  
PSE Press  
Hamilton, Ontario, Canada

Individual articles copyright © 2024 by the authors and licensed to PSEcommunity.org and PSE Press.  
Remaining text © 2024 by the PSE Press. All rights reserved.  
Cover photo credit: Photo 119650316 © Scanrail | Dreamstime.com

This book and all individual articles are released under the Creative Commons CC-BY-SA 4.0 License  
Credit must be given to creator and adaptations must be shared under the same terms.  
See <https://creativecommons.org/licenses/by-sa/4.0/> for license details.



**Living Archive for Process Systems Engineering (LAPSE) Archive ID:** LAPSE:2024.1500  
**Digital Object Identifier (DOI):** 10.69997/sct.108451

#### **Library and Archives Canada Catalogue**

A catalogue record and copy of this book is available from Library and Archives Canada

#### **Archival Permissions**

This book as well as individual articles may be deposited in digital research archives, personal archives, or institutional archives.

Electronic Book

ISBN (Volume 3) 978-1-7779403-2-4

ISSN (Series): 2818-4734

10 9 8 7 6 5 4 3 2 1

The information presented in this book was provided by individual contributors in the form of new scientific research contributions, data, and findings that were believed to be reliable and provided in good faith. Some, but not all, content has undergone peer review as marked. Neither the PSE Press, PSEcommunity.org, the American Institute of Chemical Engineers, the CACHE Foundation, Inc., the editors, nor authors guarantee the accuracy or completeness of any information provided, nor shall they be responsible for errors, omissions, incomplete information, or damages arising from its use. The information within does not consist of engineering or professional services.

Typeset in Inter

#### **Suggested Citations for Individual Articles**

##### *Paper Style*

Student MC, Supervisor QR. Title of the article. Systems & Control Trans, Vol 3, 23-28 (2024).

##### *Proceedings Style*

Student MC, Supervisor QR. Title of the article. In: Proceedings of the 10th International Conference on Foundations of Computer-Aided Process Design, July 14-18, Breckenridge, CO, USA, 2024. *Page Range* ISBN: 9781777940324

#### **Editors**

Thomas A. Adams II, PhD, PEng. *Norwegian University of Science and Technology*

Matt Bassett, PhD, *Coupa Software*

Selen Cremaschi, PhD. *Auburn University*

Monica Zanfir, PhD, *Linde*

# Table of Contents

## PREFACE

<b>Introduction</b> .....	<b>xi</b>
Thomas A. Adams II, Matt Bassett, Selen Cremaschi, and Monica Zanfir	
<b>Peer Review Policy</b> .....	<b>xii</b>
<b>International Scientific Committee</b> .....	<b>xiv</b>

## PART 1: PEER REVIEWED ARTICLES

### Section 1: Invited Plenary and Keynotes

<b>Thermo-Mechanical Exergy of a Substance in Cold Applications Approaching Absolute Zero</b> .....	<b>2</b>
Thomas A. Adams II	
<b>Designing Process Systems for Net-Zero Emissions and Nature- and People-Positive Decisions</b> .....	<b>10</b>
Bhavik R. Bakshi	
<b>From Then to Now and Beyond: Exploring How Machine Learning Shapes Process Design Problems</b> .....	<b>16</b>
Burcu Beykal	
<b>Artificial Intelligence and Machine Learning for Sustainable Molecular-to-Systems Engineering</b> .....	<b>22</b>
Alexander W. Dowling	
<b>Connecting the Dots: Push and Pull between Technology R&amp;D and Energy Transition Modeling</b> .....	<b>32</b>
Justin A. Federici, Dimitri J. Papageorgiou, Robert D. Nielsen	
<b>Sustainable Process Systems Engineering for Chemicals within Planetary Boundaries</b> .....	<b>39</b>
Gonzalo Guillén-Gosálbez	
<b>CO<sub>2</sub> Mitigation in Chemical Processes: Role of Process Recycle Optimization</b> .....	<b>44</b>
Diane Hildebrandt, James Fox, Neil Stacey, Baraka C. Sempuga	
<b>Towards a Sustainable and Defossilized/Decarbonized Chemical and Process Industry</b> .....	<b>52</b>
Mariano Martín	
<b>Life Cycle and Sustainability Analyses for Designing Chemical Circular Economy</b> .....	<b>60</b>
David Perez, John D. Chea, Jose D. Hernandez-Betancur, Gerardo J. Ruiz-Mercado	
<b>Towards the Development of Digital Twin for Pharmaceutical Manufacturing</b> .....	<b>67</b>
Katherine Raudenbush, Nikola Malinov, Jayanth V. Reddy, Chaoying Ding, Huayu Tian, Marianthi Ierapetritou	
<b>Towards 3-fold sustainability in biopharmaceutical process development and product distribution</b> .....	<b>75</b>
Miriam Sarkis, Steven Sachio, Nilay Shah, Cleo Kontoravdi, Maria M. Papathanasiou	
<b>Mining Chemical Process Information from Literature for Generative Process Design: A Perspective</b> .....	<b>84</b>
Artur M. Schweidtmann	
<b>From Process to Systems Design: A Perspective on the Future of Design Education</b> .....	<b>92</b>

Victor M. Zavala

## Section 2: Advances in PSE Design

<b>Development of Mass/Energy Constrained Sparse Bayesian Surrogate Models from Noisy Data.....</b>	<b>99</b>
Samuel Adeyemo, Debangsu Bhattacharyya	
<b>Guaranteed Error-bounded Surrogate Framework for Solving Process Simulation Problems.....</b>	<b>105</b>
Chinmay M. Aras, Ashfaq Iftakher, M. M. Faruque Hasan	
<b>Process Flowsheet Optimization with Surrogate and Implicit Formulations of a Gibbs Reactor .....</b>	<b>113</b>
Sergio I. Bugosen, Carl D. Laird, Robert B. Parker	
<b>A Novel Cost-Efficient Tributyl Citrate Production Process.....</b>	<b>121</b>
Andres F. Cabeza, Alvaro Orjuela, David E. Bernal Neira	
<b>Graph-Based Representations and Applications to Process Simulation.....</b>	<b>129</b>
Yoel R. Cortés-Peña, Victor M. Zavala	
<b>Improved Design of Flushing Process for Multi-Product Pipelines .....</b>	<b>137</b>
Barnabas Gao, Swapana Jerpoth, David Theuma, Sean Curtis, Steven Roth, Michael Fracchiolla, Robert Hesketh, C. Stewart Slater, Kirti M. Yenkie	
<b>Advances in Process Synthesis: New Robust Formulations.....</b>	<b>145</b>
Smitha Gopinath, Claire S. Adjiman	
<b>Integration of Design and Operation with Discretization Error Control.....</b>	<b>153</b>
Christian Hoffmann, Erik Esche, Jens-Uwe Repke	
<b>Simultaneous Optimization of Design and Operating Conditions for RPB-based CO<sub>2</sub> Capture Process.....</b>	<b>160</b>
Howoun Jung, NohJin Park, Jay H. Lee	
<b>Beyond Yield: Assessing Reaction System Performance using Economics .....</b>	<b>167</b>
Mary A. Katebah, Ma'moun Al-Rawashdeh, Patrick Linke	
<b>Adsorption-based Atmospheric Water Extraction Process: Kinetic Analysis and Stochastic Optimization .....</b>	<b>173</b>
Jinsu Kim, Shubham Jamdade, Yanhui Yuan, Matthew J. Realf	
<b>Optimal Process Synthesis Implementing Phenomena-based Building Blocks and Structural Screening .....</b>	<b>179</b>
David Krone, Erik Esche, Mirko Skiborowski, Jens-Uwe Repke	
<b>Cybersecurity, Image-Based Control, and Process Design and Instrumentation Selection.....</b>	<b>186</b>
Dominic Messina, Akkarakaran Francis Leonard, Ryan Hightower, Kip Nieman, Renee O'Neill, Paloma Beacham, Katie Tyrrell, Muhammad Adnan, Helen Durand	
<b>Design Space Identification of the Rotary Tablet Press.....</b>	<b>194</b>
Mohammad Shahab, Sunidhi Bachawala, Marcial Gonzalez, Gintaras Reklaitis, Zoltan Nagy	
<b>Recent Advances of PyROS: A Pyomo Solver for Nonconvex Two-Stage Robust Optimization in Process Systems Engineering.....</b>	<b>201</b>
Jason A. F. Sherman, Natalie M. Isenberg, John D. Sirola, Chrysanthos E. Gounaris	
<b>Optimal Design Approaches for Cost-Effective Manufacturing and Deployment of Chemical Process Families with Economies of Numbers.....</b>	<b>208</b>
Georgia Stinchfield, Sherzoy Jan, Joshua C. Morgan, Miguel Zamarripa, Carl D. Laird	
<b>A Study on Accelerated Convergence of Cyclic Steady State in Adsorption Process Simulations .....</b>	<b>215</b>
Sai Gokul Subraveti, Kian Karimi, Matteo Gazzani, Rahul Anantharaman	



<b>Optimal Design of Intensified Towers for CO<sub>2</sub> Capture with Internal, Printed Heat Exchangers .....</b>	<b>222</b>
Stephen Summits, Paul Akula, Debangsu Bhattacharyya, Grigorios Panagakos, Benjamin Omell, Michael Matuszewski	
<b>Design of Plastic Waste Chemical Recycling Process Considering Uncertainty .....</b>	<b>229</b>
Zhifei Yuliu, Yuqing Luo, Marianthi Ierapetritou	
<b>Section 3: Design and Emerging Fields</b>	
<b>A GRASP Heuristic for Solving an Acquisition Function Embedded in a Parallel Bayesian Optimization Framework .....</b>	<b>237</b>
R. Cory Allen, Youngdae Kim, Dimitri J. Papageorgiou	
<b>Learn-To-Design: Reinforcement Learning-Assisted Chemical Process Optimization .....</b>	<b>245</b>
Eslam G. Al-Sakkari, Ahmed Ragab, Mohamed Ali, Hanane Dagdougui, Daria C. Boffito, Mouloud Amazouz	
<b>Cost-optimal Selection of pH Control for Mineral Scaling Prevention in High Recovery Reverse Osmosis Desalination .....</b>	<b>253</b>
Oluwamayowa O. Amusat, Alexander V. Dudchenko, Adam A. Atia, Timothy Bartholomew	
<b>Hybrid Rule-based and Optimization-driven Decision Framework for the Rapid Synthesis of End-to-End Optimal (E2EO) and Sustainable Pharmaceutical Manufacturing Flowsheets.....</b>	<b>261</b>
Yash Barhate, Daniel Casas-Orozco, Daniel J. Laky, Gintaras V. Reklaitis, Zoltan K. Nagy	
<b>Neural Networks for Prediction of Complex Chemistry in Water Treatment Process Optimization.....</b>	<b>267</b>
Alexander V. Dudchenko, Oluwamayowa O. Amusat	
<b>Improving Mechanistic Model Accuracy with Machine Learning Informed Physics .....</b>	<b>275</b>
William Farlessyost, Shweta Singh	
<b>Optimal Design of Antibody Extraction Systems using Protein A Resin with Multicycling .....</b>	<b>283</b>
Fred Ghanem, Purnima M. Kodate, Gerard M. Capellades, Kirti M. Yenkie	
<b>Exploring Quantum Optimization for Computer-aided Molecular and Process Design .....</b>	<b>292</b>
Ashfaq Iftakher, M. M. Faruque Hasan	
<b>Learning Hybrid Extraction and Distillation using Phenomena-based String Representation.....</b>	<b>300</b>
Jianping Li	
<b>Optimizing Batch Crystallization with Model-based Design of Experiments .....</b>	<b>308</b>
Hailey G. Lynch, Aaron Bjarnason, Daniel J. Laky, Cameron J. Brown, Alexander W. Dowling	
<b>Machine Learning-Aided Process Design for Microwave-Assisted Ammonia Production.....</b>	<b>316</b>
Md Abdullah Al Masud, Alazar Araia, Yuxin Wang, Jianli Hu, Yuhe Tian	
<b>Fast, Accurate, and Robust Fault Detection and Diagnosis of Industrial Processes.....</b>	<b>322</b>
Alireza Miraliakbar, Zheyu Jiang	
<b>Development of Steady-State and Dynamic Mass-Energy Constrained Neural Networks using Noisy Transient Data .....</b>	<b>330</b>
Angan Mukherjee, Debangsu Bhattacharyya	
<b>Model Based Process Development and Operation of a Fluid Bed Granulation Unit to Manufacture Pharmaceutical Tablets. ....</b>	<b>338</b>
Salvador García Muñoz, Maitraye Sen, Shaswat Gupta, Ronald Ruff	

<b>Modeling hiPSC-to-Early Cardiomyocyte Differentiation Process using Microsimulation and Markov Chain Models .....</b>	<b>344</b>
Shenbageshwaran Rajendiran, Francisco Galdos, Carissa Anne Lee, Sidra Xu, Justin Harvell, Shireen Singh, Sean M. Wu, Elizabeth A. Lipke, Selen Cremaschi	
<b>Integrating Hybrid Modeling and Multifidelity Approaches for Data-Driven Process Model Discovery .....</b>	<b>351</b>
Suryateja Ravutla, Fani Boukouvola	
<b>Technoeconomic and Sustainability Analysis of Batch and Continuous Crystallization for Pharmaceutical Manufacturing .....</b>	<b>359</b>
Jungsoo Rhim, Zoltan Nagy	
<b>Enhancing Polymer Reaction Engineering Through the Power of Machine Learning .....</b>	<b>367</b>
Habibollah Safari, Mona Bavarian	
<b>Integrated Design, Control, and Techno-Ecological Synergy: Application to a Chloralkali Process .....</b>	<b>373</b>
Utkarsh Shah, Akshay Kudva, Kevin B. Donnelly, Wei-Ting Tang, Bhavik R. Bakshi, Joel A. Paulson	
<b>Use of Discrete Element Method to Troubleshoot Aesthetic Defects in Pharmaceutical Tablets.....</b>	<b>380</b>
Jerrin Job Sibychan, Nicola Sorace, Jason Melnick, Salvador Garcia Muñoz, David Mota-Aguilar, Eduardo Hernandez-Torres, David Boush	
<b>Reinforcement Learning-Driven Process Design: A Hydrodealkylation Example .....</b>	<b>387</b>
Yuhe Tian, Ayooluwa Akintola, Yazhou Jiang, Dewei Wang, Jie Bao, Miguel A. Zamarripa, Brandon Paul, Yunxiang Chen, Peiyuan Gao, Alexander Noring, Arun Iyengar, Andrew Liu, Olga Marina, Brian Koeppel, Zhijie Xu	
<b>Section 4: Design and Energy Transitions</b>	
<b>Equation-Oriented Modeling of Water-Gas Shift Membrane Reactor for Blue Hydrogen Production.....</b>	<b>395</b>
Damian T. Agi, Hani A. E. Hawa, Alexander W. Dowling	
<b>Modeling the Maximization of Waste Heat Use in a Liquid Solvent Direct Air Capture Plant Through Hydrogen Production.....</b>	<b>403</b>
Erick O. Arwa, Kristen R. Schell	
<b>Technoeconomic Analysis of Chemical Looping Ammonia Synthesis Reactors to Enable Green Ammonia Production .....</b>	<b>409</b>
Laron D. Burrows, George M. Bollas	
<b>Optimal Design and Control of Behind-the-Meter Resources for Retail Buildings with EV Fast Charging .....</b>	<b>417</b>
Gustavo Campos, Roberto Vercellino, Darice Guittet, Margaret Mann	
<b>Optimization of Retrofit Decarbonization in Oil Refineries .....</b>	<b>426</b>
Samprii Chattopadhyay, Rahul Gandhi, Ignacio E. Grossmann, Ana I. Torres.	
<b>Conceptual Design of Integrated Energy Systems with Market Interaction Surrogate Models .....</b>	<b>434</b>
Xinhe Chen, Radhakrishna Tumbalam-Gooty, Darice Guittet, Bernard Knueven, John D. Siirola, Alexander W. Dowling	
<b>Preliminary Examination of the Biogas-to-Hydrogen Conversion Process.....</b>	<b>442</b>
Hegwon Chung, Minseong Park, Jiyong Kim	
<b>Optimization of Solid Oxide Electrolysis Cell Systems Accounting for Long-Term Performance and Health Degradation .....</b>	<b>448</b>
Nishant V. Giridhar, Debansu Bhattacharyya, Douglas A. Allan, Stephen E. Zitney, Mingrui Li, Lorenz T. Biegler	
<b>Power System Design and Necessary Changes to Accommodate Future Energy Challenges.....</b>	<b>455</b>

Iiro Harjunkoski, Katarina Knezovic, Alexandre Oudalov

<b>Towards Energy and Material Transition Integration – A Systematic Multi-scale Modeling and Optimization Framework</b> .....	461
Rahul Kakodkar, Betsie Montano Flores, Marco De Sousa, Yilun Lin, Efstratios N. Pistikopoulos	
<b>Process and Network Design for Sustainable Hydrogen Economy</b> .....	469
Monzure-Khoda Kazi, Akhilesh Gandhi, M. M. Faruque Hasan	
<b>Integrated Temporal Planning for Design and Operation of the International Green Ammonia Supply Chain</b> .....	477
Sunwoo Kim, Joungho Park, Jay H. Lee	
<b>Integration of a Chemical Heat Pump with a Post- combustion Carbon Capture Sorption Unit</b> .....	484
Rajalakshmi Krishnadoss, Thomas A. Adams II	
<b>IDAES-PSE Software Tools for Optimizing Energy Systems and Market Interactions</b> .....	490
Daniel J. Laky, Radhakrishna Tumbalam Gooty, Tyler Jaffe, Marcus Holly, Adam Atia, Xinhe Chen, Alexander W. Dowling	
<b>Design and Optimization of Processes for Recovering Rare Earth Elements from End-of-Life Hard Disk Drives</b>	496
Chris Laliwala, Ana I. Torres	
<b>NMPC for Mode-Switching Operation of Reversible Solid Oxide Cell Systems</b> .....	504
Mingrui Li, Douglas A. Allan, San Dinh, Lorenz T. Biegler, Debangsu Bhattacharyya, Vibhav Dabadghao, Nishant Giridhar, Stephen E. Zitney	
<b>Integrated Design and Scheduling Optimization of Multi-product processes – case study of Nuclear-Based Hydrogen and Electricity Co-Production</b> .....	511
Ruaridh Macdonald, Dharik S. Mallapragada	
<b>Stochastic Programming Models for Long-Term Energy Transition Planning</b> .....	519
Molly A. McDonald, Christos T. Maravelias	
<b>An MINLP Formulation for Global Optimization of Heat Integration-Heat Pump Assisted Distillations</b> .....	527
Akash Nogaja, Mohit Tawarmalani, Rakesh Agrawal	
<b>Role of Hydrogen as Fuel in Decarbonizing US Clinker Manufacturing for Cement Production: Costs and CO<sub>2</sub> Emissions Reduction Potentials</b> .....	533
Ikenna J. Okeke, Sachin U. Nimbalkar, Kiran Thirumaran, Joe Cresko	
<b>Comparative Techno-economic Assessment of Hydrogen Production, Storage and Refueling Pathways</b> .....	541
Minseong Park, Hegwon Chung, Jiyong Kim	
<b>Impact of surrogate modeling in the formulation of pooling optimization problems for the CO<sub>2</sub> point sources</b> ..	546
HA Pedrozo, MA Zamarripa, JP Osorio Suárez, A Uribe-Rodríguez, MS Diaz, LT Biegler	
<b>Optimal Clustered, Multi-modal CO<sub>2</sub> Transport Considering Non-linear Costs – a Path-planning Approach</b> .....	554
Kang Qiu, Sigmund Eggen Holm, Julian Straus, Simon Roussanaly	
<b>Promising Opportunities for Improving Round-Trip Efficiencies in Liquid Air Energy Storage (LAES)</b> .....	562
Siyue Ren, Truls Gundersen, Xiao Feng	
<b>The Impact of Electrified Process Heating on Process Design, Control and Operations</b> .....	570
Jong Hyun Rho, Michael Baldea, Elizabeth E. Endler, Monica A. Herediac, Vesna Bojovic, Pejman Pajand	
<b>Biogas Valorization from a Process Synthesis Perspective: Heat and Work Integration to Maximize CO<sub>2</sub> Conversion</b> .....	578
Baraka C. Sempuga, Selusiwe Ncube	



<b>Towards Designing Sector-Coupled Energy Systems Within Planetary Boundaries .....</b>	<b>585</b>
David Y. Shu, Jan Hartmann, Christian Zibunas, Nils Baumgärtner, Niklas von der Assen, André Bardow	
<b>Sustainable Green Hydrogen Transport: A Systematic Framework for the Design of the whole Supply Chain...</b>	<b>591</b>
Elvira Spatolisano, Laura A. Pellegrini	
<b>A mathematical programming optimization framework for wind farm design considering multi-directional wake effect .....</b>	<b>598</b>
Javiera Vergara-Zambrano, Styliani Avraamidou	
<b>RiNSES4: Rigorous Nonlinear Synthesis of Energy Systems for Seasonal Energy Supply and Storage .....</b>	<b>604</b>
Yifan Wang, Marvin Volkmer, Dörthe Franzisca Hagedorn, Christiane Reinert, Niklas von der Assen	
<b>Simulation and Comparative Analysis of Conventional Steam-Methane Reforming Models for Reactor Electrification.....</b>	<b>612</b>
Yufei Zhao, Chengtian Cui, Cornelius. M. Masuku	
<b>Section 5: Design and Sustainability</b>	
<b>Machine Learning Methods for the Forecasting of Environmental Impacts in Early-stage Process Design .....</b>	<b>621</b>
Emmanuel A. Aboagye, Austin L. Lehr, Ethan Shumaker, Jared Longo, John Pazik, Robert P. Hesketh, Kirti M. Yenkie	
<b>Design and Optimization of Circular Economy Networks: A Case Study of Polyethylene Terephthalate (PET) ..</b>	<b>629</b>
Abdulhakeem Ahmed, Ana I. Torres	
<b>Economic Optimization and Impact of Utility Costs on the Optimal Design of Piperazine-Based Carbon Capture .....</b>	<b>635</b>
Ilayda Akkor, Shachit S. Iyer, John Dowdle, Le Wang, Chrysanthos Gounaris	
<b>The design and operational space of syngas production via integrated direct air capture with gaseous CO<sub>2</sub> electrolysis .....</b>	<b>641</b>
Hussain M. Almajed, Omar J. Guerra, Ana Somoza-Tornos, Wilson A. Smith, Bri-Mathias Hodge	
<b>Towards Sustainable Supply Chains for Waste Plastics through Closed-Loop Recycling: A case-study for Georgia.....</b>	<b>652</b>
Elisavet Anglou, Riddhi Bhattacharya, Patricia Stathatou, Fani Boukouvala	
<b>Environmental Impact of Simulated Moving Bed (SMB) on the Recovery of 2,3-Butanediol on an Integrated Biorefinery .....</b>	<b>660</b>
Marco E. Avendano, Jianpei Lao, Qiang Fu, Sankar Nair, Matthew J. Realff	
<b>Design and Optimization of Methanol Production using PyBOUND.....</b>	<b>668</b>
Prapatsorn Borisut, Bianca Williams, Aroonsri Nuchitprasittichai, Selen Cremaschi	
<b>Biomanufacturing in Space: New Concepts and Paradigms for Process Design .....</b>	<b>674</b>
Brenda Cansino-Loeza, Vernon McIntosh, Krista Ternus, Victor M. Zavala	
<b>Techno economical assessment of a low-carbon hydrogen production process using residual biomass gasification and carbon capture.....</b>	<b>681</b>
E.J. Carrillo, J. Lizcano-Prada, V. Kafaro, D. Rodriguez-Vallejo, A. Uribe-Rodríguez	
<b>Designing Better Plastic Management Processes Through a Systems Approach.....</b>	<b>691</b>
John D. Chea, Matthew Conway, Gerardo J. Ruiz-Mercado, Pahola Thathiana Benavides, Kirti M. Yenkie	
<b>Resilient-aware Design for Sustainable Energy Systems.....</b>	<b>698</b>

Natasha J. Chrisandina, Shivam Vedant, Catherine Nkoutche, Eleftherios Iakovou, Efstratios N. Pistikopoulos, Mahmoud M. El-Halwagi

**Design and Optimization of a Multipurpose Zero Liquid Discharge Desalination Plant ..... 705**

Dev Barochia, Hasan Nikkhah, Burcu Beykal

**Opportunities for Process Intensification with Membranes to Promote Circular Economy Development for Critical Minerals ..... 711**

Molly Dougher, Laurianne Lair, Jonathan Aubuchon Ouimet, William A. Phillip, Thomas J. Tarka, Alexander W. Dowling

**Constraint Formulations for Bayesian Optimization of Process Simulations: General Approach and Application to Post-Combustion Carbon Capture ..... 719**

Clinton M. DUEWALL, Mahmoud M. El-Halwagi

**Industrial Biosolids from Waste to Energy: Development of Robust Model for Optimal Conversion Route – Case Study ..... 726**

Hesan Elfaki, Dhabia M. Al-Mohannadi

**Analysis of Infrastructures for Processing Plastic Waste using Pyrolysis-Based Chemical Upcycling Pathways ..... 732**

Evan D. Erickson, Jiaze Ma, Philip Tominac, Horacio Aguirre-Villegas, Victor M. Zavala

**Biofuels with Carbon Capture and Storage in the United States Transportation Sector ..... 738**

Caleb H. Geissler, Christos T. Maravelias

**Sustainable Production of Fertilizers via Photosynthetic Recovery of Nutrients in Livestock Waste ..... 744**

Leonardo D. González, Celeste Mills, Aurora del C. Munguía-López, Victor M. Zavala

**Techno-Economic Analysis of Methane Production from Pulp and Paper Sludge ..... 750**

Erfan Hosseini, Selen Cremaschi, Zhihua Jiang

**Integrated Ex-Ante Life Cycle Assessment and Techno-Economic Analysis of Biomass Conversion Technologies Featuring Evolving Environmental Policies ..... 757**

Dat T. Huynh, Marianthi Ierapetritou

**Screening Green Solvents for Multilayer Plastic Films Separation ..... 763**

Ugochukwu M. Ikegwu, Victor M. Zavala, Reid C. Van Lehn

**Integration of Chemical Looping Reforming and Shift Reactors for Blue H<sub>2</sub> and N<sub>2</sub> Production ..... 771**

Adrian R. Irhamna, George M. Bollas

**Model assessment for Design of Future Manufacturing systems using Digital Twins: A case study on a single-scale pharmaceutical manufacturing unit ..... 778**

Prem Jagadeesan, Shweta Singh

**Resource Integration Across Processing Clusters: Designing a Cluster of Clusters ..... 783**

Mohammad Lameh, Dhabia Al-Mohannadi, Patrick Linke

**Evaluating Circularity and Sustainability in Plastic Waste Recycling: Open and Closed-Loop Technologies ..... 792**

Wafaa N. Majzoub, Dhabia M. Al-Mohannadi

**Uncertainty and Complexity Considerations in Food-Energy-Water Nexus Problems ..... 799**

Marcello Di Martino, Patrick Linke, Efstratios N. Pistikopoulos

**Optimal Transition of Ammonia Supply Chain Networks via Stochastic Programming ..... 807**

Ilias Mitrai, Matthew J. Palys, Prodromos Daoutidis

<b>A Fast Computational Framework for the Design of Solvent-Based Plastic Recycling Processes</b> .....	<b>814</b>
Aurora del C. Munguía-López, Panzheng Zhou, Ugochukwu M. Ikegwu, Reid C. Van Lehn, Victor M. Zavala	
<b>Optimal Design of Food Packaging Considering Waste Management Technologies to Achieve Circular Economy</b> .....	<b>820</b>
Paola A. Munoz-Briones, Aurora del C. Munguía-López, Kevin L. Sánchez-Rivera, Victor M. Zavala, George W. Huber, Styliani Avraamidou	
<b>Integrating the Design of Desalination Technologies into Produced Water Network Optimization</b> .....	<b>829</b>
Sakshi Naik, Miguel Zamarripa, Markus Drouven, Lorenz T. Biegler	
<b>Nature-inspired Bio-Mineral Refinery for Simultaneous Biofuel Feedstock production and CO<sub>2</sub> mineralization</b> .....	<b>836</b>
Pavan Kumar Narahariseti	
<b>Deciphering the Policy-Technology Nexus: Enabling Effective and Transparent Carbon Capture Utilization and Storage Supply Chains</b> .....	<b>844</b>
Manar Y. Oqbi, Dhabia M. Al-Mohannadi.	
<b>Optimal Membrane Cascade Design for Critical Mineral Recovery Through Logic-based Superstructure Optimization</b> .....	<b>853</b>
Daniel Ovalle, Norman Tran, Carl D. Laird, Ignacio E. Grossmann	
<b>Membrane-based carbon capture process optimization using CFD modeling</b> .....	<b>860</b>
Hector A. Pedrozo, Cheick Dosso, Lingxiang Zhu, Victor Kusuma, David Hopkinson, Lorenz T. Biegler, Grigorios Panagakos	
<b>Enhancing PHAs Production Sustainability: Biorefinery Design through Carbon Source Diversity</b> .....	<b>868</b>
Fernando D. Ramos, Matías H. Ramos, Vanina Estrada, M. Soledad Diaz	
<b>Computer-Aided Mixture Design Using Molecule Superstructures</b> .....	<b>876</b>
Phillipp Rehner, Johannes Schilling, André Bardow	
<b>Internally Heated Crackers for Decarbonization and Optimization of Ethylene Production</b> .....	<b>883</b>
Edwin A. Rodriguez-Gil, Rakesh Agrawal	
<b>Mathematical Optimization of Separator Network Design for Sand Management</b> .....	<b>892</b>
Pooja Zen Santhamoorthy, Selen Cremaschi	
<b>Sustainable Process Systems Engineering – You’re Doing It Wrong!</b> .....	<b>899</b>
Raymond L. Smith	
<b>Designing Reverse Electrodialysis Process for Salinity Gradient Power Generation via Disjunctive Programming</b> .....	<b>904</b>
Carolina Tristán, Marcos Fallanza, Raquel Ibáñez, Ignacio E. Grossmann, David Bernal Neira	
<b>Optimal Design of a Biogas-based Renewable Power Production System</b> .....	<b>912</b>
Vikram Uday, Sujit Jogwar	
<b>Dimensionality Reduction in Optimal Process Design with Many Uncertain Sustainability Objectives</b> .....	<b>920</b>
Hongxuan Wang, Andrew Allman	
<b>Sustainable Aviation Fuels (SAF) from Ethanol: An Integrated Systems Modeling Approach</b> .....	<b>927</b>
Madelynn J. Watson, Aline V. da Silva, Pedro G. Machado, Celma O. Ribeiro, Cláudio A.O. Nascimento, Alexander W. Dowling	



**Exploring Net-Zero Greenhouse Gas Emission Routes for Bio-Production of Triacetic Acid Lactone: An Evaluation through Techno-Economic Analysis and Life Cycle Assessment..... 933**  
Ching-Mei Wen, Charles Foster, Marianthi Ierapetritou

**An Update on Project PARETO - New Capabilities in DOE's Produced Water Optimization Framework ..... 941**  
Miguel A. Zamarripa, Elmira Shamlou, Javal Vyas, Travis Arnold, Philip Tominac, Melody H. Shellman, Markus Drouven

## Section 6: Design Education and Future of Design

**Jacobian-based Model Diagnostics and Application to Equation Oriented Modeling of a Carbon Capture System ..... 950**  
Douglas A. Allan, Anca Ostace, Andrew Lee, Brandon Paul, Anuja Deshpande, Miguel A. Zamarripa, Joshua C. Morgan, Benjamin P. Omell

**Integration of Process Design and Intensification Learning via Combined Junior Course Project ..... 959**  
Madelyn R. Ball, Oishi Sanyal, Yuhe Tian

**Model Diagnostics for Equation-Oriented Models: Roadblocks and the Path Forward ..... 966**  
Andrew Lee, Robert B. Parker, Sarah Poon, Dan Gunter, Alexander W. Dowling, Bethany Nicholson

**Analysis of Chemical Engineering Curricula Using Graph Theory..... 975**  
Blake R. Lopez, Victor M. Zavala

**Laying the foundations of Machine Learning in Undergraduate Education through Engineering Mathematics .. 983**  
Pavan Kumar Narahariseti

## PART 2: EXTENDED ABSTRACTS

**Accelerating Discovery in Consumer Product Design..... 991**  
Brian T. Gettelfinger

**Process Design for the Energy Transition: An Industrial Perspective..... 993**  
Jaffer H. Ghose

**Designing for the Future: The Role of Process Design in Decarbonization and Energy Transition..... 994**  
M. M. Faruque Hasan

**Design for Flexibility: A Robust Optimization Approach ..... 997**  
Jnana Sai Jagana, Congqin Ge, Zhihong Yuan, Satyajith Amaran, Qi Zhang

**The Current Status of Process and Product Design – What and How to Teach – and a Vision for the Future ..... 998**  
Warren D. Seider, Daniel R. Lewin

**Forces Shaping the Future of Design and Design Education ..... 999**  
Jeffrey J Sirola

**Decarbonization of Oil Refineries through Electrification and Low-Carbon Feedstocks..... 1003**  
Ana I. Torres

**Challenges in Design for Sustainability..... 1005**  
Gavin Towler

**New Design Paradigm for Integrated Energy Systems Needed for Sustainability ..... 1006**  
David Tucker, N. Farida Harun, Biao Zhang, Nana Zhou, Samuel Bayham

**Design Education Across the Curriculum for the Future of Design ..... 1009**

## INDEX

<b>Author Index .....</b>	<b>1011</b>
<b>Keyword Index .....</b>	<b>1015</b>

# Introduction

**Thomas A. Adams II, Matt Bassett, Selen Cremaschi, and Monica Zanfir**

The Foundations of Computer Aided Process Design (FOCAPD) conference is the leading conference series in the field of chemical process design. Held only once every five years, FOCAPD is not a routine research conference. It is a special time for the process design community to come together and discuss the biggest trends in the field from a big picture perspective.

FOCAPD focuses on the big picture. All speakers are selected and invited by the conference organizers, and we use a single-track approach so that attendees gather together to hear all the speakers. Speakers are asked not to focus on their own research findings and accomplishments, but rather to talk about the latest trends and forward directions in their fields more generally. We want to know what is exciting about the field, and where it is headed next. We want the whole community at each talk so we can engage in lively discussions after each one about the future of our profession.

This year's program includes 42 invited keynote and plenary speakers and 121 contributing posters, covering the most important topics to FOCAPD. The conference proceedings boasts a total of 134 peer-reviewed papers containing the latest original research from our participants in five subject areas:

- Advances in PSE Design
- Design and Emerging Fields
- Design and Energy Transitions

- Design and Sustainability
- Design Education and Future of Design

This year, FOCAPD was organized jointly between the CACHE Foundation, inc. and the American Institute for Chemical Engineers. We thank our AIChE partners for their excellent support in conference administration management.

We also thank our many sponsors. At the Conference Partners level we thank Chemstations, the Chemical Engineering department at Auburn University, and Lilly. At the Conference Supporters level, we thank the Department of Chemical and Biological Engineering at the University of Wisconsin-Madison, the Chemical and Biological Engineering department at Princeton, the Davidson School of Chemical Engineering at Purdue University, The Texas A&M Energy Institute, Linde, and Corteva Agriscience. We also thank the National Science Foundation for a grant which paid for the registration fees of 41 students, postdocs, junior faculty, and early career researchers.

This year marked a new chapter in the FOCAPD series by moving our conference proceedings to the open access Systems and Control Transactions book series, published by PSE Press, an imprint of the CACHE Foundation. We believe this will increase the impact of the conference by removing paywall barriers to our research.

We wish you an inspiring conference and a memorable stay in Breckenridge!

## Conference Chairs and Proceedings Editors

**Dr. Thomas A. Adams II**  
Professor, P.Eng.  
Norwegian University of  
Science and Technology  
(NTNU)  
Department of Energy and  
Process Engineering



**Dr. Selen Cremaschi**  
Professor  
Auburn University  
Department of Chemical  
Engineering



**Dr. Matt Bassett**  
Senior Manager  
Coupa Software  
Supply Chain Design and  
Planning



**Dr. Monica Zanfir**  
Linde Engineering R&D  
Associate Director





# Peer Review Policy

## TERMINOLOGY

We use the following terminology:

- **Conference Paper:** This is a peer-reviewed research paper describing an **original** research contribution and published in the conference proceedings.
- **Extended Abstract:** This is an optional original paper of unspecified length that is contributed by an invited speaker.

## APPLICATION

- All conference papers are subject to peer review.
- Extended abstracts are not subject to peer review.

## PUBLICATIONS

- All conference papers which have passed peer review and have been accepted by the conference chairs will be published in the conference proceedings.
- All extended abstracts which have been accepted by the conference chairs will be published in the conference proceedings.
- The conference proceedings will clearly identify which publications are conference papers and which are extended abstracts, and therefore, whether or not they were subject to peer review.

## PEER REVIEW PROCESS

### Initial Screening

The conference chairs will screen the initial submissions (one-page abstracts) of contributed papers. At their option, chairs may choose to reject these submissions for any reason, but the primary reasons for rejection at this stage are (a) subject does not match conference themes and topics; (b) in appropriate, unprofessional, unethical, or dishonest material; (c) duplicate submissions; (d) abandoned submissions; (e) multiple submissions from same presenter; (f) obvious poor quality; (g) reasonable suspicion of machine-generated text; (h) need to limit acceptances because of conference space constraints. Chairs may also transfer papers between sessions at this stage.

Conference chairs will then invite authors of the one-page abstracts that have passed the initial screening to submit a conference paper for peer review.

### Conference Paper Peer Review

Conference chairs act as editors. They will assign peer reviewers to assess the quality of the conference papers and manage the peer review process for each paper. Peer reviewers are generally selected from the International Scientific Committee who are scientists, engineers, or researchers with technical expertise in the conference topic. Peer reviewers must be considered technical experts in their field, hold a PhD, and have published peer reviewed scientific works previously. Each work must be reviewed by a minimum of two peer-reviewers.

A “single-blind” peer review system is used, in which the peer reviewers have access to the identities of the authors, but the authors are never given the identities of the peer reviewers. This is consistent with most journal peer-review procedures in our field.

Peer reviewers will be given a set of quality criteria that may include scientific and technical quality, quality of writing, originality and novelty, appropriateness for the conference topic and theme, interest to the community, and other factors of merit. Peer reviewers are expected to be rigorous and critical in their technical assessments and adhere to the highest standards in the field in order to ensure high quality. Peer reviewers are asked to comment on their assigned conference papers and issue a recommendation to the conference chairs based on this quality criteria.

Reviewers may recommend one of the following:

- **Accept Paper.** No technical changes are necessary before publication. Only typographical, spelling, grammar, or other minor changes are necessary which do not require technical review.
- **Accept Paper with Minor Revisions.** Some minor technical issues need to be addressed either through technical changes to the manuscript or through rebuttal to reviewer comments. The reviewer does not believe the issues are significant enough to require additional technical review by the reviewers. The reviewer has provided enough commentary such that the editor can decide if the minor issues have been addressed in a future revision.
- **Revise Paper with Re-Review.** Major technical issues need to be addressed either through technical changes to the manuscript or through rebuttal to reviewer comments. The reviewer believes the issues are significant enough to require that the revised manuscript receive additional review by the reviewers.
- **Reject Without Reconsideration.** The technical issues are so significant that it is unlikely that an acceptable manuscript could be produced by the deadline; or, the manuscript is out of scope, inappropriate, or computer-generated.

Peer reviewers may be asked to provide numerical scores or rankings, as well as provide written comments intended for the chairs and/or the authors. Peer reviewers are encouraged to provide specific and constructive feedback that will aid the authors in improving the work and provide advice to the chairs.

The conference chairs are responsible for making the final decision on each paper and may require several rounds of author changes if necessary to meet quality standards. Conference chairs are not required to follow the recommendations of the peer reviewers in making these decisions. It is possible that conference chairs will choose to not accept some papers that still pass through rigorous technical peer review, especially when limited by available space or when papers do not sufficiently promote conference objectives. Conference chairs may also transfer papers between sessions as desired.

## CONFLICTS OF INTEREST AND ETHICAL GUIDELINES

The chairs and peer reviewers must ensure they do not have a conflict of interest that may bias their decisions, such as ensuring that authors are at “arm’s length” and have no financial conflicts of interest. Peer reviewers and chairs must disclose if they have such a conflict of interest with a specific submission, and if so, a different chair and/or peer reviewer should be assigned to handle that submission.

To help determine conflicts of interest, and for all other ethical guidelines, we use the Systems and Control Transactions ethical guidelines described at: <https://psecommunity.org/contributor-guidelines>

# International Scientific Committee

<b>Luke Achenie</b>	Virginia Polytechnic Institute
<b>Rakesh Agrawal</b>	Purdue University
<b>Styliani Avraamidou</b>	University of Wisconsin - Madison
<b>Bhavik Bakshi</b>	The Ohio State University
<b>Michael Baldea</b>	The University of Texas at Austin
<b>Ana Barbosa-Povoa</b>	Instituto Superior Tecnico
<b>Joule Bergersen</b>	University of Calgary
<b>Burcu Beykal</b>	University of Connecticut
<b>Debangsu Bhattacharyya</b>	West Virginia University
<b>Larry Biegler</b>	Carnegie Mellon University
<b>Rahul Bindlish</b>	Dow
<b>Fani Boukouvala</b>	Georgia Tech
<b>Richard Bratz</b>	MIT
<b>Kyle Camarda</b>	University of Kansas
<b>Alejandro Cano</b>	Siemens Process Systems Enterprise
<b>Leo Chiang</b>	Dow
<b>Giancarlo Dalle Ave</b>	Hitachi Energy Research
<b>Prodromos Daoutidis</b>	University of Minnesota
<b>Maria Soledad Diaz</b>	Universidad Nacional del Sur, Planta Piloto de Ingeniería Química
<b>Alexander Dowling</b>	Notre Dame
<b>Mario Eden</b>	Auburn University
<b>Mahmoud El-Halwagi</b>	Texas A&M University
<b>Ali Elkamel</b>	University of Waterloo
<b>Nor Farida Harun</b>	NETL Morgantown
<b>Tatiana Felix Ferreira</b>	Federal Univ. of Rio de Janeiro
<b>Eric Fraga</b>	University College London
<b>Rafiqul Gani</b>	PSE For Speed
<b>Sal Garcia</b>	Lilly
<b>Emre Gençer</b>	MIT Energy Initiative
<b>Jaffer Ghouse</b>	Woodside Energy
<b>Apostolos Giovanoglou</b>	Siemens Process Systems Enterprise
<b>Vipin Gopal</b>	Walgreens Boots Alliance
<b>Chrysanthos E.Gounaris</b>	Carnegie Mellon University
<b>Ignacio Grossmann</b>	Carnegie Mellon University
<b>Kevin Harding</b>	University of Witwatersrand
<b>Faruque Hasan</b>	Texas A&M
<b>Bri-Mathias Hodge</b>	University of Colorado Border & NREL
<b>Steve Honkomp</b>	P&G
<b>Yinlun Huang</b>	Wayne State University
<b>Marianthi Ierapetritou</b>	University of Delaware
<b>Zheyu Jiang</b>	Oklahoma State University
<b>Iftekar Karimi</b>	National University of Singapore
<b>Donghoi Kim</b>	SINTEF Energy
<b>Sophie (Sun Hye) Kim</b>	Dow
<b>Antonnis Kokossis</b>	National Technical University of Athens
<b>Carl Laird</b>	Carnegie Mellon University

<b>Ajay Lakshmanan</b>	Amazon Web Services
<b>Jay Lee</b>	University of Southern California
<b>Can Li</b>	Purdue University
<b>Patrick Linke</b>	Texas A&M University at Qatar
<b>Steve Lombardo</b>	OFI
<b>Christos Maravelias</b>	Princeton University
<b>Mariano Martin</b>	University of Salamanca
<b>David Miller</b>	OLI Systems
<b>Ruth Misener</b>	Imperial College London
<b>Alexander Mitsos</b>	RWTH Aachen University
<b>Nashaat N. Nassar</b>	University of Calgary
<b>Chinedu Okoli</b>	Proctor and Gamble
<b>Costas Pantelides</b>	Siemens Process Systems Enterprise
<b>Stratos Pistikopoulos</b>	Texas A&M University
<b>Sreekanth Rajagopalan</b>	Dow
<b>Edna Soraya Rawlings</b>	Sandia National Labs
<b>Matthew Realf</b>	Georgia Tech
<b>Rex Reklaitis</b>	Purdue University
<b>Gerardo Ruiz-Mercado</b>	US EPA
<b>Luis-Ricardez Sandoval</b>	University of Waterloo
<b>Warren Seider</b>	University of Pennsylvania
<b>Jeffrey Siirola</b>	Purdue University
<b>John Siirola</b>	Sandia National Lab
<b>Shweta Singh</b>	Purdue University
<b>Eva Sorensen</b>	University College London
<b>Yuhe Tian</b>	West Virginia University
<b>Ana I. Torres</b>	Carnegie Mellon
<b>Gavin Towler</b>	Honeywell
<b>Saadet Ulas Acikgoz</b>	Honeywell UOP
<b>Fernando V. Lima</b>	West Virginia University
<b>Venkat Venkatasubramanian</b>	Columbia University
<b>Xiaonan Wang</b>	Tsinghua University
<b>John Wassick</b>	CMU/Dow
<b>Kirti Yenkie</b>	Rowan University
<b>Victor M. Zavala</b>	University of Wisconsin - Madison



# Part 1

## Peer-Reviewed Articles

Section 1: Invited Plenary and Keynotes





# Thermo-Mechanical Exergy of a Substance in Cold Applications Approaching Absolute Zero

Thomas A. Adams II\*

Norwegian University of Science and Technology (NTNU), Department of Energy and Process Engineering, Trondheim, Norway

\* Corresponding Author: [thomas.a.adams@ntnu.no](mailto:thomas.a.adams@ntnu.no)

## ABSTRACT

In this work, we consider the thermo-mechanical exergy of a substance for cold applications, even as it approaches absolute zero. This is relevant for cold-service applications such as refrigeration, liquefied natural gas, air separation, and liquid hydrogen. We demonstrate how the optimization formulation for the determination of exergy is the most suitable way for process systems engineers to think about exergy. We provide an illustrative example by computing thermo-mechanical exergy of neon approaching absolute zero. We also discuss how this result relates with the Third Law of Thermodynamics, both how it is used to compute thermo-mechanical exergy, but also what it implies about the validity of the results and the equations used to compute them.

**Keywords:** Exergy, thermo-mechanical exergy, absolute zero, neon, low temperature

## 1. INTRODUCTION

**Exergy** is a thermodynamic property of matter and energy which describes both the quantity and the quality of an energy source by combining the First and Second Laws of Thermodynamics. It is quite useful in the design of chemical process systems, especially energy conversion systems common in chemical engineering. Exergy as a metric helps in identifying process bottlenecks, improving system efficiency, making prudent design decisions, understanding energy quality, and as an approximation for value or cost. As a result, it is increasingly popular as a systems analysis tool in scientific research, with exergy analyses considered in over 2500 journal articles published in 2022 alone (Figure 1), growing exponentially at about 14% per year.

### 1.1 General Definition

Although there are several competing definitions for exergy, the Moran *et al.* [1] definition is the most useful, the most general, and the easiest to understand for chemical engineering purposes. Our emphasis is in bold, and the definition is quoted as:

#### Definition 1. (Definition of Exergy)

Exergy is the **maximum** theoretical work obtainable from an overall system consisting of a

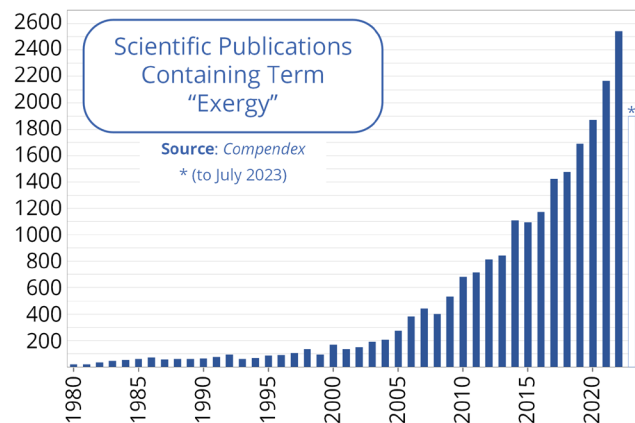


Figure 1. Scientific publications involving exergy

system and the environment as the system comes into equilibrium with the environment (passes to the dead state). [1]

This describes exergy generally, for all kinds of systems, and considering all forms of energy that contribute to the calculation of exergy. For the process systems engineer, it is convenient and useful to understand this definition in terms of an optimization problem [2]:

$$e = \max_P w \quad (1)$$

where  $e$  is the exergy of the system,  $\mathcal{P}$  is a process that brings the system into equilibrium with its environment within the set of all such processes  $\mathbb{P}$  such that  $\mathcal{P} \in \mathbb{P}$ , and  $w$  is work that is produced by  $\mathcal{P}$ . Process engineers understand optimization, and so it makes sense then to understand exergy in the form of an optimization problem. In plain language, it means that one way to compute the exergy is to find a process  $\mathcal{P}$  that produces the absolute most amount of work possible, exploiting everything possible that can be exploited, using perfect, reversible, and lossless process steps.  $\mathcal{P}$  does not have to be unique; many processes could do this in theory. This concept is general and applies to all forms of exergy.

## 1.2 Thermo-mechanical exergy

In this work, we consider only the specific thermo-mechanical exergy ( $e^{tm}$ ) of substances (noting that photons are not considered substances for this analysis), which includes the contribution of temperature, pressure, and phase of a substance. It does not consider exergy arising from chemical composition of that substance, atomic energy, motion, height, electrostatic charge, stress, strain, or other forms of energy.

As shown in [2], the following equation is one solution to the optimization problem of eq. (1) for the case of the thermo-mechanical exergy of a substance at state 1 ( $e_1^{tm}$ ) relative to environmental reference state 0:

$$e_1^{tm} = |h_1 - h_*| + (h_* - h_0) - T_0(s_1 - s_0) \quad (2)$$

where  $h$  and  $s$  are the specific enthalpy and specific entropy (on a per mass or per mole basis) of a substance at state 0, \*, or 1 [1, 2]. State 0 describes the surrounding environment, typically chosen by the person doing the analysis to be at about 1 atm and between 15–25°C for systems on the surface of the earth. State \* is an intermediate state at pressure  $p_* = p_0$  and  $s_* = s_1$ , such that the work required/produced from isentropic pressure change between state 1 and \* is  $|h_1 - h_*|$ , and the work required/produced from an isobaric reversible heat pump/reversible heat engine between state 0 and \* is  $(h_* - h_0) - T_0(s_1 - s_0)$ . Eq. (2) uses the so-called “flow exergy” form, which is exergy for systems in which mass can be exchanged with the environment. As discussed previously [3], eq. (1) differs from the classical thermo-mechanical exergy expression below ambient pressure. The classical form of thermo-mechanical exergy has historically been expressed as [4]:

$$e_1^{tm} = (h_1 - h_0) - T_0(s_1 - s_0) \quad (3)$$

Note that eq. (2) reduces to eq. (3) for  $p_1 \geq p_0$ . However, our previous work [2] showed that eq. (3) is inaccurate for  $p_1 < p_0$  because it does not solve the optimization problem in eq. (1). In short, it arises from a theoretical model that is not general enough and constrains the problem to a suboptimal solution. As a result, if eq. (3) is

used for  $p_1 < p_0$  it can result in too-low or even negative values of  $e_1^{tm}$ , which is not meaningful.

In this work, we address a much-neglected aspect of thermo-mechanical exergy of a substance, namely, its behaviour as the temperature of a substance approaches absolute zero. We discuss considerations relating to the Third Law of Thermodynamics and the Unattainability Principle regarding absolute zero. We provide a proof that  $e_1^{tm}$  is finite as  $T_1 \rightarrow 0\text{ K}$  that is general for all pressures. This work also contains enthalpy-pressure-exergy and temperature-exergy diagrams of neon, with the latter being first quantified values approaching absolute zero in the literature to the best of the author’s knowledge.

## 2. $e_1^{tm}$ AS TEMPERATURE APPROACHES ABSOLUTE ZERO

### 2.1 The Third Law of Thermodynamics

The Third Law of Thermodynamics deals with the thermodynamic properties of matter as it approaches absolute zero. A modern definition is “that any process cannot reach absolute zero temperature in a finite number of steps and within a finite time” [5]. This leads to several useful corollaries.

#### Corollary 1

The absolute entropy of a substance in thermal equilibrium, which is the total entropy of a system, tends toward zero, from above, as temperature approaches 0 K.

This is a rephrasing of the description in [6]. The exception to this rule is if the substance has a degenerate ground state at absolute zero; but nevertheless, the absolute entropy ( $s^{abs}$ ) would still be finite and zero in nearly all cases, and if not, very small [7]. This leads to a second corollary, directly quoted here:

#### Corollary 2

The heat capacity of a substance tends toward zero, from above, as temperature approaches 0 K. [7]

A simple explanation is that a substance at absolute zero cannot have any heat capacity. If it were to absorb even the slightest bit of energy, such as from a single photon collision, the temperature of the substance must increase. Otherwise, by definition, it would not have been at absolute zero in the first place. However, since it is incredibly difficult to experimentally measure heat capacity near absolute zero, models and extrapolations are needed [6].

### 2.2 $e_1^{tm}$ of a substance is finite as $T \rightarrow 0\text{ K}$

In this section, it should be noted that this proof

applies only to the thermo-mechanical exergy of substances, not matter generally, nor exergy generally.

### 2.2.1 Enthalpy

We start with eq. (3) because it is easier to follow, noting that this equation is valid for  $p_1 \geq p_0$ . We first note that  $h_0$ ,  $T_0$ , and  $s_0$  are fixed environmental reference conditions and of course finite. Therefore, to show that  $e_1^{\text{tm}}$  is finite, we need only to show that  $h_1$  and  $s_1$  are finite as  $T_1 \rightarrow 0 K$ .  $h_1$  for any pure substance is:

$$h_1 = h_{ref} + \int_{T_{ref}}^{T_1} C_p(T) dT \quad (4)$$

Where  $T_{ref}$  is some arbitrarily chosen temperature at the thermodynamic reference state such that  $h_{ref}$  is defined (usually to be 0) at  $T = T_{ref}$ . In this formulation, the heat capacity equation  $C_p(T)$  applies to the same pressure as state 1. It is important to note that the thermodynamic reference state here need not be the same as the environmental reference state. For convenience, write the equations such that the chosen reference state is at the same phase as at absolute zero (such that we do not have to consider latent heats as  $T \rightarrow 0 K$ ).

In the limit as  $T_1 \rightarrow 0 K$ , eq. (4) becomes:

$$h_{0K} = h_{ref} + \int_{T_{ref}}^{0K} C_p(T) dT \quad (5)$$

Note the integral term evaluates to a negative number for any thermodynamic reference temperature  $T_{ref} > 0 K$ .  $C_p(T)$  is finite at any  $T$  and approaches zero in the limit as  $T \rightarrow 0 K$  by Corollary 2. As a result, the integral term must be finite as well, and so must be the enthalpy.

### 2.2.2 Entropy

Corollary 1 does not imply that  $s_1 \rightarrow 0$  because  $s_1$  is not the absolute entropy, it is the thermodynamic entropy relative to the thermodynamic reference state. However, the absolute entropy  $s^{\text{abs}}$ , which is the entropy calculated on an absolute scale, and the specific entropy  $s$ , which is the entropy calculated relative to a reference condition, are easily related by a constant:

$$s_1^{\text{abs}} - s_{ref}^{\text{abs}} = s_1 - s_{ref} \quad (6)$$

The thermodynamic reference entropy  $s_{ref}$  is defined to be zero at some  $T_{ref} \geq 0 K$ , with a corresponding absolute reference entropy equal to  $s_{ref}^{\text{abs}} = s_{0K}^{\text{abs}} + \int_{0K}^{T_{ref}} \frac{C_p(T)}{T} dT$ .  $s_{0K}^{\text{abs}}$  is usually zero (but at least non-negative) and finite (corollary 1) and  $\int_{0K}^{T_{ref}} \frac{C_p(T)}{T} dT$  is non-negative and finite (corollary 2), so  $s_{ref}^{\text{abs}}$  is non-negative and finite. By eq. (6)  $s_1$  is finite and usually negative as  $T_1 \rightarrow 0K$  unless the reference condition chosen is absolute zero.

### 2.2.3 Exergy

Since  $h_1$  and  $s_1$  are finite as  $T_1 \rightarrow 0 K$ , and  $h_0$ ,  $T_0$ , and  $s_0$  are known and constant,  $e_1^{\text{tm}}$  is also finite as  $T_1 \rightarrow 0K$

(so far shown only for the  $p_1 \geq p_0$  case). Moreover, since exergy can never be negative [2],  $e_1^{\text{tm}}$  is always positive and finite, except at the environmental reference state, where it is zero. Modifications to this analysis to include enthalpy or entropy terms considering phase change or multiple species do not change the result that the exergy must be finite at absolute zero.

### 2.2.4 When $p_1 < p_0$

For sub-ambient pressures as  $T_1 \rightarrow 0 K$ ,  $h_1 \leq h_* \leq h_0$ . Since  $h_0$  and  $h_1$  are finite,  $h_*$  is always bounded between them. Therefore,  $e_1^{\text{tm}}$  in eq. (1) is always finite in the limit as  $T_1 \rightarrow 0 K$ , at any pressure.

### 2.2.5 Remarks

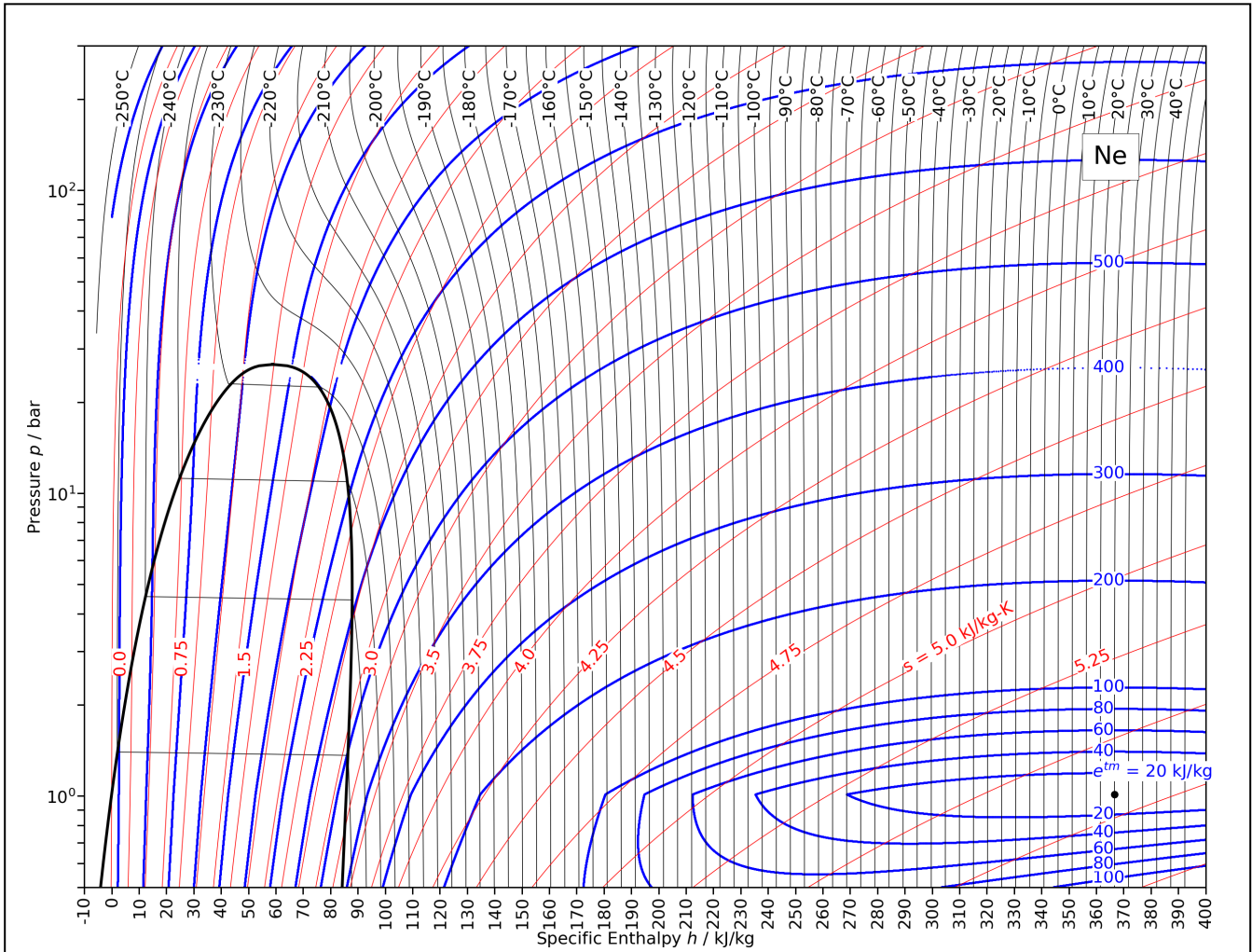
Petela [8] presented a similar proof that  $e_1^{\text{tm}}$  is finite as  $T_1 \rightarrow 0 K$ , but the proof required that the  $C_p(T)$  equation had a particular form ( $C_p(T) = mT^3$ ). It also used eq. (3) and so did not apply to pressures below  $p_0$ . In fact, the author used weak language concerning eq. 3, declaring that  $e_1^{\text{tm}}$  “**might** be expressed as [eq. (3)]” (emphasis ours) [8]. This kind of language is strangely common in the exergy literature when referring to eq. (3), and perhaps indicates a general uneasiness when using it. Therefore the above proof is stronger and applies to all pressures.

It is also interesting to note also that for a pure substance,  $\left(\frac{\partial s}{\partial p}\right)_h < 0$  and  $\left(\frac{\partial s}{\partial p}\right)_h$  tends toward zero from above as  $T_1 \rightarrow 0 K$ . This is visually apparent as nearly vertical iso-entropy lines on a pressure-enthalpy diagram below atmospheric pressure and at low temperature, as shown repeatedly in [3], and also evident on the example for neon in Figure 2, discussed in the next section. As a result,  $h_* \rightarrow h_1$  from above as  $T_1 \rightarrow 0K$ . So, eq. (2) tends toward to eq. (3) as  $T_1 \rightarrow 0 K$  at low pressures.

## 3. EXAMPLE WITH NEON

Neon is a noble gas used as a cryogenic refrigerant, especially in refrigerant blends, at temperatures potentially down to its normal boiling point near 27 K ( $-300^\circ\text{C}$ ) [9, 10]. It also forms diverse crystalline structures below its fusion point (near 24.5 K). As temperatures approach absolute zero, many different crystalline configurations have been observed or predicted using quantum mechanics-based modelling approaches. For example, as solid neon is cooled, it can in theory undergo many transitions between crystalline configurations, including core melting. This can cause the heat capacity curves to be non-monotonic in the solid region, but they still tend toward zero as temperature approaches absolute zero [11].

A pressure-enthalpy-exergy diagram of fluid-phase neon is shown in Figure 2. The enthalpy, entropy, phase, and temperature lines were computed using CoolProp [12]. The iso-exergy lines are computed using eq. (2) with



**Figure 2:** Pressure-enthalpy-exergy diagram of neon using eq. (2) for cold applications. This diagram applies to fluid phases only within model limits. Closed circle (•): Environmental reference state at  $T_0 = 25^\circ\text{C}$ ,  $p_0 = 1.01325$  bar. Blue lines (—):  $e^{tm}$  in kJ/kg. Red lines (—): specific entropy in kJ/kg-K. Black lines (—): temperature in  $^\circ\text{C}$ .

an environmental reference state of  $(T_0, p_0) = (25^\circ\text{C}, 1.01325 \text{ bar})$ .  $e_1^{tm}$  increases as neon gets colder or as the pressure deviates (above or below) the reference pressure. Due to the limits of the model used [12], the figure cannot be extended below  $-250^\circ\text{C}$ .

### 3.1 Heat Capacity Curves for Crystalline Neon

To produce a meaningful thermo-mechanical exergy curve for neon down to absolute zero, we used experimentally determined [13] heat capacity values for  $^{20}\text{Ne}$  and  $^{22}\text{Ne}$  from 3 to 23 K, which are only slightly different. Recent statistical moment method computations explain these curves well, showing that heat capacity should smoothly approach zero as temperature approaches zero [14] (see also Corollary 2).

Therefore, we used the average experimental heat capacity of the  $^{20}\text{Ne}$  and  $^{22}\text{Ne}$  isotopes to represent the solid phase neon more generally for the purposes of diagram construction. We then constructed a simple model

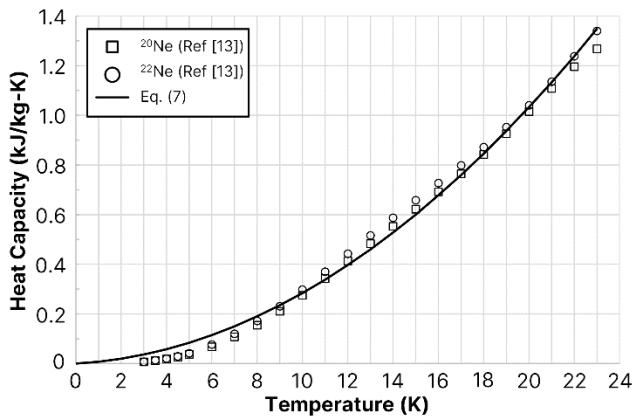
valid between  $0 < T < 23 \text{ K}$ :

$$C_{p,Ne}(T) = 2.3246 T^2 + 5.1511 T \quad (7)$$

where the heat capacity of bulk crystalline neon  $C_{p,Ne}$  is in J/kg-K. We note that it is an extrapolation to use this curve below 3 K but it evaluates to zero at 0 K and follows the quadratic-like trend in that region predicted by the statistical moment method models [14]. The resulting curve is shown in Figure 3.

We chose to use the as-reported experimentally determined values of  $h$  [13] as a function of temperature in the 3 K to 23 K region because of the known effect of core melting and various transitions on the heat capacity function [11] which are not reflected in eq. (7). However, eq. (7) was used to estimate the entropy from 0 K to 23 K analytically. The enthalpy of fusion at 1.01325 bar was taken to be 16.259 kJ/kg [15] which when using a melting point of 24.56 K for that pressure gives an entropy of fusion of 0.662 kJ/kg-K. All entropy and enthalpy

calculations above the melting point were performed using an equation of state for neon as computed by [12].



**Figure 3.** The heat capacity model for solid neon used in computing thermo-mechanical exergy at very low temperatures. Black lines (—): model. Circle: (○) Experimental data for <sup>22</sup>Ne. Square: (□) Experimental data for <sup>20</sup>Ne. Data from [13].

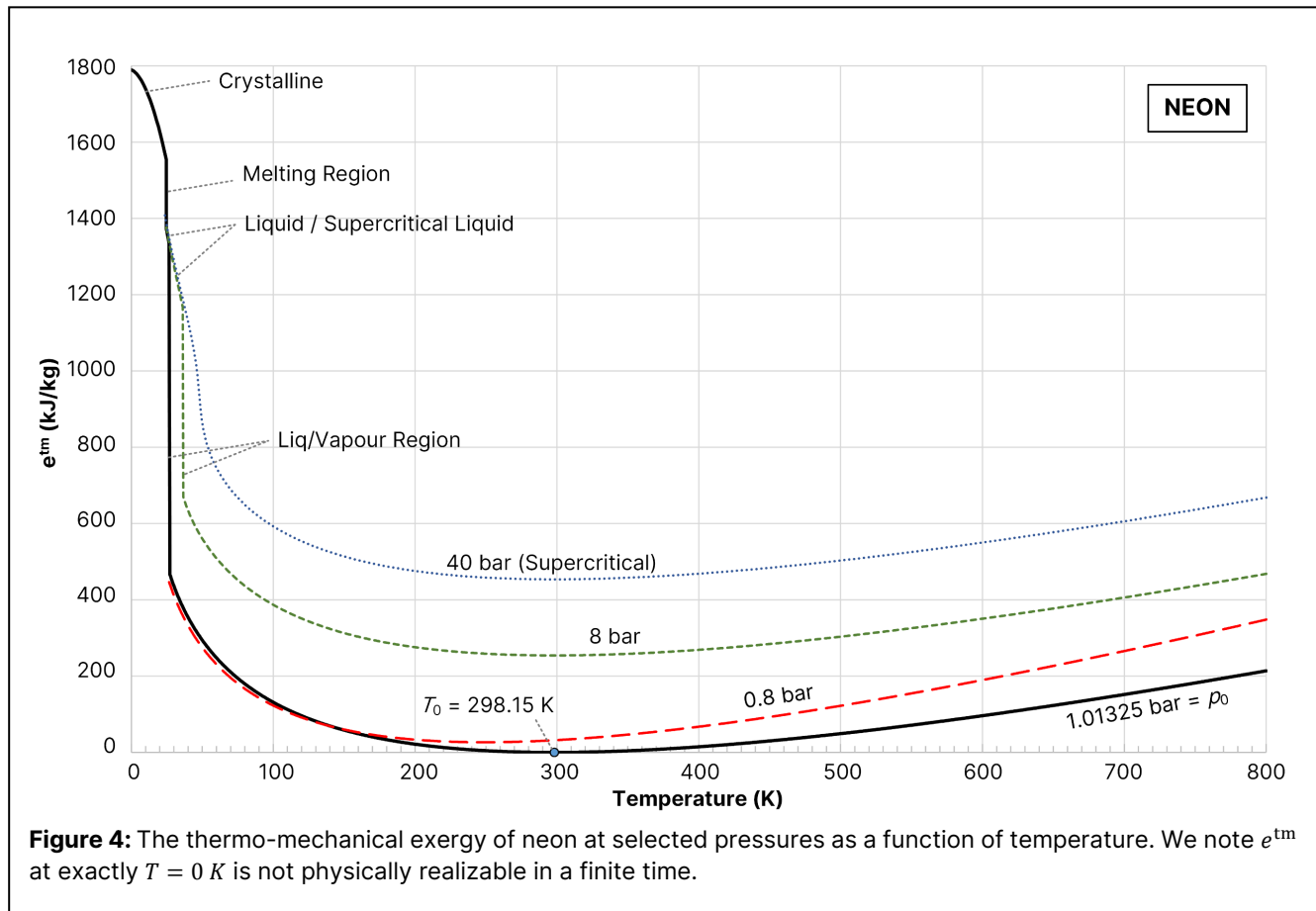
The final temperature-exergy diagram, computed using eq. (2), is shown in Figure 4. Four lines are shown for four selected pressures: one at the reference

pressure  $p_0$  (1.01325 bar), one below the reference pressure (0.8 bar), one above it (8 bar) but below the critical point of neon, and one above (40 bar) the critical point. Thermo-mechanical exergy was only computed below the melting point for the 1.01325 bar pressure case since heat of fusion data were only available at that pressure, even though it is expected to be very similar at other pressures. Thermo-mechanical exergy could only be calculated for the gas phase for the 0.8 bar state due to limitations in the equation of state model.

## 4. DISCUSSION

### 4.1 Qualitative Assessment

As temperature moves down from  $T_0$ , decreasing toward the melting point, all four pressure lines curve smoothly upward, forming a convex shape. For the subcritical pressure lines, this ends upon reaching the two-phase region, where it jumps sharply upward, reflecting the heat of condensation. The supercritical pressure case does not experience a jump; instead, the slope of the exergy curve transitions to something concave, with a less steep slope, closely matching the liquid curves for the 1.01325 bar and 8 bar cases in both quantity and slope. The liquid phase region is small for the atmospheric



**Figure 4:** The thermo-mechanical exergy of neon at selected pressures as a function of temperature. We note  $e^{tm}$  at exactly  $T = 0 K$  is not physically realizable in a finite time.



pressure case (only about 2.5 K). Essentially, the lines for the compressible fluids all collapse onto each other as they transition into the crystalline phases. As noted previously, this is because  $\lim_{T \rightarrow 0K} \frac{\partial s}{\partial p} \rightarrow 0$ .

In the crystalline region, exergy continues to increase as  $T \rightarrow 0K$  but with decreasing slope as the heat capacity is likewise approaching zero. The final estimate for thermo-mechanical exergy approaching absolute zero for neon at atmospheric pressure relative to 25°C and 1.01325 bar reference conditions is 1,790 kJ/kg. It should be noted that it does not make sense to speak of  $e_1^{\text{tm}}$  at 0 K exactly because it is unreachable in practice.

Finally, there is an open argument as to whether the “flow” form of thermo-mechanical exergy (eq. (2)) or the “closed” form of thermo-mechanical exergy is the more useful choice to describe the exergy of a substance for chemical process systems engineers for use as an analysis tool. The closed form is:

$$e_1^{\text{tm}} = (u_1 - u_0) + p_0(v_1 - v_0) - T_0(s_1 - s_0) \quad (8)$$

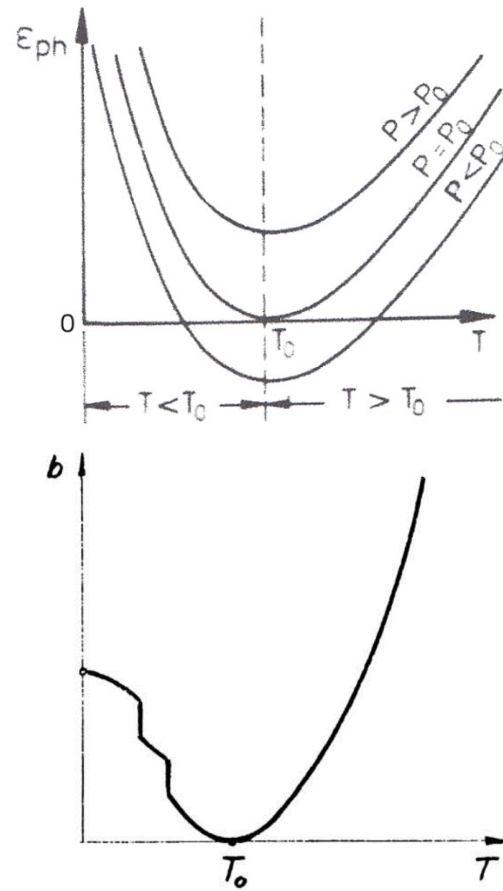
where  $u$  and  $v$  are the specific internal energy and specific volume, respectively. The closed form and flow forms differ by a term  $v_1(p_1 - p_0)$  sometimes called the “flow work” [1], which is work associated with the substance flowing into the environment and/or the boundary of the system expanding or contracting. However, for  $p_1 = p_0$ , this term is zero, making the flow and closed forms identical. Moreover, because there is very little influence of pressure on exergy in the crystalline state, the  $e_1^{\text{tm}}$  values calculated by eqs. (2), (3), and (8) converge to the same trajectory as  $T_1 \rightarrow 0K$ .

## 4.2 Comparison to Literature Predictions

Although Figure 4 is the first such quantitative diagram in the open literature that we can find, it is interesting to compare against two previous hypothetical sketches of what such a diagram would look like. Figure 5 shows two previously published curves from two previous works, for an arbitrary substance.

In Figure 5 (top),  $e_1^{\text{tm}}$  in the vacuum pressure ( $p < p_0$ ) case is sometimes negative, in contrast to Figure 4 where  $e_1^{\text{tm}} > 0$  always. This is because that work [4] assumes the form of eq. (3) to compute  $e_1^{\text{tm}}$ , which as shown previously cannot be correct precisely because it can result in negative values [2]. It is also interesting to note that the thermo-mechanical exergy of the vacuum pressure case in Figure 5 (top) has the same qualitative shape as the high-pressure case, whereas in Figure 4, the slope of the line in the vacuum pressure case increases more quickly as temperature increases, while the  $p \geq p_0$  cases have very similar slopes at all temperatures as  $T$  increases. As discussed in [2], the behaviour shown in Figure 4 is because the mechanics of mass exchange with the environment for systems above atmospheric pressure are different than those for systems below it. Also,

Figure 5 (top) does not consider phase change, nor does it account for the switch to a concave shape as  $T \rightarrow 0K$ .



**Figure 5:** Two hypothesized sketches of  $e_1^{\text{tm}}$  as a function of  $T$  and  $p$  in previous works. Top: Reprinted from Kotas [4] with digital enhancements. Bottom: Reprinted from Petela [8] with digital enhancements, noting the open circle at  $T = 0K$ . Note that  $\varepsilon_{ph}$  and  $b$  in the sketches are equivalent to our notation  $e_1^{\text{tm}}$ .

Figure 5 (bottom) [8] shows a more accurate prediction below  $T_0$ . However, in Figure 5 (bottom)  $e_1^{\text{tm}}$  increases very quickly as  $T$  increases above  $T_0$ , rapidly overtaking the maximum cold  $e_1^{\text{tm}}$  at  $T \rightarrow 0K$ . In the final result of Figure 4,  $e_1^{\text{tm}}$  for neon grows very quickly as  $T \rightarrow 0K$  and far more slowly as  $T$  increases above  $T_0$ . This prediction in Figure 5 is curious. It arises because that work assumes that heat capacity near 0 K is cubic ( $C_p(T) = kT^3$  for some constant  $k$ ), which is quite reasonable—in fact we considered a cubic version of eq. (7) and achieved almost identical results. However, in making the drawing, it assumes that  $C_p(T) = kT^3$  for all  $T$ , with no balancing factors of lower order terms, which is not reasonable. Thus in the drawing of Figure 5 (bottom),  $e_1^{\text{tm}}(T)$  is quartic ( $\sim T^4$ ) and this term dominates. This was thought to be correct because the exergy of heat radiation also grows  $\sim T^4$  above  $T_0$  and follows a very similar shape to

Figure 5 (bottom) except without the jumps due to phase change, and so it was thought to be justified by analogy. However, the real example of Figure 4 shows that this is not the case.

### 4.3 $e^{\text{tm}}$ in cryogenic applications and relationship to the Unattainability Principle

Figure 4 is useful because it shows a clear upper bound on thermo-mechanical exergy of neon as temperature approaches absolute zero. In the context of the definition of exergy, this is the thermodynamically limited maximum work that could be produced by very cold substances **from thermo-mechanical processes**. However, there are some key points for discussion.

It is well known that the processes which produce the maximum theoretical work in eq. (1) must be reversible [1] and in fact, some theorists have gone so far as to require reversibility in the very definition of exergy itself [16]. From the optimization perspective, this means the optimization problem formulation for exergy can be posed in the reverse [2]:

$$e = \min_Q w \quad (9)$$

where  $Q$  is a process that brings the system **from** equilibrium with its environment **to** the state of interest within the set of all such processes  $\mathbb{Q}$  such that  $Q \in \mathbb{Q}$ , and  $w$  is work that is required by  $Q$ . By computing a finite value for  $e^{\text{tm}}$  as a substance approaches absolute zero, this implies that this is the **minimum** amount of work that must be consumed in order to bring a substance to absolute zero, which would be finite.

However, consider the Principle of Unattainability, which Kieu [17] words as “cooling any system to absolute zero temperature in a finite number of steps and within a finite time is physically impossible by any procedure, no matter how idealized the procedure.” Razek [18] states that “it is impossible to cool, using finite resources, any finite quantum system below a certain minimal temperature  $T_{\text{min}}$ .” This is a stronger claim because it means that not only is absolute zero unattainable, but for quantum systems at least, there are temperatures above 0 K that are also unattainable. Shargut and Brodyanskii [19] reject the application of eq. (3) “in the region close to” absolute zero as being meaningless (and therefore by extension also rejecting the more general form eq. (2) or the closed form eq. (8) as noted in section 4.1) precisely because it yields something not infinite. They claim this violates the Third Law of Thermodynamics (which may not be strictly equal to the Unattainability Principle [17], discussed later).

How then can we interpret these apparent contradictions? How useful is the computed value of  $e_1^{\text{tm}}$  in cryogenic applications? First, it should be noted that the Third Law was used both in the proof of the finiteness of  $e_1^{\text{tm}}$  approaching absolute zero (namely that the absolute entropy of a substance is zero at absolute zero) and in its

computation (that the heat capacity of a substance must be zero at absolute zero), not only in this work but also through alternative approaches in [8]. So, the Third Law itself cannot imply infinite exergy.

Second, when considering the definition of exergy and its subsequent optimization formulations, there is no constraint that the theoretical process that produces the maximum work requires **finite** time. In fact, it is theoretically possible to cool a substance to absolute zero using finite energy in an infinite amount of time [18, 20]. In optimization terms, the process  $Q$  which solves eq. (9) takes infinite time but results in a finite  $w$ . In fact, some recent work provides an interesting proof that the Third Law and the Principle of Unattainability are not strictly identical, and therefore the “possibility, logically and physically speaking, of attainability of absolute zero...by non-adiabatic means” remains [17]. So perhaps the solution  $Q$  is non-adiabatic. However, we note that there is ongoing disagreement in the equivalency of the Third Law and the Principle of Unattainability.

Finally, there is a more existential issue, namely that definition 1 applies to exergy generally, not thermo-mechanical exergy specifically. This distinction is important. Most process systems engineers find it very practical to decompose exergy into various types that can be summed together to get a final exergy value, using a variety of different taxonomies to do this [3]. For example, the exergy owing to chemical potential (molecular bonds and concentration differences), height above a reference point (potential energy due to gravity), kinetic energy of motion, temperature, pressure, and phase, are conceptually decomposable in many real applications. However, it has also been shown that these different kinds of exergies are strongly interlinked, and such decompositions are not applicable in many cases. For example, the exergy associated with phase (e.g. saturated vapour vs. saturated liquid) is not isolatable because it affects both thermo-mechanical and chemical exergy [3]. Marmolejo-Correa [21, 22] also showed that even the decomposition into temperature and pressure-based contributions was not meaningful because the solutions to eq. (1) are not unique, so it is not possible to take any given thermo-mechanical exergy and uniquely describe it as the sum of a temperature and a pressure component. Therefore, the quantification of exergy at absolute zero in general outside of thermo-mechanical contributions remains an open problem. Nevertheless, the practicality of computing  $e_1^{\text{tm}}$  down to absolute zero remains because it is a good and meaningful descriptor, useful in the analysis of real cryogenic processes that considers the First, Second, **and** Third Laws of Thermodynamics.

## 5. CONCLUSIONS

Using neon as an illustrative example, we



demonstrated how the thermo-mechanical exergy of substances changes due to the impacts of pressure and phase across subcritical gas, subcritical liquid, supercritical fluid, and crystalline phases as temperature decreases toward absolute zero. Although pressure impacts exergy in the gas phase, it has almost no impact on substances in compressed phases. For example, liquid and supercritical liquid phases at the same temperature have virtually indistinguishable thermo-mechanical exergies. Furthermore, contrary to previous predictions, calculations using real data for neon shows that  $e_1^{\text{tm}}$  grows much more slowly as temperature increases away from  $T_0$  than it does as temperature decreases away from  $T_0$ .

For practicing engineers, the general form of thermo-mechanical exergy in eq. (2) can be used even for very cold situations. The most common applications in chemical process systems engineering are in assessing the quality and performance of refrigeration systems, organic Rankine cycles, and in work recovery from stored very cold substances, such as in the regasification of liquified natural gas or liquid hydrogen.

## ACKNOWLEDGMENTS

No funding was used for this work. Helpful discussions with Professors Truls Gundersen (NTNU), Bjørn Austbø (NTNU), Nadia Shardt (NTNU), and Nidret Ibric (Univ. Tuzla) are gratefully acknowledged.

## REFERENCES

- Moran MJ, Shapiro HN, Boettner DD, Bailey MB. Fundamentals of Engineering Thermodynamics, 8<sup>th</sup> Edition. Wiley: New Jersey (2014)
- Adams TA II, Gundersen T. Thermo-mechanical exergy of a substance below environmental pressure. *Ind Eng Chem Res* in press, DOI: 10.1021/acs.iecr.4c00007 (2024)
- Deng L, Adams TA II, Gundersen T. Exergy Tables. McGraw-Hill: New York (2023)
- Kotas TJ. The Exergy Method of Thermal Plant Analysis, 3<sup>rd</sup> Edition. Exergon: London (2012)
- Masanés L, Oppenheim J. A general derivation and quantification of the third law of thermodynamics. *Nature Commun.* 8:14538 (2017)
- Steane AM. On determining absolute entropy without quantum theory or the third law of thermodynamics. *New J Physics* 18:043022 (2016)
- Cross PC, Eckstrom HC. The third law of classical thermodynamics. *J Chem Phys* 10:287-291 (1942)
- Petela R. Exergy of heat radiation. *J Heat Transfer* 86(2):187-192 (1964)
- Quack H, Haberstroh C, Seeman I, Kalus M. Neonium, a refrigerant with high potential for the temperature range between 27 and 70 K. *Phys Procedia* 67:176-

- 182 (2015)
- Lee J, Oh H, Jeong S. Investigation of neon-nitrogen mixed refrigerant Joule-Thomson cryocooler operating below 70K with precooling at 100K. *Cryogenics* 61:55-62 (2014)
- Frantsuzov PA, Meluzzi D, Mandelshtam VA. Structural transformations and melting in neon clusters: Quantum versus classical mechanics. *Phys Rev Lett* 96:113401 (2006)
- Hill IH, Wronski J, Quoilin S, Lemort V. Pure and pseudo-pure fluid thermophysical property evaluation and the open-source thermophysical property library CoolProp. *Ind Eng Chem Res* 53:2498-2508 (2014)
- Somoza E, Fenichel H. Isotope effects in the specific heat of solid neon. *Phys. Review B* 3:3434 (1971)
- Ho KH, Nguyen PT, Huynh NT, Dang TH, Vu VH. Isotopic effect in thermodynamic properties of solid neon. *Physica A: Stat Mech App* 630:12927 (2023).
- Air Liquide. "Neon". *Gas Encyclopedia*. [Accessed October 2023] <https://encyclopedia.airliquide.com/neon>
- Szargut J. International progress in second law analysis. *Energy* 5:709-718 (1980)
- Kieu TD. Principle of Unattainability of absolute zero temperature, the Third Law of Thermodynamics, and projective quantum measurements. *Phy Lett A* 383:125848 (2019)
- Razek Y. Heat machines and quantum systems: Towards the Third Law of quantum thermodynamics. Hebrew Univ. of Jerusalem (2011)
- Shargut Y, Brodyanskii VM. Exergy and unattainability of absolute zero. *Izvestiya Vysshikh Uchebnykh Zavedenij I Energeticheskikh Ob"edinenij Sng. Energetika* 6:134-137 (1973)
- Masanés L, Oppenheim J. A general derivation and quantification of the third law of thermodynamics. *Nat Commun* 8:14538 (2017)
- Marmolejo-Correa D, Gundersen T. A new efficiency parameter for exergy analysis in low temperature processes. *Int J Exergy* 17:135-170 (2015)
- Marmolejo-Correa D. Analysis and Design of Low Temperature Processes with Focus on LNG. Doctoral Thesis: Norwegian University of Science and Technology (2013)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Designing Process Systems for Net-Zero Emissions and Nature- and People-Positive Decisions

Bhavik R. Bakshi<sup>a,b,c\*</sup>

<sup>a</sup> School for Engineering of Matter, Transport and Energy

<sup>b</sup> School of Sustainability

<sup>c</sup> School of Complex Adaptive Systems, Arizona State University, Tempe, AZ 85281, USA

\* Corresponding Author: [bhavik.bakshi@asu.edu](mailto:bhavik.bakshi@asu.edu)

---

## ABSTRACT

Sustainability of the chemical and materials industry (CMI) requires it to achieve net-zero emissions of greenhouse gases and other resources while making decisions that have a net-positive impact on nature and society. Many corporations, nations, and universities have pledged to meet such goals but systematic models, methods, and tools to guide this transition are missing. We present a framework to meet this need. It involves developing a comprehensive, open access model of the global CMI. In addition to existing technologies, this model includes emerging alternatives for renewable energy, circularization, and carbon capture, utilization and storage. Systematic methods help identify innovation opportunities and develop roadmaps that account for long-term changes such as technology evolution and climate change. Meeting the goal of net-zero emissions requires inclusion of life cycle impacts. Nature-positive decisions need to encourage ecological protection and restoration. This is enabled by a multiscale framework for determining the absolute environmental sustainability of products and processes by accounting for the availability of ecosystem services and their carrying capacities at multiple spatial scales. People-positive decisions need to account for the benefits to society versus harm. Issues of social justice and equity also need to be included in the decisions. More work has focused on the goal of net-zero greenhouse gas emissions but the need for better models, methods and applications remains. Nature- and people-positive decisions need to consider spatial and temporal variation of ecological and social systems. Meeting these challenges presents many novel opportunities for socially-relevant process systems engineering.

---

**Keywords:** Environment, Life Cycle Analysis, Process Design, Process Synthesis, Interdisciplinary, Net-zero, Ecosystem services, Social equity

## MOTIVATION

The last few decades have witnessed a growing interest and urgency in incorporating sustainability in all aspects of human activities including process systems engineering. In process design, early efforts focused on meeting regulations expanded to waste minimization from individual processes and then the entire life cycle. From being a necessary evil, when the environment was included as a constraint in process design, environmental protection is becoming a source of competitive advantage, and is routinely included as an objective along with conventional economic goals. These efforts have resulted in methods for reducing the environmental impact

of manufacturing processes [1]. Expanding the system boundary to include the life cycle has helped in reducing the chance of unintended harm due to emissions shifting along the life cycle. However, these efforts are not enough for ensuring sustainability for the following reasons.

- Reducing emissions is necessary but not good enough. This is because for many categories, it is urgent that we reduce emissions to zero or even negative. The most important in this category is GHG emissions, but zero emissions are also needed for water, particularly in arid regions of the world, and eventually for all other emissions and

resources as well.

- Environmental sustainability requires all human activities to respect nature’s capacity, but most approaches do not include ecosystems and their capacity in system design. These approaches need to not only account for and respect nature’s limits but also encourage protection and restoration of ecosystems.
- Chemical processes have substantial positive and negative impacts on society. The positive impact includes improvement in our standard of living, employment opportunities and economic growth, while negative impacts are due to resource use, emissions, and other side effects. The net impacts on society need to be positive while preventing social inequities and injustice.

To address these shortcomings, human activities need to have net-zero emissions and resource use while resulting in nature- and people-positive decisions [2]. The goal of net-zero emissions of greenhouse gases (GHG) is currently the most active area of global effort across disciplines. It is covered by the United Nations Framework Convention on Climate Change and most corporations, university campuses, and nations have pledged to achieve net-zero GHG emissions within a few decades. Nature-positive means that decisions should result in restoration and protection of natural ecosystems and their biodiversity. This will result in the production of ecosystem goods and services, which are essential for sustaining human activities and well-being. Many nations and corporations are pledging to make nature-positive decisions so that biodiversity loss can be stopped and reversed by 2050 [3,4]. People-positive means that any negative impact of human activities should be less than the positive impacts. In addition, the impact should also be socially just, that is, the negative impact should not be a function of factors such as race or class [5]. Studies of the positive economic benefits and negative health impacts and climate change due to air pollution of sectors in the U.S. economy identify sectors with net-negative impact.

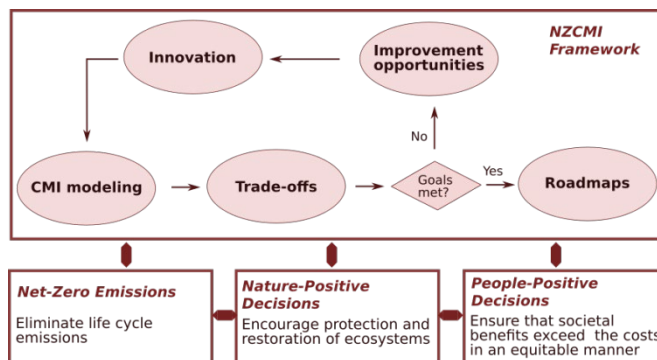
These goals are analogous to the triple bottom line [1] but take this farther by accounting for specific goals of the Paris Accord for reducing GHG emissions, respecting nature’s capacity, and doing more good than harm to people. Such pledges and efforts toward meeting them are commendable. However, achieving these goals in an economically feasible manner poses a formidable challenge since it will require transformation of industry and its supply chain in a manner and at a scale that is likely to be unprecedented. It will require innovation, and while accounting for diverse stakeholders, may result in compromise or win-win solutions. The need for such transformation is urgent since the impacts of the current

approach continues to worsen the effects of climate change, ecological degradation, and societal inequities. As conveyed in a recent emissions gap report, “the international community is falling far short of the Paris goals, with no credible pathway to 1.5°C in place. Only an urgent system-wide transformation can avoid climate disaster” [6].

Process Systems Engineering (PSE) has a unique opportunity to contribute to meeting these challenges since they require a systems view and can benefit from its methods and tools [7]. For industry, campuses, and nations to meet their pledges, they need data, models and methods to evaluate current activities, identify emerging options, and to guide the transition. This paper defines the problem and identifies the challenges in becoming net-zero, nature-positive and people-positive. We present a framework for meeting these goals, summarize current efforts and identify future needs and opportunities.

## FRAMEWORK FOR GUIDING THE TRANSITION

A general framework for guiding industry transition to a net-zero, nature-positive and people-positive future is shown in Figure 1. It relies on advanced PSE methods for modeling the current and future chemical and materials industry (CMI) pathways, evaluating trade-offs, guiding innovation and developing roadmaps to meet the specified pledges. More details about each step are provided in the rest of this section.



**Figure 1.** Framework for guiding the transition to net-zero emissions, nature- and people-positive decisions.

### Modeling the CMI

Transforming the chemicals and materials industry (CMI) to meet the goals of net-zero, nature-positive and people-positive decisions requires models of the industrial activities and their life cycles. Ideally, such a model should include details about the underlying stoichiometry, reactions, cost, and technologies of current and emerging technologies. Such models have been

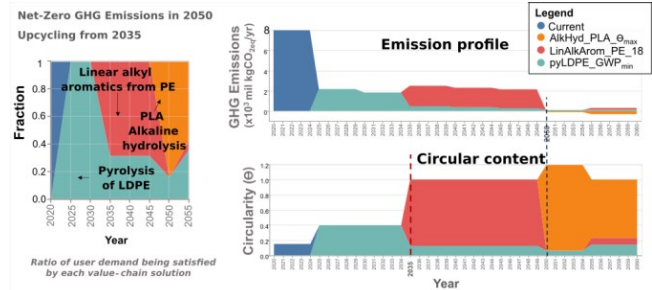
developed by companies such as S&P Global and used in academic research [8,9]. However, such models do not permit widespread academic research due to their licensing constraints and high cost. A recent model of the CMI [10] overcomes these barriers due to its open access license and no cost for non-commercial use. This model can provide the superstructure of current and emerging technologies relevant to CMI. The matrix representation enables easy linking with life cycle inventory datasets from which the trade-offs between economic, environmental, social, and circularity objectives for different pathways may be identified.

**Table 1.** Ranking innovations for sustainable circular economy of grocery bags based on improvement potential (UU\*) and readiness level (RL) [11].

Rank	Innovation	RL	UU*	Ranking criterion
1	Catalytic pyrolysis of segregated LDPE	7	0.85	0.621
2	Alkaline hydrolysis of PLA to LA using ionic liquids	2	1.83	0.611
3	Linear alkyl benzenes from sorted PE	2	1.83	0.610
4	Bio-polyethylene from sugarcane based bio-ethanol harvested in Brazil	7	0.8	0.603

### Guiding Innovation

The desired transition to net-zero, nature-positive and people-positive is unlikely with currently used alternatives. Therefore, innovations in technologies, policies, and behavior are needed. Approaches such as hotspot analysis and sensitivity optimization can be used to determine where to focus future research efforts. Results from applying such an approach to alternatives relevant to develop sustainable and circular grocery bags are shown in Table 1 [11]. Here, the improvement metric, UU\* is the distance between utopia points of the Pareto surface with current alternatives and the surface after including the new alternative. The ranking criterion is the geometric mean of the readiness level and UU\*. Finding potential innovations can rely on fundamental knowledge in chemistry databases, trade and patent literature and journals [12]. Artificial Intelligence methods for text mining and large language models can also help [13].



**Figure 2.** Roadmapping results for transforming grocery bags to net-zero GHG emissions [14].

### Roadmapping

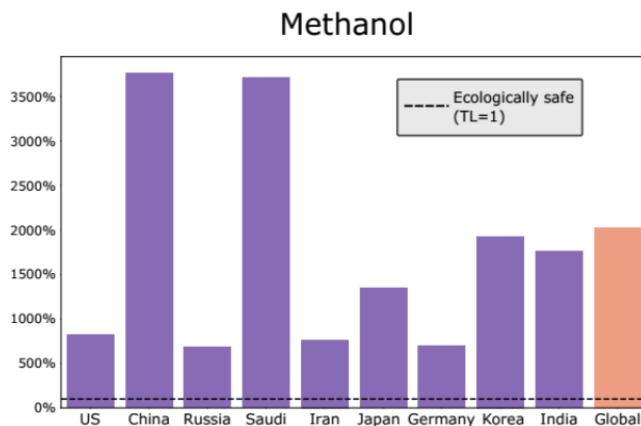
Once adequate alternatives are available for meeting a goal such as net-zero emissions, a roadmap needs to be developed to guide the transition. Given the long time horizon of the roadmap, the roadmapping approach should account for changes over this time period in aspects like technology evolution, climate change, improvement in the energy grid and changes in the economy. Such a problem has recently been formulated as a long-term planning problem which can be solved by methods such as robust optimization and multi-period stochastic programming [14]. Typical results from application of this approach to roadmapping for grocery bags with net-zero emissions is shown in Figure 2. This framework models technology evolution as continuous time Markov chains and may be linked with integrated assessment models of climate change and other large-scale changes.

### NET-ZERO EMISSIONS

Net-zero emissions are achieved when the direct and indirect emissions over the entire life cycle are zero. Some definitions consider a 90% reduction to be adequate for making claims of “net-zero”. Accounting for net-zero emissions relies on the approach of life cycle assessment and life cycle inventory databases. However, unlike the broad focus on diverse resources and impacts in LCA, most current efforts toward net-zero emissions focus only on GHG emissions. For industry to reach net-zero emissions, available alternatives belong to three broad categories: input-side technologies such as use of renewable sources of energy, circularity technologies such as mechanical and chemical recycling, and output side technologies such as carbon capture, utilization and storage (CCUS). For the hard-to-decarbonize CMI sector, circularity technologies are of critical importance since along with reducing GHG emissions, they can simultaneously reduce consumption of resources and pollution due to plastics. Reliance on CCUS may be essential for meeting goals of the Paris Accord. Reliance on renewable resources is also critical, but such technologies



are much more materials intensive than nonrenewable technologies. This may result in larger land use change and biodiversity loss due to agriculture, mining, and other land-intensive activities. Land use change can also contribute to achieving net-zero emissions by the implementation of nature-based solutions such as reforestation. For such solutions, it may be best to avoid tree plantations but rely on restoring native vegetation and biodiversity. This is likely to simultaneously contribute to all three goals of net-zero, nature-positive and people-positive solutions.



**Figure 3.** Transgression of the ecological boundary for climate change for methanol in selected nations [17]. Orange bar shows the transgression determined by direct downscaling without considering geographical variation in nature’s capacity [16].

## NATURE-POSITIVE DECISIONS

To determine whether decisions are nature-positive requires knowing the ecological impact of an activity and ensuring that ecological protection and restoration exceeds this impact. Such decisions can benefit from knowledge about the goods and services provided by nature, since that can help quantify ecological impact and restoration. A popular approach related to nature-positive decisions is based on the concept of planetary boundaries [15], which identifies the “safe operating space” for humanity. This space is based on ensuring that human activities do not exceed global ecological limits for impacts like climate change, biodiversity loss, water scarcity etc. This approach has been used to define metrics for absolute environmental sustainability, and such metrics have been calculated for various processes and products [16,17]. The result is insight into the extent to which a specific human activity transgresses planetary boundaries. Most such approaches rely on normative methods to downscale global planetary boundaries to local scales in proportion to quantities such as economic value addition, emissions, population, etc. Unfortunately, approaches based on direct downscaling of planetary

boundaries [16] need not encourage ecosystem restoration and protection because such an approach can “give away” ecosystem capacity from privately owned land to others due to their larger contribution in terms of the selected downscaling approach [17]. This approach utilizes ecological data and models to determine ecological boundaries at multiple spatial scales can provide insight into the degree of transgression of ecological boundaries and encourage.

## PEOPLE-POSITIVE DECISIONS

For a decision to be people-positive, its benefits to society should outweigh any harm. Such quantification may be done by representing the benefits and harm in monetary units. The benefits may be indicated by the economic value added by the activity while the harm can be quantified by the monetary value of lost time due to illness, disability or death. Such a study of the US economy found that for sectors such as fossil-based electricity generation, sewage treatment, and stone quarrying, the negative impact from air pollution and climate change are larger than the economic value addition from these sectors [18]. Using the framework of techno-ecological synergy for designing integrated networks of industrial and ecological systems can convert a net-negative manufacturing process into a net-positive TES system. This is by planting native and biodiverse trees in the vicinity of the manufacturing process to reduce the impact of air pollution while also obtaining other ecosystem services [19].

Another aspect of people-positive decisions involves accounting for social justice and equity. This involves at least two aspects: Industry has contributed to social inequity due to the establishment of regions where the negative impacts of industrial activity are borne disproportionately by disadvantaged communities resulting in the formation of regions such as “cancer alleys” [20]. Secondly, a minimum consumption of resources and emission of pollutants is necessary to ensure that the basic needs of a community or region are met. Thus, for decisions to be ecologically safe and socially just, they need to respect ecological limits while meeting basic needs. This operating space is referred to as the “safe and just space” (SJS) for humanity. People-positive engineering decisions require quantification of the SJS and its inclusion in process engineering methods. A recent step in this direction considers specific ecosystem services and determines the safe space or “ecological ceiling” using the TES framework described in the previous subsection. The just threshold or “social foundation” is determined as a function of the minimum food-energy-water consumption of the population in a selected region that is needed for their well-being, and the technologies used for meeting their needs. Based on this information,

the SJS has been determined for various nations in terms of the ecosystem services of carbon sequestration and water provisioning [21]. Such results are being incorporated in process and supply chain design problems and will be included in the revised version of this manuscript.

## CONCLUSIONS

Rather than just reducing emissions and environmental impact, designs for sustainability need to also meet goals of net-zero emissions particularly for greenhouse gases, respect nature's capacity while protecting and restoring ecosystems, and contribute to societal well-being and equity. Many corporations, universities and nations have pledged to meet such goals. This work describes a framework to guide industry transformation toward meeting these goals of net-zero emissions, and nature- and people-positive designs and decisions. A basic requirement for such work is the availability of a comprehensive superstructure model of the chemicals and materials industry that includes current and emerging technologies. The effect of business strategies and economic policies should also be included in such models. Developing the roadmap to meet corporate goals requires advanced optimization methods and approaches for dealing with uncertainties. Discovering innovations may be enabled by mining chemistry and engineering databases. Quantifying the metrics of net-zero emissions relies on life cycle assessment. Metrics for nature-positive decisions can use ecological models and data at multiple spatial scales, while those of people-positive decisions need to consider demographics and social equity. Meeting these goals is an urgent need that requires engineering to go beyond its traditional narrow, technocentric boundary. Companies that are successful at navigating these challenges are more likely to grow and prosper while addressing economic, environmental and social needs.

## ACKNOWLEDGEMENTS

Partial funding was provided by the National Science Foundation, The Global Kaiteki Center at Arizona State University and the Gretchen and Michael Mavrovouniotis Philanthropic Fund.

## REFERENCES

1. Bakshi, B. R. Toward Sustainable Chemical Engineering: The Role of Process Systems Engineering. Annual Review of Chemical and Biomolecular Engineering 10:265-288, (2019)
2. Schlosser, P., Rockström, J., Edwards, C., Mirazo, P., Heilemann, A., Kitzmann, N. H., Krobjinski, S. L. Accelerating transformations for a just, sustainable future: 10 'Must Haves'. Global Sustainability 6 (2023)
3. Locke, H., Rockström, J., Bakker, P., Bapna, M., Gough, M., Hilty, J., Lambertini, M., Morris, J., Polman, P., Rodriguez, C. M. et. al. A nature-positive world: the global goal for nature. Global Goal for Nature Group, (2021)
4. zu Ermgassen, S. O., Howard, M., Bennun, L., Addison, P. F., Bull, J. W., Loveridge, R., Pollard, E., Starkey, M. Are corporate biodiversity commitments consistent with delivering 'nature-positive' outcomes? A review of 'nature-positive' definitions, company progress and challenges. Journal of Cleaner Production 379:134798 (2022)
5. Obura, D. O., DeClerck, F., Verburg, P. H., Gupta, J., Abrams, J. F., Bai, X., Bunn, S., Ebi, K. L., Gifford, L., Gordon, C. et. al. Achieving a nature- and people-positive future. One Earth, 6:105-117 (2023)
6. United Nations Environment Program. Emissions Gap Report. <https://www.unep.org/resources/emissions-gap-report-2022> (2022)
7. Stephanopoulos, G., Bakshi, B. R., Basile, G. Reinventing the Chemicals/Materials Company: Transitioning to a Sustainable Circular Enterprise. In Yamashita, Y., Kano, M. (Eds.) 14th International Symposium on Process Systems Engineering, 49:67-72 (2022)
8. DeRosa, S. E., Kimura, Y., Stadtherr, M. A., McGaughey, G., McDonald-Buller, E., & Allen, D. T. Network modeling of the US petrochemical industry under raw material and Hurricane Harvey disruptions. Industrial & Engineering Chemistry Research, 58(28), 12801-12815 (2019).
9. Meys, R., Kätelhön, A., Bachmann, M., Winter, B., Zibunas, C., Suh, S., Bardow, A. Achieving net-zero greenhouse gas emission plastics by a circular carbon economy. Science 374:71-76 (2021)
10. Sen, A.; Stephanopoulos, G. & Bakshi, B. R. Mapping Anthropogenic Carbon Mobilization through Chemical Process and Manufacturing Industries. In Yamashita, Y. & Kano, M. (Eds.) 14th International Symposium on Process Systems Engineering, 49:553-558 (2022)
11. Thakker, V., Bakshi, B. R. Ranking Eco-innovations to Enable a Sustainable Circular Economy with Net-Zero Emissions. ACS Sustainable Chemistry and Engineering 11:1363-1374 (2023)
12. Coley, C. W., Green, W. H., Jensen, K. F. Machine learning in computer-aided synthesis planning. Accounts of chemical research 51:1281-1289 (2018)
13. Kumar, A., Bakshi, B., Ramteke, M., Kodamana, H. Recycle-BERT: Extracting Knowledge about Plastic

Waste Recycling by Natural Language Processing, ACS Sustainable Chemistry and Engineering, 11, 12123-12134 (2023)

14. Thakker, V., Bhavik Bakshi. Mapping the Path to a Net-Zero Chemicals Industry by Long-Term Planning with Changes in Technologies and Climate. Authorea Preprints (2023)
15. Rockström, J., Steffen, W., Noone, K., Persson, A., Chapin, F.S., 3rd, Lambin, E.F., Lenton, T.M., Scheffer, M., Folke, C., Schellnhuber, H.J., et al. A safe operating space for humanity. *Nature* 461:472–475 (2009)
16. Bachmann, M., Zibunas, C., Hartmann, J., Tulus, V., Suh, S., Guillen-Gos albez, G., Bardow, A. Towards circular plastics within planetary boundaries. *Nature Sustainability*, 6(5):599–610 (2023)
17. Xue, Y., Bakshi, B. R., Need science-based metrics to guide plastics within planetary boundaries, *Nature Sustainability*, submitted (2023)
18. Muller, N. Z., Mendelsohn, R., Nordhaus, W. Environmental Accounting for Pollution in the United States Economy. *American Economic Review* 101:1649-1675 (2011)
19. Shah, U., Bakshi, B. R. Toward Nature-Positive Manufacturing by Adapting Industrial Processes to Pollution Uptake by Vegetation. *ACS Sustainable Chemistry and Engineering*, 9:16709-16718 (2021)
20. Singer, Merrill. "Down cancer alley: the lived experience of health and environmental suffering in Louisiana's chemical corridor." *Medical Anthropology Quarterly* 25.2 (2011): 141-163.
21. Aleissa, Y. M., Bakshi, B. R. Possible but Rare: Safe and Just Satisfaction of National Human Needs in Terms of Ecosystem Services. *One Earth* 6:409-418 (2023)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# From Then to Now and Beyond: Exploring How Machine Learning Shapes Process Design Problems

Burcu Beykal<sup>a,b,\*</sup>

<sup>a</sup> Department of Chemical & Biomolecular Engineering, University of Connecticut, Storrs, CT, USA

<sup>b</sup> Center for Clean Energy Engineering, University of Connecticut, Storrs, CT, USA

\* Corresponding Author: [beykal@uconn.edu](mailto:beykal@uconn.edu)

---

## ABSTRACT

Following the discovery of the least squares method in 1805 by Legendre and later in 1809 by Gauss, surrogate modeling and machine learning have come a long way. From identifying patterns and trends in process data to predictive modeling, optimization, fault detection, reaction network discovery, and process operations, machine learning became an integral part of all aspects of process design and process systems engineering. This is enabled, at the same time necessitated, by the vast amounts of data that are readily available from processes, increased digitalization, automation, increasing computation power, and simulation software that can model complex phenomena that span over several temporal and spatial scales. Although this paper is not a comprehensive review, it gives an overview of the recent history of machine learning models that we use every day and how they shaped process design problems from the recent advances to the exploration of their prospects.

---

**Keywords:** Surrogate modeling, Artificial Intelligence, Historical view, Data-driven analysis, Process synthesis

## A BRIEF HISTORY OF MACHINE LEARNING

The roots of machine learning (ML) can be traced back to the early 19<sup>th</sup> century when the method of least squares was first discovered by Legendre in 1805 and later by Gauss in 1809 [1]. However, the main concept of computers learning from experience without explicitly being programmed has roots tracing back to more recent history, the mid-20<sup>th</sup> century.

The foundational idea of neural networks emerged in the 1940s and 1950s when researchers began exploring mathematical models inspired by the structure and functioning of the human brain. Warren McCulloch and Walter Pitts' paper, "A Logical Calculus of Ideas Immanent in Nervous Activity," was published in 1943, where they proposed a mathematical model of an artificial neuron which was the first idea of using a computational model for neural networks [2]. This foundational paper laid the groundwork for subsequent developments in neural network theory. The term "neural network" itself gained prominence in the 1950s and 1960s as researchers like Frank Rosenblatt developed the perceptron, an early

form of a neural network designed for pattern recognition tasks. While the perceptron had limitations, the idea of using computational models to simulate neural processes became a cornerstone in the evolution of artificial neural networks (ANNs) and ML. Around the same timeline, response surface methodology was introduced by Box and Wilson [3], and the term "machine learning" was coined by Arthur Samuel [4].

Throughout the following decades, various modeling approaches and algorithms, including Gaussian process (GP) regression, backpropagation algorithm, support vector machines (SVMs), and Random Forest (RF), emerged in the ML landscape. Especially, the establishment of the backpropagation algorithm was a pivotal moment in the resurgence and widespread adoption of neural works starting 1980s, enabling researchers to revisit the complex problems that were not possible to address before. This ultimately led to the application of neural networks in various scientific and engineering domains, including process design and operations, and to the development of more complex algorithms. These include reinforcement learning, deep learning, natural language processing, and generative artificial intelligence (AI)

which are emerging areas of research within process systems engineering [5,6]. Although NN models have been the primary modeling strategy in process design problems due to their ability to capture nonlinearities very accurately, we will also demonstrate that SVMs and tree-based ensemble models like RF and gradient boosted trees are also studied in depth.

## PAST APPLICATIONS OF ML IN PROCESS DESIGN

With these developments underway, it was also imperative to revisit optimal design problems from the lens of ML models and process synthesis. Regression analysis and parameter estimation for kinetic or thermodynamic models were already performed for process design, dating back to the 1960s. However, with the increasing computation power and the development of process simulation software, optimal design problems recognized the need for surrogate ML models due to: (1) the “black-box” nature of the simulation software that lacks the derivative information that is imperative for optimization; (2) the computational expense associated with sample-based derivative-free optimization techniques; and (3) the high mathematical complexity of process synthesis problems (mixed-integer nonlinear program – MINLP) that become intractable with high number of variables and constraints. Hence, earlier introduction of ML techniques in process design focused on replacing highly complex and/or noisy simulations with relatively simpler representations, especially within optimization frameworks to alleviate the mathematical complexity.

For example, Caballero and Grossmann used kriging surrogate models to replace noisy unit operations in modular flowsheet optimization [7]. Likewise, Davis and Ierapetritou used kriging surrogates for tertbutyl methacrylate production design and process synthesis [8]. Henao and Maravelias used ANN surrogate models trained using the data collected from the process simulator to replace complex unit operations (*e.g.*, distillation column, expansion valves, heaters/coolers, flash vessels, absorption columns) and reformulated these ANNs to incorporate within their superstructure optimization framework [9]. Fahmi and Cremaschi also used ANNs to substitute for thermodynamics and mixing models, as well as unit operations for process synthesis of biodiesel production [10]. One of the key challenges using ANNs was also noted in this work, where these models were “data hungry” (*i.e.*, large amounts of data were required to train accurate ANN representations). The modeling complexity of the ANNs also made it challenging to incorporate them in large-scale optimization problems without any efficient reformulation strategies. Equation 1 shows the mathematical structure of a general feed-forward NN, represented by a repeated composition of functions,

$$y = \mathbf{f}(\mathbf{x}; \boldsymbol{\theta}) = f_L \circ f_{L-1} \circ \dots \circ f_2 \circ f_1(\mathbf{x}; \boldsymbol{\theta}) \quad (1)$$

where  $y$  is the output of the network,  $\mathbf{x}$  are the inputs to the network,  $f_i$  is a layer in the neural network with transformations applied by the activation functions, and  $\boldsymbol{\theta}$  are the weights and biases for the entire network. In simpler terms, this mathematical structure generates highly nonlinear expressions (except for purely linear activation functions) that create additional complexities for optimization algorithms to handle (Equation 2).

$$y = f_L(f_{L-1}(f_{L-2}(\dots f_1(\mathbf{x})))) \quad (2)$$

Especially, within a global optimization framework, this nested functional form can be intractable as well. Motivated by this, most recent progress focused on using more simplified surrogate models for process design and synthesis problems, as well as developing novel reformulation strategies that exploit the mathematical properties of activation functions. Next, we discuss these developments and other key progress in this area.

## CURRENT PROGRESS

### Reformulation of ML Models

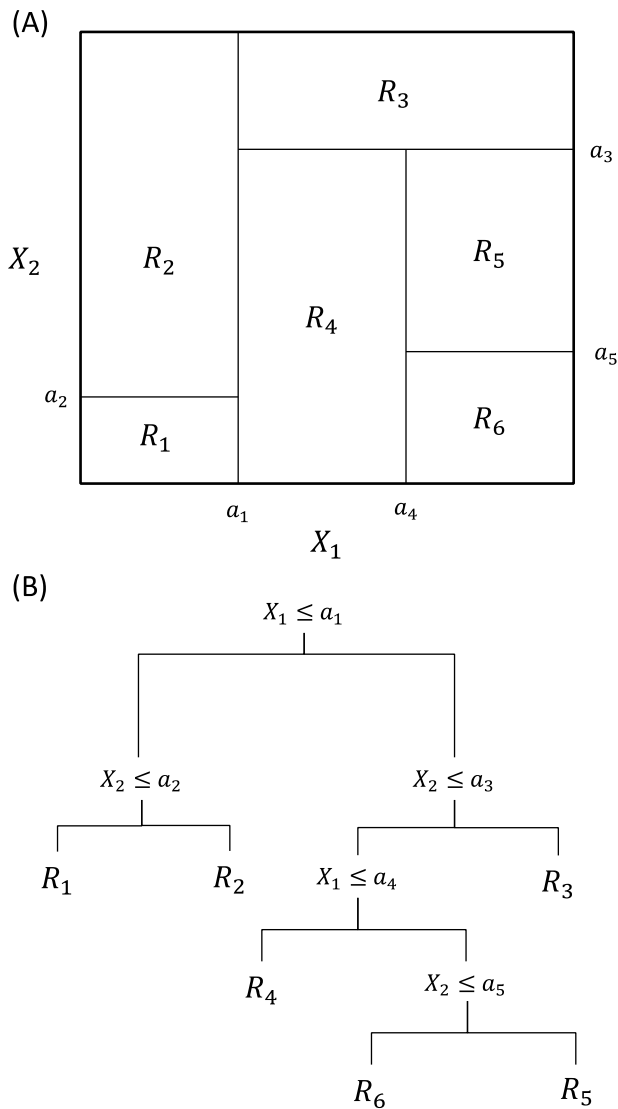
One of the most recent key breakthroughs in using ML in any optimization framework (*e.g.*, process design, synthesis, or operations) is the ability to reformulate deep NNs with rectified linear unit (ReLU) activation functions into a mixed-integer linear program (MILP) [11-13]. By recognizing ReLU activation functions as max-affine spline operators that are piecewise linear (Equation 3), ANNs can be exactly reformulated with big-M constraints to create a MILP that can be solved to global optimality with off-the-shelf solvers.

$$y = \max\{0, z\} = \begin{cases} 0 & \text{if } z < 0 \\ z & \text{otherwise} \end{cases} \quad (3)$$

This enabled MILP-reformulated ANNs to be embedded in a variety of problems, including optimizing the extractive distillation process [14], sustainable hydrogen production using sorption enhanced steam methane reforming [15], as well as for modeling flexibility index constraints in biorefinery design by superstructure optimization [16]. This technique is also extended to other activation functions that are nonlinear [17,18], and different modeling strategies such as tree-based ML models, as they also partition the modeling space with piecewise linear models (Figure 1).

Mišić [19] and Mistry *et al.* [20] encoded trained gradient boosted regression trees, which are ensemble decision tree models, to MILP models that are later embedded into optimization problems. This technique is further extended as a black-box optimization algorithm in the ENTMOOT framework [21] and implemented as an open-source software package named OMLT [22]. The applicability of the tree-based reformulation is also

demonstrated on an optimal layout design problem of an offshore windfarm [23].



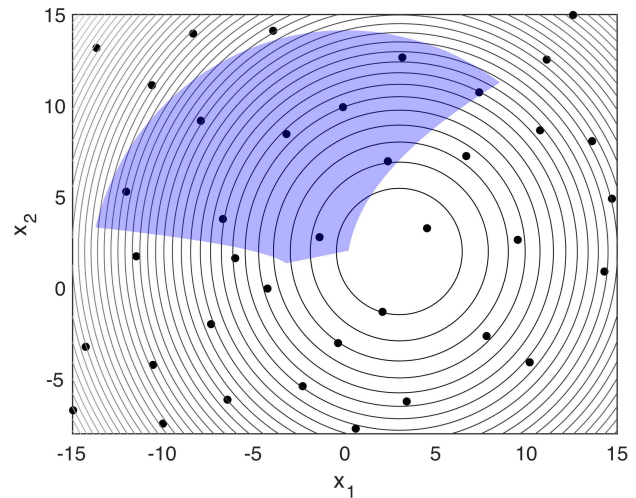
**Figure 1.** A demonstration of how decision tree models (A) partition the space with piecewise linear models; and (B) map this partitioning onto decision trees for a visual representation.

Despite reformulation strategies alleviating a portion of the nonlinearity issues in NN and tree-based ensemble models, we also observe that large-scale process synthesis problems still rely on more simplistic models. For instance, Demirhan *et al.* used linear surrogate models to model the conversion of a Haber-Bosch reactor within the renewable ammonia process synthesis problem that has 18,573 continuous, 38 binary variables, and 18,924 constraints [24]. Under such large-scale global optimization problems, reformulating MILP representations of NN or tree-based models of individual units will amplify the

number of binary variables, which will further increase the complexity of the overall optimization model. Hence, the use of ML in large-scale synthesis problems is still contingent on the overall problem complexity, even when ML models offer highly accurate predictions.

### ML Algorithms as Constraints

Nowadays, the use of ML is not limited to modeling individual unit operations or an entire flowsheet, but it can also serve as constraints to process design problems. Especially, process simulations are typically subject to black-box constraints that lack explicit analytical expressions relating the decision variables to constraint violations (*i.e.*, the constraint violations can only be obtained once the simulation run is completed). Constraint handling can be achieved in many ways, including augmented Lagrangian formulations, penalty, filter, or barrier methods [25].

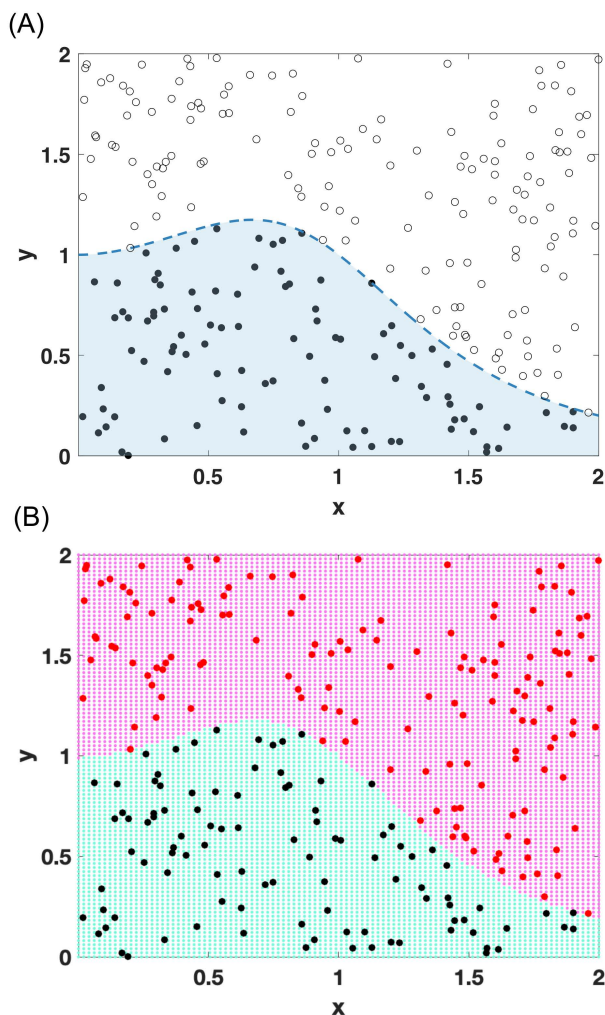


**Figure 2.** The feasible region derived by fitted surrogates using regression analysis (purple) is shown on a contour plot of the objective function. Black dots show the sampling points for the input space. The constraints are:  $x^2 + y^2 - 200 \leq 0$ ;  $x - 5y + 10 \leq 0$ ;  $25x - 2y^2 + 4y - 5 \leq 0$ ;  $-x - 2y^2 + 4y - 5 \leq 0$ .

On the other hand, ML tools can be leveraged to handle constraints individually as a regression task [26,27], where constraint violations are modeled as less than or equal to constraints with surrogate models (Figure 2), or holistically as a classification task [28-30], where a separating model between feasible and infeasible solutions are established. A conceptual demonstration of a nonlinear constraint being modeled as a classifier using SVM with Gaussian radial basis function kernel is provided in Figure 3.

This is achieved by using a dataset of simulated samples with their binary outcome (feasible/infeasible) to train one or more classification models instead of

individually modeling actual violation values as done in the regression analysis. The trained classifier can then be used with a data-driven optimizer or implemented in a superstructure model using the aforementioned reformulation strategies. In that respect, SVM classifiers are shown to effectively model the implicit constraints of numerically infeasible differential algebraic equations of a steam cracker reactor design problem [28,29]. SVM classifiers are also used for modeling feasibility constraints in the vertex formulation of modular design problems [30].



**Figure 3.** An SVM-based classifier trained to mimic the nonlinear constraint,  $y < 1/(x^3 - x^2 + 1)$ : (A) The original constraint within the bounded space; (B) The map of the feasible region captured by SVM. The predictive performance of the classifier on a blind testing set: Accuracy = 100%; Sensitivity = 100%; Specificity = 100%; F<sub>1</sub> score = 100%.

While these studies show promise for using classifiers, the offline model training is still time-consuming (*i.e.*, several thousand samples are collected from the simulator to create the model) with no efficient way to

recycle or integrate the already collected data into the decision-making process. Also, understanding the uncertainties surrounding these models as well as their misclassification rate is of utmost importance for constraint modeling, as misclassifications can lead to infeasible solutions or designs, whereas in regression-based models, such misviolations are less likely to happen.

## Large Language Models & Generative AI for Process Design

Choosing the right sequence of unit operations and connections to create a flowsheet is a fundamental practice in process design, whether it is done heuristically or through superstructure optimization. With the launch of ChatGPT, the key question becomes whether natural language processing or generative techniques can be utilized for process design and discovery. As a language model, ChatGPT is designed to understand and generate human-like text based on the input it receives. However, chemical engineering problems, such as process flowsheet generation, are based on recognizing the sequence of unit operations. This comes with the caveat of lack of suitable data to be able to train the large language models that can generate a flowsheet automatically [31]. For instance, there could be 20+ different representations for the same unit [32] or process flowsheets are most likely proprietary or unavailable to extract the information necessary for the AI model development [33].

To overcome these limitations, Vogel *et al.* developed SFILES 2.0 [33] to represent flowsheets using a graph notation, analogously similar to the text-based SMILES notation for representing chemical structures. This was primarily developed to topologically describe a flowsheet with the disadvantage of not storing any information about the sizing or the operating conditions of the units. Along the idea of ChatGPT, Vogel *et al.* also investigated the automatic completion of flowsheets using causal language modeling [34]. Their results show that the generative AI model can learn the topological patterns in flowsheet data and can automatically complete flowsheets. However, like the issues faced in ChatGPT, the generated flowsheet may not make practical sense, as the developed model does not intake contextual information about the process. Using the SFILES notation, Hirtreiter *et al.* also investigated the automatic generation of control structures for flowsheets [35]. While the predictive accuracy of the trained models was relatively high, significant limitations are also noted by the authors. Especially, concerns regarding safety indicators, understanding the process dynamics and operational objectives, and lack of information on the equipment sizing and operating conditions for the units pose major questions. These promising developments show that natural language processing and large language models can provide a “warm start” for flowsheet generation and facilitate



some of the time-consuming tasks. However, more research needs to be done in this area to be able to improve confidence in model predictions while providing a holistic view of process design beyond just the topological investigation.

## CONCLUSIONS

Artificial Intelligence and machine learning (ML) models are now an essential component of process design with efficient model integration and reformulation strategies paving the way. Constraint handling with ML models to flowsheet generation using large language models, we see new and innovative ways of how ML is used for process design problems. While the main motivation for using ML models is to alleviate the model complexities and such models have proven to be successful over the course of decades, domain knowledge and model interpretability will still play the most important role despite the promise these models hold. The ability to understand why a model makes a particular prediction and to reason if predicted results are physically sound or whether a generative model-derived process flow diagram is safe to implement becomes critically important. This judgment requires a deep fundamental understanding of the process, engineering expertise, and other relevant domain knowledge. Incorporating safety and risk measures, and combining ML models with first-principles information to create hybrid models are a few avenues that researchers are currently investigating. Nevertheless, the importance of interpretability and understanding the process relevance of the predictions will persist in this field.

## ACKNOWLEDGEMENTS

This work was supported by the National Institutes of Health [NIH P42 ES027704] and University of Connecticut. Portions of this research were conducted with the advanced computing resources provided by the University of Connecticut Storrs High Performance Computing facility. The manuscript contents are solely the responsibility of the grantee and do not necessarily represent the official views of the NIH. Further, NIH does not endorse the purchase of any commercial products or services mentioned in the publication.

## REFERENCES

1. Stigler SM. Gauss and the invention of least squares. *Ann Stat* 9(3):465-474 (1981)
2. McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys* 5:115-133 (1943)
3. Box GEP, Wilson KB. On the Experimental Attainment of Optimum Conditions. *J R Stat Soc Series B Stat Methodol* 13(1):1-45 (1951)
4. Samuel AL. Some studies in machine learning using the game of checkers. *IBM J Res Dev* 3(3):210-229 (1959)
5. Daoutidis P, Lee JH, Rangarajan S, Chiang L, Gopaluni B, Schweidtmann AM, Harjunkoski I, Mercangöz M, Mesbah A, Boukouvala F, Lima FV. Machine learning in process systems engineering: Challenges and opportunities. *Comput Chem Eng* 181:108523 (2024)
6. Schweidtmann AM, Esche E, Fischer A, Kloft M, Repke JU, Sager S, Mitsos A. Machine learning in chemical engineering: A perspective. *Chem Ing Tech* 93(12):2029-2039 (2021)
7. Caballero JA, Grossmann IE. An algorithm for the use of surrogate models in modular flowsheet optimization. *AIChE J* 54(10):2633-2650 (2008)
8. Davis E, Ierapetritou M. A kriging-based approach to MINLP containing black-box models and noise. *Ind Eng Chem Res* 47(16):6101-6125 (2008)
9. Henao CA, Maravelias CT. Surrogate-based superstructure optimization framework. *AIChE J* 57(5):1216-1232 (2011)
10. Fahmi I, Cremaschi S. Process synthesis of biodiesel production plant using artificial neural networks as the surrogate models. *Comp Chem Eng* 46:105-123 (2012)
11. Lomuscio A, Maganti L. An approach to reachability analysis for feed-forward relu neural networks. *arXiv preprint arXiv:1706.07351* (2017)
12. Fischetti M, Jo J. Deep neural networks and mixed integer linear optimization. *Constr* 23(3):296-309 (2018)
13. Grimstad B, Andersson H. ReLU networks as surrogate models in mixed-integer linear programs. *Comp Chem Eng* 131:10658 (2019)
14. Ma K, Sahinidis NV, Bindlish R, Bury SJ, Haghpanah R, Rajagopalan S. Data-driven strategies for extractive distillation unit optimization. *Comp Chem Eng* 167:107970 (2022)
15. Arora A, Zantye MS, Hasan MF. Sustainable hydrogen manufacturing via renewable-integrated intensified process for refueling stations. *Appl Energy* 311:118667 (2022).
16. Luo Y, Ierapetritou M. Multifedstock and multiproduct process design using neural network surrogate flexibility constraints. *Ind Eng Chem Res* 62(5):2067-2079 (2023)
17. Schweidtmann AM, Mitsos A. Deterministic global optimization with artificial neural networks embedded. *J Optim Theory Appl* 180(3):925-948 (2019)
18. Wilhelm ME, Wang C, Stuber MD. Convex and concave envelopes of artificial neural network

- activation functions for deterministic global optimization. *J Glob Optim* 85(3):569-594 (2023)
19. Mišić VV. Optimization of tree ensembles. *Oper Res* 68(5):1605-1624 (2020)
  20. Mistry M, Letsios D, Krennrich G, Lee RM, Misener R. Mixed-integer convex nonlinear optimization with gradient-boosted trees embedded. *INFORMS J Comput* 33(3):1103-1119 (2021)
  21. Thebelt A, Kronqvist J, Mistry M, Lee RM, Sudermann-Merx N, Misener R. ENTMOOT: a framework for optimization over ensemble tree models. *Comp Chem Eng* 151:107343 (2021)
  22. Ceccon F, Jalving J, Haddad J, Thebelt A, Tsay C, Laird CD, Misener R. OMLT: Optimization & machine learning toolkit. *J Mach Learn Res* 23(1):15829-15836 (2022)
  23. Thebelt A, Tsay C, Lee RM, Sudermann-Merx N, Walz D, Tranter T, Misener R. Multi-objective constrained optimization for energy applications via tree ensembles. *Appl Energy* 306:118061 (2022)
  24. Demirhan CD, Tso WW, Powell JB, Pistikopoulos EN. Sustainable ammonia production through process synthesis and global optimization. *AIChE J* 65(7):e16498 (2019)
  25. Beykal B, Pistikopoulos EN. Data-Driven Optimization Algorithms. In: Artificial Intelligence in Manufacturing. Ed: Soroush M, Braatz RD. Elsevier (2024)
  26. Beykal B, Boukouvala F, Floudas CA, Sorek N, Zalavadia H, Gildin E. Global optimization of grey-box computational systems using surrogate functions and application to highly constrained oil-field operations. *Comp Chem Eng* 114:99-110 (2018)
  27. Beykal B, Boukouvala F, Floudas CA, Pistikopoulos EN. Optimal design of energy systems using constrained grey-box multi-objective optimization. *Comp Chem Eng* 116:488-502 (2018)
  28. Beykal B, Onel M, Onel O, Pistikopoulos EN. A data-driven optimization algorithm for differential algebraic equations with numerical infeasibilities. *AIChE J* 66(10):e16657 (2020)
  29. Beykal B, Aghayev Z, Onel O, Onel M, Pistikopoulos EN. Data-driven stochastic optimization of numerically infeasible differential algebraic equations: an application to the steam cracking process. *Comput Aided Chem Eng* 49:1579-1584 (2022)
  30. Bhosekar A, Ierapetritou M. Modular design optimization using machine learning-based flexibility analysis. *J Process Control*, 90:18-34 (2020)
  31. Xuan J, Daniel T. The Future of Chemical Engineering in the Era of Generative AI. The Chemical Engineer. <https://www.thechemicalengineer.com/features/the-future-of-chemical-engineering-in-the-era-of-generative-ai/>
  32. Theisen MF, Flores KN, Balhorn LS, Schweidtmann AM. Digitization of chemical process flow diagrams using deep convolutional neural networks. *Digit Chem Eng* 6:100072 (2023)
  33. Vogel G, Hirtreiter E, Balhorn LS, Schweidtmann AM. SFILES 2.0: an extended text-based flowsheet representation. *Optim Eng* 1-23 (2023)
  34. Vogel G, Balhorn LS, Schweidtmann AM. Learning from flowsheets: A generative transformer model for autocompletion of flowsheets. *Comp Chem Eng* 171:108162 (2023)
  35. Hirtreiter E, Schulze Balhorn L, Schweidtmann AM. Toward automatic generation of control structures for process flow diagrams with large language models. *AIChE J* e18259 (2023)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Artificial Intelligence and Machine Learning for Sustainable Molecular-to-Systems Engineering

Alexander W. Dowling\*

University of Notre Dame, Department of Chemical and Biomolecular Engineering, Notre Dame, IN, USA

\* Corresponding Author: [adowling@nd.edu](mailto:adowling@nd.edu).

---

## ABSTRACT

Sustainability encompasses many *wicked problems* involving complex interdependencies across social, natural, and engineered systems. We argue holistic multiscale modeling and decision-support frameworks are needed to address multifaceted interdisciplinary aspects of these wicked problems. This review highlights three emerging research areas for artificial intelligence (AI) and machine learning (ML) in molecular-to-systems engineering for sustainability: (1) molecular discovery and materials design, (2) automation and self-driving laboratories, (3) process and systems-of-systems optimization. Recent advances in AI and ML are highlighted in four contemporary application areas in chemical engineering design: (1) equitable energy systems, (2) decarbonizing the power sector, (3) circular economies for critical materials, and (4) next-generation heating and cooling. These examples illustrate how AI and ML enable more sophisticated interdisciplinary multiscale models, faster optimization algorithms, more accurate uncertainty quantification, smarter and faster data collection, and incorporation of diverse stakeholders into decision-making processes, improving the robustness of engineering and policy designs while focusing on the multifaceted goals and constraints in wicked problems.

**Keywords:** Artificial Intelligence, Machine Learning, Multiscale Modelling, Interdisciplinary, Optimization

## INTRODUCTION

Creating engineered solutions to help achieve UN sustainable development goals (e.g., clean water and sanitation, affordable and clean energy, responsible consumption and production),<sup>1</sup> requires managing complex trade-offs across diverse molecular, material, device, process, and infrastructure scales<sup>2</sup>. As such, breakthroughs at a single scale are often insufficient to realize global impact. Moreover, these *wicked problems*<sup>3,4</sup> require interdisciplinary teams to manage interdependencies across social, natural, and engineered complex systems.

Using four contemporary sustainability challenges in chemical engineering, this short paper argues recent advances in artificial intelligence (AI) and machine learning (ML) methods for product and process design offer new capabilities to accelerate decision-making across molecular-to-system length and timescale. For brevity, we direct the reader to several excellent review articles and editorials for technical overviews of AI<sup>5,6</sup>, ML<sup>7-11</sup>, and data

science (DS)<sup>12-14</sup> methods used in chemical engineering. This paper focuses on advances and opportunities for AI, ML, and DS for product and materials design, which we argue is inherently multiscale and interdisciplinary. Many now consider AI as an academic discipline, and ML is a field within AI<sup>15</sup>. Similarly, DS is an interdisciplinary field focused on extracting knowledge from data, which historical roots in applied statistics<sup>16</sup>. We do not dwell on the formal distinction between AI, ML, and DS.

## MOTIVATING APPLICATIONS

We start by framing four sustainability challenges, each of which presents many opportunities for innovations in designing materials, products, and processes. However, fully realizing the benefits of technology development requires holistic systems engineering approaches, which AI and ML help facilitate.

### Equitable Energy Transitions

Social, natural, and engineered systems around the



globe must adapt to climate change<sup>17</sup>. While the poorest and most marginalized have contributed minimally to global emissions, they are impacted the most by climate change<sup>18</sup>. Thus, it is critical to consider equity and environmental justice dimensions of technologies and policies to transition the global energy economy from fossil to renewable resources. Decision-making must consider many objectives and perspectives from diverse stakeholders<sup>19</sup> to manage complex interdependencies across social, natural, and engineered systems. Best practices to broaden representation in policy processes<sup>18</sup> can be integrated with systems thinking to enable participatory research on materials and chemical products focused on the authentic needs of poor and marginalized communities.

## Decarbonizing the Power Sector

A key aspect of climate adaptation is decarbonizing electricity generation through greater adoption of carbon-neutral or renewable energy sources. One key technical challenge is that non-dispatchable renewable sources, carbon capture systems, and nuclear generators are often less flexible than fossil fuel generators. New sources of dynamic flexibility are needed to continuously balance electricity generation and production and ensure resiliency to extreme events (e.g., weather). Technoeconomic analysis of grid-connect systems, including power generators, storage systems, building, and (chemical) manufacturing processes, must consider the time-varying value of electrical energy, ancillary services, and demand response incentives<sup>20-23</sup> while considering uncertainty<sup>24</sup> and trade-offs between flexible operation and emissions<sup>25</sup>. Moreover, systems modeling can help inform the value of flexibility<sup>26,27</sup> and set performance targets for materials development, e.g., degradation in energy storage systems<sup>28</sup>. However, introducing new generators, energy storage, or integrated energy systems can distort prices in energy markets<sup>29,30</sup>, which emphasizes the need for system-of-systems modeling to assess the economics of new technologies properly. Expansion planning<sup>31,32</sup> and similar decision-support tools<sup>33</sup> are needed to establish multi-decade decarbonization pathways<sup>34,35</sup> that ultimately inform materials design and technology development.

## Circular Economies for Critical Materials

Reducing society's carbon footprint requires robust supply chains of critical minerals and materials (CMMs), including rare-earth elements (REEs)<sup>36</sup>. CMMs and REEs are essential to modern technologies, including energy storage, permanent magnetism in high-efficiency motors, and wind turbines. However, primary sources of REEs (and many CMMs) are environmentally costly to extract, separate, and refine<sup>37</sup>. Moreover, global pandemics, policies, and geopolitics can complicate CMM and REE

supply chains<sup>38,39</sup>. Thus, there is a need for new materials that enable new separations and processes to recycle CMMs and REEs from distributed sources such as used batteries and consumer electronics<sup>40</sup>. In these proceedings, Dougher et al.<sup>41</sup> elaborates on opportunities for process systems engineering to accelerate membrane separations for CMMs and REEs.

## Next Generation Heating and Cooling

As part of climate adaptation, recent international agreements mandate the phaseout of hydrofluorocarbon (HFC) refrigerants due to their high global warming potential. However, recycling existing HFCs requires new materials such as membranes<sup>42,43</sup>, sorbents<sup>44</sup>, ionic liquids<sup>45-51</sup>, and deep eutectic solvents<sup>52</sup> to separate (near)-azeotropic mixtures. Beyond vapor recompression systems, new technologies such as solid-state devices may improve performance and lower carbon footprints<sup>53</sup>. Similarly, thermoelectric materials and devices can be engineered to power distributed electronics or convert waste heat into electrical energy<sup>54</sup>. These technologies require sophisticated co-optimization of materials composition<sup>55</sup>, manufacturing processes<sup>56,57</sup>, and device designs while considering performance targets from systems models.

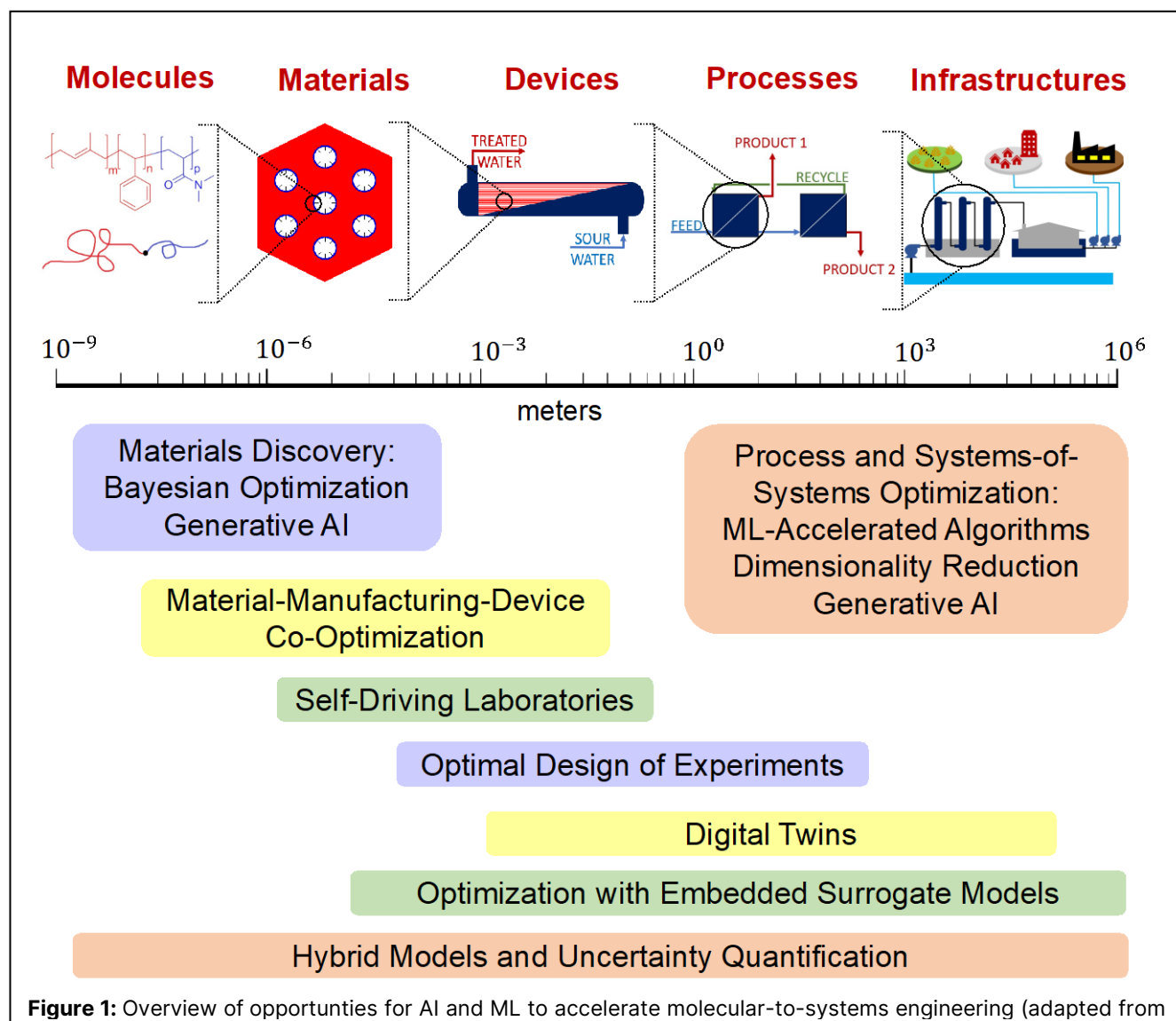
Policy-mandated phase-outs and transformations in the heating and cooling section provide another example of complex interactions between social, natural, and engineered systems<sup>58</sup>. Systems-of-system models are needed to understand the development of HFC recycling supply chains and consumer behavior, e.g., adopting new technologies to establish a circular economy, especially considering changing climates. For example, policy interventions may be needed to ensure that next-generation heating and cooling technologies remain affordable for everyone. New materials and technologies must be benchmarked using systems models that capture these complexities.

## AI & ML FOR MOLECULAR-TO-SYSTEMS

These four motivating examples emphasize the complexities of engineering new materials, chemical products, and processes to help meet sustainable development goals. This section highlights recent advances in AI and ML to address these challenges.

## Molecular Discovery and Materials Design

Computer-aided molecular design (CAMD) methods integrate predictive models to engineer novel molecules and materials, often in conjunction with processes<sup>59,60</sup>. Classic examples include co-optimizing refrigerants and refrigeration cycles, creating *designer solvents* (e.g., ionic liquids) for a wide range of separations and reactions (e.g., CO<sub>2</sub> capture, crystallization, pharmaceutical manufacturing), and engineering new materials for



**Figure 1:** Overview of opportunities for AI and ML to accelerate molecular-to-systems engineering (adapted from

energy storage (e.g., electrolytes). Advances in deep learning<sup>61</sup> combined with massive datasets enabled by new data/text mining<sup>62</sup> can provide CAMD approaches with more accurate models to predict thermophysical and mixture properties as a function of molecular/materials structure.

Molecular simulations accelerate the engineering of new molecules and materials by providing new scientific insights or predicting properties at conditions impractical to measure. AI and ML offer several opportunities to improve the speed and accuracy of molecular simulations<sup>63</sup>. For example, ML surrogate models and Bayesian Optimization (BO)<sup>64</sup> can improve the speed and accuracy of parameter estimation for force fields<sup>65,66</sup>. We proposed<sup>67</sup> and refined<sup>68</sup> an ML-assisted optimization workflow to develop new force fields for seven refrigerants. Moreover, these physically interpretable force fields accurately predict properties not considered in calibration<sup>69</sup>. An alternative approach is to replace the force field

expression with an ML model<sup>70,71</sup>. Both ML approaches are important because the lack of accurate force fields has historically hindered the use of molecular simulations to engineer some new classes of chemical products and functional materials.

Generative AI methods are emerging to engineer new molecules, materials, and synthesis pathways<sup>72,73</sup>. One popular approach uses variational autoencoders to reduce molecular descriptors into a latent space. BO<sup>64</sup> or reinforcement learning<sup>74</sup> can then optimize molecules in the latent space. ML frameworks are also emerging to co-optimize the material design, manufacturing, and end-use. For thermoelectric materials, we recently developed data science and BO methods to optimize photonic sintering<sup>57</sup> (advanced manufacturing), aerosol jet printing<sup>56</sup> (additive manufacturing), and dopant composition<sup>55</sup> (material design). These studies combined ML with expert intuition and laboratory experiments.

## Automation and Self-Driving Laboratories

Self-driving laboratories (SDLs) are revolutionizing molecular discovery and materials design by combining recent advances in robotic automation with ML<sup>75,76</sup>. Most SDLs use black-box data-driven ML models to predict experiment outcomes (e.g., material properties, synthesis yields) from input decisions. AI methods such as BO or active learning plan the next sequence or batch of experiments. SDLs have successfully optimized materials and chemical syntheses, including learning the Pareto trade-offs between competing objectives, e.g., material properties<sup>77</sup>.

Many functional materials for unit operations need to be optimized in the context of broader systems, which is likely too complex for SDLs with only data-driven ML models. For example, designing new membranes is not as simple as maximizing one property, such as selectivity or permeability, but instead, deciding how to balance material properties in the context of the broader separation system<sup>78,79</sup>. Instead, we propose combining automation and dynamic experiments for membrane characterization with science-based mathematical models<sup>80</sup>. Then model-based design of experiments methods<sup>81</sup> can be used to first distinguish between competing transport mechanisms and then optimize experiments to reduce model uncertainty. ML surrogate models can help reduce the computational burden of optimal experiment design<sup>82</sup> and model calibration<sup>83</sup>. Ultimately, this approach results in mathematical models with quantified uncertainty, i.e., digital twins<sup>84</sup>, for process and infrastructure scale optimization.

## Process and System-of-Systems Optimization

AI and ML provide new capabilities to improve the computational tractability, accuracy, and ease of implementation of integrated multiscale optimization across molecular, material, device, process, and infrastructure scales.

ML hybrid<sup>83,85</sup> or reduced-order surrogate models provide new scale-bridging approaches, especially considering the recent advances in computational optimization with embedded ML models<sup>86,87</sup>. For example, Rall et al.<sup>88</sup> used artificial neural network (ANN) surrogate models to incorporate high-fidelity ion transport membrane model performance predictions into the global superstructure optimization of a separation process. This methodology is especially powerful because it integrates rigorous models from another discipline (e.g., membrane science) with process systems engineering analyses. There is a significant opportunity for surrogate models to quantitatively establish (membrane) material property and device performance targets using process and systems models<sup>78,79</sup>.

Similarly, ML surrogate models can enhance systems-of-systems modeling. For example, Jalving et al.<sup>89</sup> trained ANN surrogate models to predict how replacing/retrofitting a generator impacts transmission network-wide outcomes in a wholesale electricity market. These ANNs were then embedded into the steady-state co-optimization of generator design and operation. This example highlights the complex interactions between individual agents (e.g., generators) in the context of a larger system (electric market). In the broader sustainability context, systems-of-systems models combine economic, societal, and environmental submodels<sup>90,91</sup>. ML surrogate models provide new opportunities to further integrate these with technology submodels (e.g., materials, processes, supply chains, infrastructure) while leveraging multi-objective computation optimization. Moreover, ML surrogate models can enable next-generation decision-support tools where diverse stakeholders interact in real time with system-of-system optimization models to facilitate negotiations. Data reduction and visualization approaches are critical to understanding and explaining trade-offs between tens or more (often correlated) objectives in sustainability problems<sup>25</sup>.

ML also provides new methods to quantify and mitigate uncertainty in decision-making. For example, Kennedy and O'Hagan hybrid models<sup>92</sup> can quantify the epistemic (model-form) uncertainty from simplifications in multiscale models<sup>93,94</sup>. We argue that Bayesian inference methods, which interpret probability as a belief, are conceptually aligned with stochastic and chance-constrained programming, which optimize over a probability distribution, i.e., the posterior from Bayesian model calibration. In contrast, frequentist statistical methods interpret probability as long-term error rates. Frequentist methods output confidence regions conceptually aligned with uncertainty sets for robust optimization. This perspective helps align the foundations of the data science and ML methods used to characterize uncertainty with the choice of optimization under uncertainty paradigm. Computational tractability remains a significant challenge to incorporating uncertainty in multiscale simulation and optimization problems. ML facilitates improved decomposition methods<sup>95</sup>, aggregation and clustering<sup>96</sup>, and branching strategies<sup>97</sup>. New generative AI methods are emerging to either accelerate or replace classical mixed integer optimization algorithms, e.g., fast sensitivity analysis<sup>98</sup>.

Finally, generative AI can accelerate the problem formulation and time to solution. For example, generative AI for flowsheet synthesis<sup>99</sup> can dramatically reduce the time to screen novel materials (e.g., new catalyst) in the context of a chemical manufacturing process. Likewise, generative AI can propose model reformulations, variable and constraint scaling, and initialization to improve model diagnostics<sup>100,101</sup> to reduce the barriers to nonlinear

optimization. These advances would make sophisticated optimization paradigms more accessible to general chemical engineering practitioners instead of highly specialized experts.

## CONCLUDING REMARKS

AI and ML improve the tractability of multiscale optimization, use data more effectively to quantify and mitigate uncertainty, and facilitate automation and faster time to solutions. These new capabilities complement recent trends in modeling software<sup>102</sup> that enable process systems engineers to guide the development of novel molecules, materials, devices, processes, and systems to address global grand challenges such as sustainable development. AI and ML provide methods that help facilitate bidirectional feedback across diverse scales and disciplines<sup>2</sup>. The described themes and opportunities are broadly relevant to sustainability challenges beyond the four motivating examples, such as recycling plastics, remediating legacy pollutants (e.g., perfluoroalkyl and polyfluoroalkyl substances (PFAS), lead), and decarbonizing chemical production (e.g., H<sub>2</sub>, ammonia, biofuels).

## ACKNOWLEDGEMENTS

The U.S. National Science Foundation supported this work via grant CBET-1941596.

## REFERENCES

1. *UN Sustainable Development Goals (SDG)*. <https://sdgs.un.org/goals> (accessed 2023-11-26).
2. Eugene, E. A.; Phillip, W. A.; Dowling, A. W. Data Science-Enabled Molecular-to-Systems Engineering for Sustainable Water Treatment. *Current Opinion in Chemical Engineering* **2019**, *26*, 122–130. <https://doi.org/10.1016/j.coche.2019.10.002>.
3. Rittel, H. W. J.; Webber, M. M. Dilemmas in a General Theory of Planning. *Policy Sci* **1973**, *4* (2), 155–169. <https://doi.org/10.1007/BF01405730>.
4. Lönngrén, J.; Van Poeck, K. Wicked Problems: A Mapping Review of the Literature. *International Journal of Sustainable Development & World Ecology* **2021**, *28* (6), 481–502. <https://doi.org/10.1080/13504509.2020.1859415>.
5. Thon, C.; Finke, B.; Kwade, A.; Schilde, C. Artificial Intelligence in Process Engineering. *Advanced Intelligent Systems* **2021**, *3* (6), 2000261. <https://doi.org/10.1002/aisy.202000261>.
6. Venkatasubramanian, V. The Promise of Artificial Intelligence in Chemical Engineering: Is It Here, Finally? *AIChE Journal* **2019**, *65* (2), 466–478. <https://doi.org/10.1002/aic.16489>.
7. Dobbelaere, M. R.; Plehiers, P. P.; Van De Vijver, R.; Stevens, C. V.; Van Geem, K. M. Machine Learning in Chemical Engineering: Strengths, Weaknesses, Opportunities, and Threats. *Engineering* **2021**, *7* (9), 1201–1211. <https://doi.org/10.1016/j.eng.2021.03.019>.
8. Haghghatlari, M.; Hachmann, J. Advances of Machine Learning in Molecular Modeling and Simulation. *Current Opinion in Chemical Engineering* **2019**, *23*, 51–57. <https://doi.org/10.1016/j.coche.2019.02.009>.
9. Lee, J. H.; Shin, J.; Realff, M. J. Machine Learning: Overview of the Recent Progresses and Implications for the Process Systems Engineering Field. *Computers & Chemical Engineering* **2018**, *114*, 111–121. <https://doi.org/10.1016/j.compchemeng.2017.10.008>.
10. Schweidtmann, A. M.; Esche, E.; Fischer, A.; Kloft, M.; Repke, J.; Sager, S.; Mitsos, A. Machine Learning in Chemical Engineering: A Perspective. *Chemie Ingenieur Technik* **2021**, *93* (12), 2029–2039. <https://doi.org/10.1002/cite.202100083>.
11. Daoutidis, P.; Lee, J. H.; Rangarajan, S.; Chiang, L.; Gopaluni, B.; Schweidtmann, A. M.; Harjunkoski, I.; Mercangöz, M.; Mesbah, A.; Boukouvala, F.; Lima, F. V.; del Rio Chanona, A.; Georgakis, C. Machine Learning in Process Systems Engineering: Challenges and Opportunities. *Computers & Chemical Engineering* **2024**, *181*, 108523. <https://doi.org/10.1016/j.compchemeng.2023.108523>.
12. Gao, H.; Zhu, L.-T.; Luo, Z.-H.; Fraga, M. A.; Hsing, I.-M. Machine Learning and Data Science in Chemical Engineering. *Ind. Eng. Chem. Res.* **2022**, *61* (24), 8357–8358. <https://doi.org/10.1021/acs.iecr.2c01788>.
13. Qin, S. J.; Chiang, L. H. Advances and Opportunities in Machine Learning for Process Data Analytics. *Computers & Chemical Engineering* **2019**, *126*, 465–473. <https://doi.org/10.1016/j.compchemeng.2019.04.003>.
14. Thebelt, A.; Wiebe, J.; Kronqvist, J.; Tsay, C.; Misener, R. Maximizing Information from Chemical Engineering Data Sets: Applications to Machine Learning. *Chemical Engineering Science* **2022**, *252*, 117469. <https://doi.org/10.1016/j.ces.2022.117469>.
15. Machine Learning. *Wikipedia*; 2024.
16. Data Science. *Wikipedia*; 2024.
17. Bierbaum, R.; Smith, J. B.; Lee, A.; Blair, M.; Carter, L.; Chapin, F. S.; Fleming, P.; Ruffo, S.; Stults, M.; McNeeley, S.; Wasley, E.; Verduzco, L. A Comprehensive Review of Climate Adaptation in

- the United States: More than before, but Less than Needed. *Mitig Adapt Strateg Glob Change* **2013**, *18* (3), 361–406. <https://doi.org/10.1007/s11027-012-9423-1>.
18. Shi, L.; Chu, E.; Anguelovski, I.; Aylett, A.; Debats, J.; Goh, K.; Schenk, T.; Seto, K. C.; Dodman, D.; Roberts, D.; Roberts, J. T.; VanDeveer, S. D. Roadmap towards Justice in Urban Climate Adaptation Research. *Nature Clim Change* **2016**, *6* (2), 131–137. <https://doi.org/10.1038/nclimate2841>.
  19. Dowling, A. W.; Ruiz-Mercado, G.; Zavala, V. M. A Framework for Multi-Stakeholder Decision-Making and Conflict Resolution. *Computers & Chemical Engineering* **2016**, *90*, 136–150. <https://doi.org/10.1016/j.compchemeng.2016.03.034>.
  20. Dowling, A. W.; Zavala, V. M. Economic Opportunities for Industrial Systems from Frequency Regulation Markets. *Computers & Chemical Engineering* **2018**, *114*, 254–264. <https://doi.org/10.1016/j.compchemeng.2017.09.018>.
  21. Dowling, A. W.; Kumar, R.; Zavala, V. M. A Multi-Scale Optimization Framework for Electricity Market Participation. *Applied Energy* **2017**, *190*, 147–164. <https://doi.org/10.1016/j.apenergy.2016.12.081>.
  22. Dowling, A. W.; Zheng, T.; Zavala, V. M. Economic Assessment of Concentrated Solar Power Technologies: A Review. *Renewable and Sustainable Energy Reviews* **2017**, *72*, 1019–1032. <https://doi.org/10.1016/j.rser.2017.01.006>.
  23. Allman, A.; Zhang, Q. Distributed Cooperative Industrial Demand Response. *Journal of Process Control* **2020**, *86*, 81–93. <https://doi.org/10.1016/j.jprocont.2019.12.011>.
  24. Elmore, C. T.; Dowling, A. W. Learning Spatiotemporal Dynamics in Wholesale Energy Markets with Dynamic Mode Decomposition. *Energy* **2021**, *232*, 121013. <https://doi.org/10.1016/j.energy.2021.121013>.
  25. Wang, H.; Allman, A. Analysis of the Correlating or Competing Nature of Cost-Driven and Emissions-Driven Demand Response. *Computers & Chemical Engineering* **2024**, *181*, 108520. <https://doi.org/10.1016/j.compchemeng.2023.108520>.
  26. Dowling, A. W.; Zheng, T.; Zavala, V. M. A Decomposition Algorithm for Simultaneous Scheduling and Control of CSP Systems. *AIChE Journal* **2018**, *64* (7), 2408–2417. <https://doi.org/10.1002/aic.16101>.
  27. Tang, X.; Baldea, M. A Grid View on the Dynamics of Processes Participating in Demand Response Programs. *Computers & Chemical Engineering* **2023**, *169*, 108070. <https://doi.org/10.1016/j.compchemeng.2022.108070>.
  28. Sorourifar, F.; Zavala, V. M.; Dowling, A. W. Integrated Multiscale Design, Market Participation, and Replacement Strategies for Battery Energy Storage Systems. *IEEE Transactions on Sustainable Energy* **2020**, *11* (1), 84–92. <https://doi.org/10.1109/TSTE.2018.2884317>.
  29. Gao, X.; Knueven, B.; Sirola, J. D.; Miller, D. C.; Dowling, A. W. Multiscale Simulation of Integrated Energy System and Electricity Market Interactions. *Applied Energy* **2022**, *316*, 119017. <https://doi.org/10.1016/j.apenergy.2022.119017>.
  30. Tsay, C.; Qvist, S. Integrating Process and Power Grid Models for Optimal Design and Demand Response Operation of Giga-Scale Green Hydrogen. *AIChE Journal* **2023**, *69* (12), e18268. <https://doi.org/10.1002/aic.18268>.
  31. Gacitua, L.; Gallegos, P.; Henriquez-Auba, R.; Lorca, Á.; Negrete-Pincetic, M.; Olivares, D.; Valenzuela, A.; Wenzel, G. A Comprehensive Review on Expansion Planning: Models and Tools for Energy Policy Analysis. *Renewable and Sustainable Energy Reviews* **2018**, *98*, 346–360. <https://doi.org/10.1016/j.rser.2018.08.043>.
  32. Koltsaklis, N. E.; Dagoumas, A. S. State-of-the-Art Generation Expansion Planning: A Review. *Applied Energy* **2018**, *230*, 563–589. <https://doi.org/10.1016/j.apenergy.2018.08.087>.
  33. Demirhan, C. D.; Tso, W. W.; Powell, J. B.; Pistikopoulos, E. N. A Multi-Scale Energy Systems Engineering Approach towards Integrated Multi-Product Network Optimization. *Applied Energy* **2021**, *281*, 116020. <https://doi.org/10.1016/j.apenergy.2020.116020>.
  34. Plazas-Niño, F. A.; Ortiz-Pimiento, N. R.; Montes-Páez, E. G. National Energy System Optimization Modelling for Decarbonization Pathways Analysis: A Systematic Literature Review. *Renewable and Sustainable Energy Reviews* **2022**, *162*, 112406. <https://doi.org/10.1016/j.rser.2022.112406>.
  35. Verástegui, F.; Lorca, Á.; Olivares, D.; Negrete-Pincetic, M. Optimization-Based Analysis of Decarbonization Pathways and Flexibility Requirements in Highly Renewable Power Systems. *Energy* **2021**, *234*, 121242. <https://doi.org/10.1016/j.energy.2021.121242>.
  36. Balaram, V. Rare Earth Elements: A Review of Applications, Occurrence, Exploration, Analysis, Recycling, and Environmental Impact. *Geoscience Frontiers* **2019**, *10* (4), 1285–1303. <https://doi.org/10.1016/j.gsf.2018.12.005>.
  37. Opore, E. O.; Struhs, E.; Mirkouei, A. A Comparative State-of-Technology Review and Future Directions



- for Rare Earth Element Separation. *Renewable and Sustainable Energy Reviews* **2021**, *143*, 110917. <https://doi.org/10.1016/j.rser.2021.110917>.
38. Hammond, D. R.; Brady, T. F. Critical Minerals for Green Energy Transition: A United States Perspective. *International Journal of Mining, Reclamation and Environment* **2022**, *36* (9), 624–641. <https://doi.org/10.1080/17480930.2022.2124788>.
  39. Srivastava, N.; Kumar, A. Minerals and Energy Interface in Energy Transition Pathways: A Systematic and Comprehensive Review. *Journal of Cleaner Production* **2022**, *376*, 134354. <https://doi.org/10.1016/j.jclepro.2022.134354>.
  40. Lair, L.; Ouimet, J. A.; Dougher, M.; Boudouris, B.; Dowling, A.; Phillip, W. Critical Mineral Separations: Opportunities for Membrane Materials and Processes to Advance Sustainable Economies and Secure Supplies. *under revision*.
  41. Dougher, M.; Lair, L.; Ouimet, J. A.; Phillip, W. A.; Dowling, A. Opportunities for Process Intensification with Membranes to Promote Circular Economy Development for Critical Minerals. In *Systems & Control Transactions*; PSE Press: Breckenridge, Colorado, USA.
  42. Harders, A. N.; Sturd, E. R.; Vallier, J. E.; Corbin, D. R.; White, W. R.; Junk, C. P.; Shiflett, M. B. Selective Separation of HFC-32 from R-410A Using Poly(Dimethylsiloxane) and a Copolymer of Perfluoro(Butenyl Vinyl Ether) and Perfluoro(2,2-Dimethyl-1,3-Dioxole). *Journal of Membrane Science* **2022**, *652*, 120467. <https://doi.org/10.1016/j.memsci.2022.120467>.
  43. Pardo, F.; Gutiérrez-Hernández, S. V.; Rodríguez-San Miguel, P.; Zarca, G.; Urtiaga, A. Polymer/Ionic Liquid Pilot Scale Membrane Prototype for the Recovery of Difluoromethane (R-32) from Refrigerant Mixtures. *Separation and Purification Technology* **2023**, *320*, 124115. <https://doi.org/10.1016/j.seppur.2023.124115>.
  44. Yancey, A. D.; Terian, S. J.; Shaw, B. J.; Bish, T. M.; Corbin, D. R.; Shiflett, M. B. A Review of Fluorocarbon Sorption on Porous Materials. *Microporous and Mesoporous Materials* **2022**, *331*, 111654. <https://doi.org/10.1016/j.micromeso.2021.111654>.
  45. Asensio-Delgado, S.; Pardo, F.; Zarca, G.; Urtiaga, A. Absorption Separation of Fluorinated Refrigerant Gases with Ionic Liquids: Equilibrium, Mass Transport, and Process Design. *Separation and Purification Technology* **2021**, *276*, 119363. <https://doi.org/10.1016/j.seppur.2021.119363>.
  46. Viar, M.; Asensio-Delgado, S.; Pardo, F.; Zarca, G.; Urtiaga, A. In the Quest for Ionic Liquid Entrainers for the Recovery of R-32 and R-125 by Extractive Distillation under Rate-Based Considerations. *Separation and Purification Technology* **2023**, *324*, 124610. <https://doi.org/10.1016/j.seppur.2023.124610>.
  47. Finberg, E. A.; May, T. L.; Shiflett, M. B. Multicomponent Refrigerant Separation Using Extractive Distillation with Ionic Liquids. *Ind. Eng. Chem. Res.* **2022**, *61* (27), 9795–9812. <https://doi.org/10.1021/acs.iecr.2c00937>.
  48. Monjur, M. S.; Iftakher, A.; Hasan, M. M. F. Separation Process Synthesis for High-GWP Refrigerant Mixtures: Extractive Distillation Using Ionic Liquids. *Ind. Eng. Chem. Res.* **2022**, *61* (12), 4390–4406. <https://doi.org/10.1021/acs.iecr.2c00136>.
  49. Baca, K. R.; Olsen, G. M.; Matamoros Valenciano, L.; Bennett, M. G.; Haggard, D. M.; Befort, B. J.; Garcadiago, A.; Dowling, A. W.; Maginn, E. J.; Shiflett, M. B. Phase Equilibria and Diffusivities of HFC-32 and HFC-125 in Ionic Liquids for the Separation of R-410A. *ACS Sustainable Chem. Eng.* **2022**, *10* (2), 816–830. <https://doi.org/10.1021/acssuschemeng.1c06252>.
  50. Garcadiago, A.; Mazumder, M.; Befort, B. J.; Dowling, A. W. Modeling and Optimization of Ionic Liquid Enabled Extractive Distillation of Ternary Azeotrope Mixtures. In *Computer Aided Chemical Engineering*; Yamashita, Y., Kano, M., Eds.; 14 International Symposium on Process Systems Engineering; Elsevier, 2022; Vol. 49, pp 307–312. <https://doi.org/10.1016/B978-0-323-85159-6.50051-8>.
  51. Morais, A. R. C.; Harders, A. N.; Baca, K. R.; Olsen, G. M.; Befort, B. J.; Dowling, A. W.; Maginn, E. J.; Shiflett, M. B. Phase Equilibria, Diffusivities, and Equation of State Modeling of HFC-32 and HFC-125 in Imidazolium-Based Ionic Liquids for the Separation of R-410A. *Ind. Eng. Chem. Res.* **2020**, *59* (40), 18222–18235. <https://doi.org/10.1021/acs.iecr.0c02820>.
  52. Codera, V.; Clijnk, D.; Pou, J. O.; Fernandez-Garcia, J.; Llovel, F.; Gonzalez-Olmos, R. Process Design for the Recovery of Waste Refrigerants Using Deep Eutectic Solvents. *Journal of Environmental Chemical Engineering* **2023**, *11* (3), 110255. <https://doi.org/10.1016/j.jece.2023.110255>.
  53. Aprea, C.; Greco, A.; Maiorino, A.; Masselli, C. Solid-State Refrigeration: A Comparison of the Energy Performances of Caloric Materials Operating in an Active Caloric Regenerator. *Energy* **2018**, *165*, 439–455. <https://doi.org/10.1016/j.energy.2018.09.114>.
  54. He, J.; Tritt, T. M. Advances in Thermoelectric Materials Research: Looking Back and Moving Forward. *Science* **2017**, *357* (6358), eaak9997.

- <https://doi.org/10.1126/science.aak9997>.
55. Shang, W.; Zeng, M.; Tanvir, A. N. M.; Wang, K.; Saeidi-Javash, M.; Dowling, A.; Luo, T.; Zhang, Y. Hybrid Data-Driven Discovery of High-Performance Silver Selenide-Based Thermoelectric Composites. *Advanced Materials* **2023**, *35* (47), 2212230. <https://doi.org/10.1002/adma.202212230>.
  56. Wang, K.; Zeng, M.; Wang, J.; Shang, W.; Zhang, Y.; Luo, T.; Dowling, A. W. When Physics-Informed Data Analytics Outperforms Black-Box Machine Learning: A Case Study in Thickness Control for Additive Manufacturing. *Digital Chemical Engineering* **2023**, *6*, 100076. <https://doi.org/10.1016/j.dche.2022.100076>.
  57. Saeidi-Javash, M.; Wang, K.; Zeng, M.; Luo, T.; Dowling, A. W.; Zhang, Y. Machine Learning-Assisted Ultrafast Flash Sintering of High-Performance and Flexible Silver-Selenide Thermoelectric Devices. *Energy Environ. Sci.* **2022**, *15* (12), 5093–5104. <https://doi.org/10.1039/D2EE01844F>.
  58. Trencher, G.; Rinscheid, A.; Rosenbloom, D.; Truong, N. The Rise of Phase-out as a Critical Decarbonisation Approach: A Systematic Review. *Environ. Res. Lett.* **2022**, *17* (12), 123002. <https://doi.org/10.1088/1748-9326/ac9fe3>.
  59. Austin, N. D.; Sahinidis, N. V.; Trahan, D. W. Computer-Aided Molecular Design: An Introduction and Review of Tools, Applications, and Solution Techniques. *Chemical Engineering Research and Design* **2016**, *116*, 2–26. <https://doi.org/10.1016/j.cherd.2016.10.014>.
  60. Chemmangattuvalappil, N. G. Development of Solvent Design Methodologies Using Computer-Aided Molecular Design Tools. *Current Opinion in Chemical Engineering* **2020**, *27*, 51–59. <https://doi.org/10.1016/j.coche.2019.11.005>.
  61. Alshehri, A. S.; Gani, R.; You, F. Deep Learning and Knowledge-Based Methods for Computer-Aided Molecular Design—toward a Unified Approach: State-of-the-Art and Future Directions. *Computers & Chemical Engineering* **2020**, *141*, 107005. <https://doi.org/10.1016/j.compchemeng.2020.107005>.
  62. Krallinger, M.; Rabal, O.; Lourenço, A.; Oyarzabal, J.; Valencia, A. Information Retrieval and Text Mining Technologies for Chemistry. *Chem. Rev.* **2017**, *117* (12), 7673–7761. <https://doi.org/10.1021/acs.chemrev.6b00851>.
  63. Noé, F.; Tkatchenko, A.; Müller, K.-R.; Clementi, C. Machine Learning for Molecular Simulation. *Annual Review of Physical Chemistry* **2020**, *71* (1), 361–390. <https://doi.org/10.1146/annurev-physchem-042018-052331>.
  64. Wang, K.; Dowling, A. W. Bayesian Optimization for Chemical Products and Functional Materials. *Current Opinion in Chemical Engineering* **2022**, *36*, 100728. <https://doi.org/10.1016/j.coche.2021.100728>.
  65. Befort, B. J.; DeFever, R. S.; Maginn, E. J.; Dowling, A. W. Machine Learning-Enabled Optimization of Force Fields for Hydrofluorocarbons. In *Computer Aided Chemical Engineering*, Yamashita, Y., Kano, M., Eds.; 14 International Symposium on Process Systems Engineering; Elsevier, 2022; Vol. 49, pp 1249–1254. <https://doi.org/10.1016/B978-0-323-85159-6.50208-6>.
  66. C. Madin, O.; R. Shirts, M. Using Physical Property Surrogate Models to Perform Accelerated Multi-Fidelity Optimization of Force Field Parameters. *Digital Discovery* **2023**, *2* (3), 828–847. <https://doi.org/10.1039/D2DD00138A>.
  67. Befort, B. J.; DeFever, R. S.; Tow, G. M.; Dowling, A. W.; Maginn, E. J. Machine Learning Directed Optimization of Classical Molecular Modeling Force Fields. *J. Chem. Inf. Model.* **2021**, *61* (9), 4400–4414. <https://doi.org/10.1021/acs.jcim.1c00448>.
  68. Wang, N.; Carlozo, M. N.; Marin-Rimoldi, E.; Befort, B. J.; Dowling, A. W.; Maginn, E. J. Machine Learning-Enabled Development of Accurate Force Fields for Refrigerants. *J. Chem. Theory Comput.* **2023**, *19* (14), 4546–4558. <https://doi.org/10.1021/acs.jctc.3c00338>.
  69. Agbodekhe, B.; Marin-Rimoldi, E.; Zhang, Y.; Dowling, A. W.; Maginn, E. J. Assessment and Ranking of Difluoromethane (R32) and Pentafluoroethane (R125) Interatomic Potentials Using Several Thermophysical and Transport Properties Across Multiple State Points. *J. Chem. Eng. Data* **2023**. <https://doi.org/10.1021/acs.jced.3c00379>.
  70. Unke, O. T.; Chmiela, S.; Sauceda, H. E.; Gastegger, M.; Poltavsky, I.; Schütt, K. T.; Tkatchenko, A.; Müller, K.-R. Machine Learning Force Fields. *Chem. Rev.* **2021**, *121* (16), 10142–10186. <https://doi.org/10.1021/acs.chemrev.0c01111>.
  71. Kocer, E.; Ko, T. W.; Behler, J. Neural Network Potentials: A Concise Overview of Methods. *Annual Review of Physical Chemistry* **2022**, *73* (Volume 73, 2022), 163–186. <https://doi.org/10.1146/annurev-physchem-082720-034254>.
  72. Bilodeau, C.; Jin, W.; Jaakkola, T.; Barzilay, R.; Jensen, K. F. Generative Models for Molecular Discovery: Recent Advances and Challenges. *WIREs Computational Molecular Science* **2022**, *12* (5), e1608. <https://doi.org/10.1002/wcms.1608>.
  73. Walters, W. P.; Murcko, M. Assessing the Impact of Generative AI on Medicinal Chemistry. *Nat Biotechnol* **2020**, *38* (2), 143–145.



- <https://doi.org/10.1038/s41587-020-0418-2>.
74. Sui, F.; Guo, R.; Zhang, Z.; Gu, G. X.; Lin, L. Deep Reinforcement Learning for Digital Materials Design. *ACS Materials Lett.* **2021**, *3* (10), 1433–1439.  
<https://doi.org/10.1021/acsmaterialslett.1c00390>.
  75. Abolhasani, M.; Kumacheva, E. The Rise of Self-Driving Labs in Chemical and Materials Sciences. *Nat. Synth* **2023**, *2* (6), 483–492.  
<https://doi.org/10.1038/s44160-022-00231-0>.
  76. Seifrid, M.; Pollice, R.; Aguilar-Granda, A.; Morgan Chan, Z.; Hotta, K.; Ser, C. T.; Vestfrid, J.; Wu, T. C.; Aspuru-Guzik, A. Autonomous Chemical Experiments: Challenges and Perspectives on Establishing a Self-Driving Lab. *Acc. Chem. Res.* **2022**, *55* (17), 2454–2466.  
<https://doi.org/10.1021/acs.accounts.2c00220>.
  77. MacLeod, B. P.; Parlane, F. G. L.; Rupnow, C. C.; Dettelbach, K. E.; Elliott, M. S.; Morrissey, T. D.; Haley, T. H.; Proskurin, O.; Rooney, M. B.; Taherimaksousi, N.; Dvorak, D. J.; Chiu, H. N.; Waizenegger, C. E. B.; Ocean, K.; Mokhtari, M.; Berlinguette, C. P. A Self-Driving Laboratory Advances the Pareto Front for Material Properties. *Nat Commun* **2022**, *13* (1), 995.  
<https://doi.org/10.1038/s41467-022-28580-6>.
  78. Wamble, N. P.; Eugene, E. A.; Phillip, W. A.; Dowling, A. W. Optimal Diafiltration Membrane Cascades Enable Green Recycling of Spent Lithium-Ion Batteries. *ACS Sustainable Chem. Eng.* **2022**, *10* (37), 12207–12225.  
<https://doi.org/10.1021/acssuschemeng.2c02862>.
  79. Eugene, E. A.; Phillip, W. A.; Dowling, A. W. Material Property Targets to Enable Adsorptive Water Treatment and Resource Recovery Systems. *ACS EST Eng.* **2021**, *1* (8), 1171–1182.  
<https://doi.org/10.1021/acsestengg.0c00046>.
  80. Ouimet, J. A.; Liu, X.; Brown, D. J.; Eugene, E. A.; Popp, T.; Muetzel, Z. W.; Dowling, A. W.; Phillip, W. A. DATA: Diafiltration Apparatus for High-Throughput Analysis. *Journal of Membrane Science* **2022**, *641*, 119743.  
<https://doi.org/10.1016/j.memsci.2021.119743>.
  81. Wang, J.; Dowling, A. W. Pyomo.DOE: An Open-Source Package for Model-Based Design of Experiments in Python. *AIChE Journal* **2022**, *68* (12), e17813. <https://doi.org/10.1002/aic.17813>.
  82. Stuckner, J.; Piekenbrock, M.; Arnold, S. M.; Ricks, T. M. Optimal Experimental Design with Fast Neural Network Surrogate Models. *Computational Materials Science* **2021**, *200*, 110747.  
<https://doi.org/10.1016/j.commatsci.2021.110747>.
  83. Bradley, W.; Kim, J.; Kilwein, Z.; Blakely, L.; Eydenberg, M.; Jalvin, J.; Laird, C.; Boukouvala, F. Perspectives on the Integration between First-Principles and Data-Driven Modeling. *Computers & Chemical Engineering* **2022**, *166*, 107898.  
<https://doi.org/10.1016/j.compchemeng.2022.107898>.
  84. Kuchemüller, K. B.; Pörtner, R.; Möller, J. Digital Twins and Their Role in Model-Assisted Design of Experiments. In *Digital Twins: Applications to the Design and Optimization of Bioprocesses*; Herwig, C., Pörtner, R., Möller, J., Eds.; Advances in Biochemical Engineering/Biotechnology; Springer International Publishing: Cham, 2021; pp 29–61.  
[https://doi.org/10.1007/10\\_2020\\_136](https://doi.org/10.1007/10_2020_136).
  85. Sansana, J.; Joswiak, M. N.; Castillo, I.; Wang, Z.; Rendall, R.; Chiang, L. H.; Reis, M. S. Recent Trends on Hybrid Modeling for Industry 4.0. *Computers & Chemical Engineering* **2021**, *151*, 107365.  
<https://doi.org/10.1016/j.compchemeng.2021.107365>.
  86. Ceccon, F.; Jalving, J.; Haddad, J.; Thebelt, A.; Tsay, C.; Laird, C. D.; Misener, R. OMLT: Optimization & Machine Learning Toolkit. *J. Mach. Learn. Res.* **2022**, *23* (1), 349:15829–349:15836.
  87. Schweidtmann, A. M.; Bongartz, D.; Grothe, D.; Kerkenhoff, T.; Lin, X.; Najman, J.; Mitsos, A. Deterministic Global Optimization with Gaussian Processes Embedded. *Math. Prog. Comp.* **2021**, *13* (3), 553–581. <https://doi.org/10.1007/s12532-021-00204-y>.
  88. Rall, D.; Schweidtmann, A. M.; Kruse, M.; Evdochenko, E.; Mitsos, A.; Wessling, M. Multi-Scale Membrane Process Optimization with High-Fidelity Ion Transport Models through Machine Learning. *Journal of Membrane Science* **2020**, *608*, 118208.  
<https://doi.org/10.1016/j.memsci.2020.118208>.
  89. Jalving, J.; Ghouse, J.; Cortes, N.; Gao, X.; Knueven, B.; Agi, D.; Martin, S.; Chen, X.; Guittet, D.; Tumbalam-Gooty, R.; Bianchi, L.; Beattie, K.; Gunter, D.; Sirola, J. D.; Miller, D. C.; Dowling, A. W. Beyond Price Taker: Conceptual Design and Optimization of Integrated Energy Systems Using Machine Learning Market Surrogates. *Applied Energy* **2023**, *351*, 121767.  
<https://doi.org/10.1016/j.apenergy.2023.121767>.
  90. Tolaymat, T.; El Badawy, A.; Sequeira, R.; Genaidy, A. A System-of-Systems Approach as a Broad and Integrated Paradigm for Sustainable Engineered Nanomaterials. *Science of The Total Environment* **2015**, *511*, 595–607.  
<https://doi.org/10.1016/j.scitotenv.2014.09.029>.
  91. Hadian, S.; Madani, K. A System of Systems Approach to Energy Sustainability Assessment: Are All Renewables Really Green? *Ecological Indicators* **2015**, *52*, 194–206.  
<https://doi.org/10.1016/j.ecolind.2014.11.029>.

92. Kennedy, M. C.; O'Hagan, A. Bayesian Calibration of Computer Models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **2001**, *63* (3), 425–464. <https://doi.org/10.1111/1467-9868.00294>.
93. Bhat, K. S.; Mebane, D. S.; Mahapatra, P.; Storlie, C. B. Upscaling Uncertainty with Dynamic Discrepancy for a Multi-Scale Carbon Capture System. *Journal of the American Statistical Association* **2017**, *112* (520), 1453–1467. <https://doi.org/10.1080/01621459.2017.1295863>.
94. Eugene, E. A.; Jones, K. D.; Gao, X.; Wang, J.; Dowling, A. W. Learning and Optimization under Epistemic Uncertainty with Bayesian Hybrid Models. *Computers & Chemical Engineering* **2023**, *179*, 108430. <https://doi.org/10.1016/j.compchemeng.2023.108430>.
95. Allen, R. C.; Iseri, F.; Demirhan, C. D.; Pappas, I.; Pistikopoulos, E. N. Improvements for Decomposition Based Methods Utilized in the Development of Multi-Scale Energy Systems. *Computers & Chemical Engineering* **2023**, *170*, 108135. <https://doi.org/10.1016/j.compchemeng.2023.108135>.
96. Teichgraeber, H.; Brandt, A. R. Time-Series Aggregation for the Optimization of Energy Systems: Goals, Challenges, Approaches, and Opportunities. *Renewable and Sustainable Energy Reviews* **2022**, *157*, 111984.
97. Bayramoglu, S.; Nemhauser, G.; Sahinidis, N. *Learning to Branch with Interpretable Machine Learning Models*; National Energy Technology Laboratory (NETL), Pittsburgh, PA, Morgantown, WV, and Albany, OR (United States), 2023. <https://www.osti.gov/biblio/1987701> (accessed 2023-12-01).
98. Anderson, L.; Turner, M.; Koch, T. Generative Deep Learning for Decision Making in Gas Networks. *Math Meth Oper Res* **2022**, *95* (3), 503–532. <https://doi.org/10.1007/s00186-022-00777-x>.
99. Vogel, G.; Schulze Balhorn, L.; Schweidtmann, A. M. Learning from Flowsheets: A Generative Transformer Model for Autocompletion of Flowsheets. *Computers & Chemical Engineering* **2023**, *171*, 108162. <https://doi.org/10.1016/j.compchemeng.2023.108162>.
100. Dowling, A. W.; Biegler, L. T. Degeneracy Hunter: An Algorithm for Determining Irreducible Sets of Degenerate Constraints in Mathematical Programs. In *Computer Aided Chemical Engineering*, Gernaey, K. V., Huusom, J. K., Gani, R., Eds.; 12 International Symposium on Process Systems Engineering and 25 European Symposium on Computer Aided Process Engineering; Elsevier, 2015; Vol. 37, pp 809–814. <https://doi.org/10.1016/B978-0-444-63578-5.50130-4>.
101. Parker, R. B.; Nicholson, B. L.; Sirola, J. D.; Biegler, L. T. Applications of the Dulmage–Mendelsohn Decomposition for Debugging Nonlinear Optimization Problems. *Computers & Chemical Engineering* **2023**, *178*, 108383. <https://doi.org/10.1016/j.compchemeng.2023.108383>.
102. Agi, D. T.; Jones, K. D.; Watson, M. J.; Lynch, H. G.; Dougher, M.; Chen, X.; Carlozo, M. N.; Dowling, A. W. Computational Toolkits for Model-Based Design and Optimization. *Current Opinion in Chemical Engineering* **2024**, *43*, 100994. <https://doi.org/10.1016/j.coche.2023.100994>.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Connecting the Dots: Push and Pull between Technology R&D and Energy Transition Modeling

Justin A. Federici<sup>a</sup>\*, Dimitri J. Papageorgiou<sup>a</sup>, and Robert D. Nielsen<sup>a</sup>

<sup>a</sup> ExxonMobil Technology and Engineering Company, Annandale, NJ, USA

\* Corresponding Author: [justin.a.federici@exxonmobil.com](mailto:justin.a.federici@exxonmobil.com)

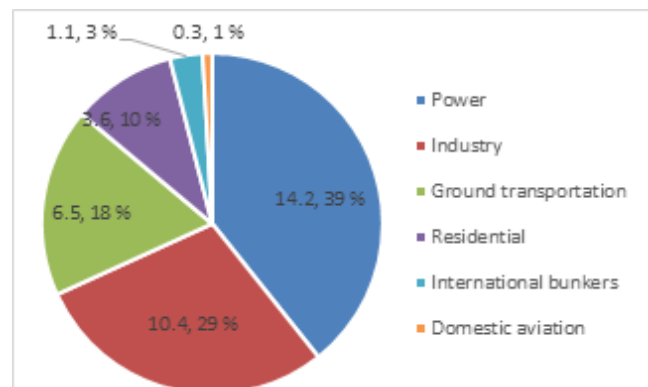
## ABSTRACT

This paper discusses the symbiotic relationship between technology research and development (R&D) and energy transition modeling. On the one hand, energy system modeling has a noteworthy history of providing macroscopic views and critical insights concerning the role that myriad technologies may play in the future energy system. On the other hand, R&D can lead to both incremental and disruptive technological advances that can shape energy transition planning. In this work, we focus on the bidirectional flow of information between the two with a particular focus on highlighting the potential role of carbon capture, storage, and sequestration technology.

**Keywords:** Energy, Multiscale Modelling, Carbon Capture, Carbon Dioxide Sequestration

## INTRODUCTION

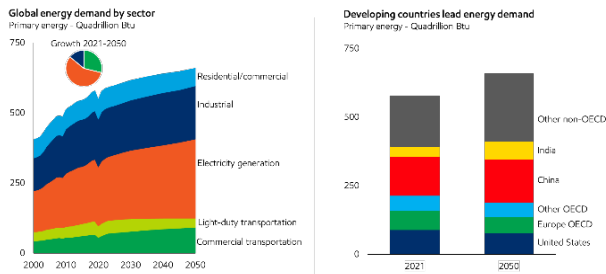
As nations, communities, and organizations continue to determine their role in the energy transition, energy system modeling offers a means for these entities to understand and appreciate the intricacies of today's massive energy system. This is particularly relevant as there has been increasing interest in understanding how the world will address the so-called "energy trilemma," which involves providing low-emissions energy, at low cost, while ensuring energy security from geopolitical risk and economic hardship. Thanks to their macroscopic purview, energy system models can help policymakers understand the trade-offs associated with different energy transition pathways so that they can strike a balance amongst different objectives. At the corporate level, they can inform the direction of a company's technology R&D portfolio as certain technologies may play to the organizations strengths and capabilities and appear more profitable under a wide range of energy transition scenarios while others may only appear profitable under a narrow range of scenarios or may differ widely from the corporate focus. Meanwhile, technology R&D plays a significant role in shaping the assumptions and parameters that populate an energy system model. Improvements in cost and efficiency may allow for certain technologies to become more favorable under a range of societal energy transition goals.



**Figure 1.** 2022 Global CO<sub>2</sub> emissions [Gt] and corresponding percentage. Source: (Liu et al., 2023)

In 2022, global greenhouse gas emissions totalled roughly 36.1 Gt (Liu et al., 2023); see Figure 1. While there exist competitive, low- to zero-carbon technologies to highly decarbonize sectors such as power generation and light-duty vehicle transportation, there remain "difficult-to-eliminate emissions related to aviation, long-distance transportation, and shipping; structural materials; and highly reliable electricity total[ing] ... 27% of global CO<sub>2</sub> emissions from all fossil fuel and industrial sources" (Davis et al., 2018). Zero-carbon technologies may exist for these sectors, but their competitiveness and availability are far less established relative to the incumbent

technologies and infrastructures (Fankhauser et al., 2022). Meanwhile, as shown in Figure 2, global commercial (i.e., non-light-duty) transportation, industrial, and power demand are projected to increase significantly by 2050, particularly in China, India, and other non-OECD countries.



**Figure 2.** Global energy demand (past, present, and projected) by sector and country. Source: (ExxonMobil Global Outlook, 2023)

Numerous organizations indicate that society should pursue a multitude of technologies to decarbonize the energy sector. For example, the United Nations Intergovernmental Panel on Climate Change reports that “global modelled mitigation pathways reaching net zero CO<sub>2</sub> and GHG emissions include transitioning from fossil fuels without carbon capture and storage (CCS) to very low- or zero-carbon energy sources, such as renewables or fossil fuels with CCS, demand-side measures and improving efficiency” (IPCC, 2023, B.6.3). Individual corporations must decide which options within this portfolio they will pursue based on their insights and analysis.

In what follows, we discuss the interplay between energy system modeling and process modeling as they relate to the energy transition and technology R&D. We first discuss current trends in energy system modeling and then offer potential directions for future research. We then link this discussion with the role of process modeling for technology R&D for a specific CCS application. We conclude by discussing opportunities and challenges associated with the bidirectional flow of information between energy system and process models.

## ENERGY SYSTEM MODELING

For decades, energy system models have assisted planners, policymakers, and stakeholders gain insights related to short- and long-term policy decisions. Surveying over 30 energy system models, Lopion et al. (2018), Groissböck (2019), Prina et al. (2020), and Fodstad et al. (2022) review a diverse set of reputable decision support tools based on attributes including data requirements, foresight, methodology (e.g., simulation, optimization), operational/process granularity, sectoral coverage, spatial and temporal resolution, transparency, uncertainty, and more.

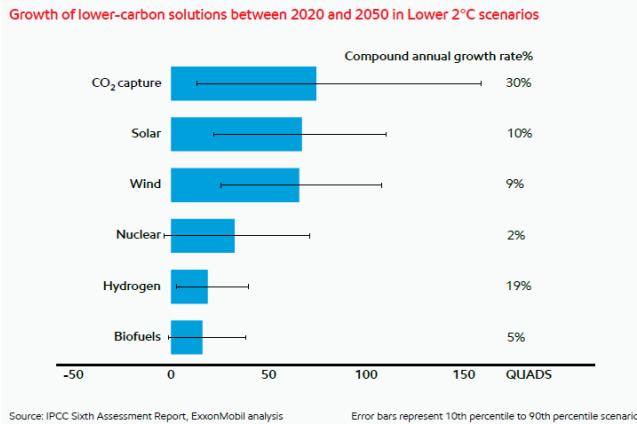
Energy system models are often partitioned into top-down and bottom-up models. Top-down models emphasize connections between the energy system and the larger economy. Due to their simplified energy system representation, they are not always appropriate to analyze sector-specific policies (Prina et al., 2020). In contrast, bottom-up models seek to integrate highly resolved components and interconnections between energy sub-sectors. “These detailed models from a techno-economic point of view allow the user to compare the impact of different technologies on the energy system and to evaluate the best future alternatives to lower GHG emissions for the achievements of the energy targets. However, the bottom-up approach does not take into account the connections between the energy system and the macro-economic sectors, thus neglecting the impacts on these sectors” (Prina et al., 2020). Hybrid models have therefore become popular in recent years to strike a balance between the two approaches.

State-of-the-art energy system models operate at multiple scales (Kakodkar et al., 2023) capturing global detail (as found in integrated assessment), national detail to investigate domestic trends and impacts, and sectoral detail to explore technological transitions in fundamental end-use sectors. Optimization- and simulation-based models are the most common tools, each with their own idiosyncrasies to handle behavioral assumptions of individual actors (i.e., energy sector supply and demand participants). Model outputs typically provide strategic insights to policymakers concerning equity and cost-effective outcomes of alternative energy transitions. We discuss how individual corporations may consume model outputs below.

## Influence of Energy System Models on Technology R&D

In 2022, the IPCC published the Working Group III contribution to the Sixth Assessment Report on Mitigation of Climate Change (IPCC AR6, 2022) and utilized more than 1,200 potential pathways with underlying socioeconomic development assumptions, energy system transformations, and land use change until the end of the century. The IPCC report identified 311 scenarios as “Lower 2°C,” which are defined as pathways with a 67% likelihood of limiting peak warming to below 2°C throughout the 21st century. These scenarios imply a range of lower-emission growth opportunities as highlighted in Figure 3, which looks across the IPCC Lower 2°C scenarios and illustrates the average (blue bars) growth potential of various lower-emission solutions. While all these solutions are needed, the black bars represent the wide range of growth potential across the IPCC Lower 2°C scenarios. To support further deployment of these technologies at scale, additional policies and technology advancements are needed to incentivize investments and

influence consumer behavior. Striking the right balance in investments at a pace consistent with policy support and technology advancements is crucial (P.26 & 28, ExxonMobil ACS 2023).



**Figure 3.** Growth of lower-carbon solutions between 2020 and 2050 in IPCC “Lower” 2°C scenarios. Source: (ExxonMobil ACS, 2023)

Figure 3 offers a concrete example of how energy system models and integrated assessment models can influence technology R&D decisions. The figure summarizes outcomes from models in the open literature from diverse groups with different technology performance and cost assumptions bases. These outcomes give rise to a wide range of potential technology adoption, with certain technologies that are commonly called on at scale, such as CCS. Corporations may use this macroscopic analysis to guide the resources that they allocate towards certain decarbonization technologies that they believe will be most competitive in a low-carbon future.

Technology R&D can also benefit from more microscopic energy system model output, that includes sector and sub-sector modeling. While large scale modeling identifies key technology types, it does not assess the performance at a resolution necessary to evaluate important process trade offs. Resilience and robustness of a technology mix may depend on geographic or temporal factors arising from demand variation and fuel supply. Examples include electricity grid power system modeling that includes weather and seasonal demand inputs. Industrial end-use planning that places a premium on continuity of operation may contemplate exposure to fuel price variation favoring diversity or flexibility of fuel type.

### Open Research Questions on Energy System Models

We conclude this section on energy system modeling with future research opportunities and directions. Prina et al. 2020 highlight four “engineering” challenges for the community: resolution in time, space, techno-economic detail, and sectoral coupling. While these points

are valid, numerous “economic” and “policy” challenges exist and are ripe for future research. Therefore, in contrast to Prina et al. 2020, we highlight orthogonal dimensions including behavioral economics (including endogenous demand modeling), policy uncertainty, technology “spillover” effects, applications of artificial intelligence/machine learning, and sustainability objectives.

While a rich literature exists on how to model and algorithmically handle high-resolution spatio-temporal representations of supply (Kakodkar et al., 2023), Nikas et al. (2020) assert that demand-side modeling is far less mature. “Values, choices, cohesion, culture, and lifestyle shifts in society are indirectly narrated as assumptions,” but not interwoven into the model with the same granularity as given to the supply side (O’Neill et al., 2017). Rebound effects, whereby improvements in technology efficiency lead to an overall increase in energy consumption, are well documented, but rarely modeled. Individuals do not always behave rationally or make decisions based solely on cost optimization. Li (2017) and the references therein highlight “the energy efficiency gap” as a prime example of how energy efficiency measures, which could offer a “cost-effective contribution to GHG mitigation, fuel poverty reduction, and energy security objectives,” have historically witnessed non-cost barriers prevent or diminish their widespread adoption. In short, greater attention to endogenous demand modeling is needed.

Rarely does a policy get implemented without unintended consequences. Well-intentioned policies may fail to deliver desired results. In Europe, inexpensive Asian palm oil is being passed off as used cooking oil because it receives additional credits as the latter; meanwhile, palm oil extraction may be causing deforestation abroad (Reuters 2023). Following a cautionary introduction entitled “The EV transition is harder than anyone thinks,” Charette (2023) highlights recent examples of unintended policy consequences in the EV transition and other sectors. While vehicle-to-grid (V2G) capabilities have the potential to offer energy and reserves to the power grid, modeling V2G policies and vehicle owners’ V2G willingness and behavior remains a challenge.

Technology “spillover” effects may also have a manifest impact on energy system model output. Nemet (2012) presents examples and evidence of inter-technology spillovers for energy technologies in which seemingly unrelated innovations and improvements in one technology impact those of another. “For example, the very commonly used General Electric LM6000 50 MW gas turbine is directly descended from the TF39 high-bypass turbo-fan engine developed for military aircraft, such as the C5 Galaxy, in the 1960s” (Nemet 2012). Meanwhile, spatial spillover effects, in which innovations and cost reductions in one region influence those of another, are well documented (Stephan et al., 2019).



Although some research has been done to address endogenous experience curves and spillover in energy system modeling domain, e.g., (Louwen et al., 2020) and (Straus et al., 2021), more research is warranted.

AI is emerging as a potential tool to assist key steps in energy system modeling. At the front end, applications include capturing information and synthesizing data sets to improve input to energy system models. Improvements can be expected in terms of the fidelity and geographic uniformity of data sets for existing energy generators and infrastructure, as well as sharpening estimates of future demand inputs. Examples include matching separate cost and production data sets where asset identifiers differ; combining satellite and areal data to identify capacity in regions where information is incomplete, e.g., powerlines or residential solar use in emerging countries; demand prediction that combines smart meter, satellite and survey information. Downscaling of highly-resolved data sets and data imputation for incomplete data sets are other possibilities. AI tools may find application in terms of efficient algorithm and constraint development. Examples include efficient approximation to expensive non-linear non-convex problems; formulating constraints that involve multi-objective input, e.g., estimating the total available land resource for renewables based on resource quality and societal or stakeholder goals. Ultimately, AI tools may lead to the synthesis of disparate data to provide system-level case sets that reflect a wholistic picture of energy system supply and demand.

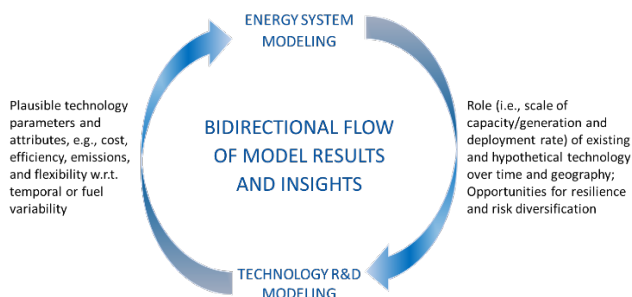
In certain settings, energy system models should attempt to incorporate multiple objectives. Current energy systems are typically constructed with an emphasis on total cost minimization and bounds on greenhouse gas emissions (Algunaibet et al., 2019). Thakker and Bakshi (2022) present a multi-scale multi-objective optimization framework capable of designing chemical reaction pathways, life-cycle value-chains, and economic cash flows in an integrated manner from a superstructure network containing alternative solutions. Sustainability, poverty, affordability, education, and other metrics should be considered. How do we model, or at least incorporate, sustainable development goals (SDGs) (IPCC AR6, 2022)? We need new approaches to include SDG-based metrics in quantitative assessments in energy system modeling.

## TECHNOLOGY R&D

As shown in Figure 4, energy system models and technology R&D have a symbiotic relationship. Energy modeling informs R&D possibilities by helping to identify what role a technology type could play in terms of deployment scale and rate, given an assessment of performance and cost attributes; in other words what technologies are robustly important. Conversely, energy

modeling can illustrate how targeted R&D improvements in technology performance and cost could lead to a wider or more rapid adoption; in other words what it would take for a technology to play a larger role.

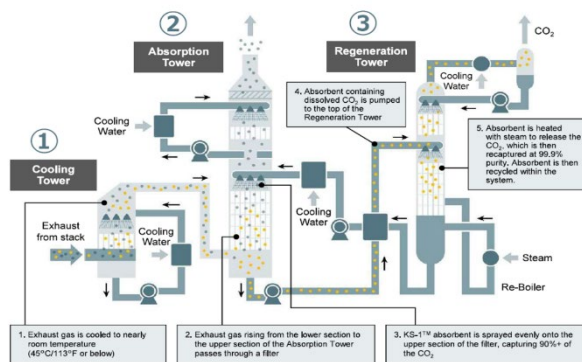
Figure 3 illustrates that CCS technology is expected to play a major role in IPCC “Lower” 2°C energy transition scenarios. More granular modeling could help differentiate specific forms of CCS technology in terms of potential impact. Applications include point source capture as well as distributed opportunities, requiring different modeling approaches. In the case of CCS there is a common underlying technology that could be used for a wide range of applications: amine-based capture. Process modeling bridges to the final step of how to implement a specific technology.



**Figure 4.** Bidirectional flow of information between energy system models and technology R&D models.

## Amine Based Carbon Capture Process

Based on the need to decarbonize various sectors of the global economy, a great deal of R&D has focused on point source post-combustion carbon capture. While there are many different avenues to take CO<sub>2</sub> out of a flue gas stream, to date aqueous amine-based process technology has been proven to be the simplest and most cost-effective method (Herzog 2018). One of these amine-based technologies can be seen in Figure 5.



**Figure 5.** KM CDR Process™ for Postcombustion CO<sub>2</sub> Capture. Courtesy of MHI.

As seen in Figure 5, the process is quite simple. Flue



gas is cooled via direct contact cooling with water (1). It is then sent to an absorption tower (2) where it is contacted with a proprietary, industry performance leading amine-based solvent (KS21) (Kamijo 2023). The flue gas is decarbonized via an intrinsic chemical reaction between the CO<sub>2</sub> acid gas and the basic amine solution. Depending on the design, the percentage CO<sub>2</sub> removal can be anywhere from 90% to 99% removal. Based on recent IEA reports, higher capture rates should be investigated based on the need to reduce CO<sub>2</sub> avoided (still CO<sub>2</sub> leaving the top of the tower) (Thambimuthu 2023). In addition to the amine treating, there are additional clean up steps that occur at the top of the tower (wash sections). The CO<sub>2</sub> that was captured in the aqueous amine solvent moves with the liquid through heat exchangers and is sent to the regeneration tower (3). It is here that steam is used to heat and strip off the CO<sub>2</sub> as a product gas where it can either be compressed/stored or utilized. The regenerated amine is then sent back through the other side of the heat exchanger and the cycle starts again in the absorption tower (2). While a relatively simple process, substantial effort is required to model and predict the process performance accurately.

### Process Modeling Methodologies

When it comes to process modeling, it is always important to understand the model objective. Depending on these objectives, multiple modeling platforms and strategies can be used. Generation of heat and material balances are very useful for understanding steady state energy consumption and steam balance of a CCS plant. These flowrates are also important when designing detailed equipment lists. By combining the two, OPEX and CAPEX can be estimated, financial models developed, and the famous “\$/tonne CO<sub>2</sub>” capture metric can be calculated. It is vital that the underlying assumptions and approximations employed are understood when reporting out such values.

For example, for scoping level estimation of process performance, one may use equilibrium based models for the three towers in Figure 5. This requires an understanding of how to translate a theoretical stage into a specific packing height. Another more accurate method would be to explicitly solve for the packing height via rate based tower models. Software platforms such as ProMax, ProTreat, and AspenPlus all have these features. The Rochelle research group and CCSI are two excellent examples on how to develop these models. Surrogate models can also be developed for detailed process models. Platforms such as Aspen and IDAES are very good for building such reduced order models.

### Linking Process with Plant Economics

While developing models that capture the fundamentals is important for optimization of technical

parameters, it is not enough. Economics must also be understood to make CCS technology investment decisions. This can be a challenge as the skillsets of a research or process engineer may lack project or capital cost estimation. Thus, there has been development on the linkage between process simulators and economic evaluation. For example, the IDAES equation-oriented modeling platform (Miller et al., 2018) has the capability to integrate with cost estimation tools from the Department of Energy (DOE) and the National Energy Technology Laboratory (NETL) as seen in Figure 6.

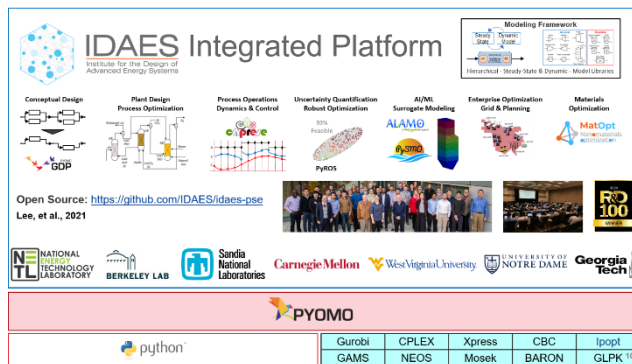


Figure 6. Multiscale IDAES Modeling Platform.

Another example for linking process flowsheet simulation with capital estimation is from AspenTech. Through the concurrent engineering concept of digital thread, a single modeling platform can integrate multidisciplinary teams all working on the same platform.

Through the use of phenomenological, process scale up models, researchers are able to evaluate novel improvement concepts for active materials (amine solvents) at the molecular level. For process engineers, optimal contactor/tower design can be evaluated for these new solvents and integrated into a process design. Various trade offs can be explored on either technical (steam balance) or economic (\$/tonne CO<sub>2</sub> captured) metrics. The end result is a tailored design based on a firm set of model assumptions. This is critical when utilizing these models for energy system modeling.

## BIDIRECTIONAL FLOW AND FUTURE DIRECTIONS

Although there has been extensive work done in the post carbon capture technology space, there is still more work to do. To have an advantaged amine technology, a multiscale, model-centric approach can be used to look at materials (amines), contactors (traditional or intensified), and process configurations to best integrate with current and future power generation and manufacturing plants. As stated before, there is also a need to look at the energy systems scale of states, countries, or even

globally. However, computational intensity becomes an issue based on the shear difference in magnitudes of these systems (molecules to global energy economy). Thus, more academic, government lab, and industrial work should focus on the development of fit for purpose models. These can include: 1) transient models to understand load following and 2) reduced order models for integration in more complex modeling systems (such as IDAES).

A few examples of bidirectional flow of information from post-combustion capture with aqueous amines and energy systems modeling include: 1) reduced-order models (ROMs) as surrogates; 2) transient response of power plants (non-steady operation); and 3) degradation of amine solvent solutions in the presence of flue gas impurities. It is in these three areas that the authors feel like future work should be directed. One great example of this type of work is the FOCUSS (Flexibility Operated Capture Using Solvent Storage) project with SSE Thermal, AECOM, and the University of Sheffield.

## ACKNOWLEDGEMENTS

The authors would like to acknowledge David Tremblay and Mark Lauritsen from AspenTech as well as David Miller from NETL for their help in many technical discussions.

## REFERENCES

- Algunaibet, I.M., Pozo, C., Galán-Martín, Á., Huijbregts, M.A., Mac Dowell, N. and Guillén-Gosálbez, G., 2019. Powering sustainable development within planetary boundaries. *Energy & environmental science*, 12(6), pp.1890-1900.
- Charette R.N., *The EV Transition Explained*. IEEE Spectrum. 2023.
- Davis, S.J., Lewis, N.S., Shaner, M., Aggarwal, S., Arent, D., Azevedo, I.L., Benson, S.M., Bradley, T., Brouwer, J., Chiang, Y.M. and Clack, C.T., 2018. Net-zero emissions energy systems. *Science*, 360(6396).
- ExxonMobil. *Advancing Climate Solutions 2023 Progress Report*. 2023.
- ExxonMobil. *ExxonMobil Global Outlook: Our view to 2050*. 2023.
- Fankhauser, S., Smith, S.M., Allen, M. et al. The meaning of net zero and how to get it right. *Nature Climate Change* 12, 15–21 (2022). <https://doi.org/10.1038/s41558-021-01245-w>
- Fodstad, M., del Granado, P.C., Hellemo, L., Knudsen, B.R., Pisciella, P., Silvast, A., Bordin, C., Schmidt, S. and Straus, J., 2022. Next frontiers in energy system modelling: A review on challenges and the state of the art. *Renewable and Sustainable Energy Reviews*, 160, p.112246.
- Groissböck, M., 2019. Are open source energy system optimization tools mature enough for serious use? *Renewable and Sustainable Energy Reviews*, 102, pp.234-248.
- Howard J Herzog. *Carbon Capture*. Chapter 3. Pages 39 – 46. 20018. The MIT Press Eessential Knowledge Series.
- IPCC, 2022: *Climate Change 2022: Mitigation of Climate Change*. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [P.R. Shukla, J. Skea, R. Slade, A. Al Khourdajie, R. van Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz, J. Malley, (eds.)]. Cambridge University Press, Cambridge, UK and New York, NY, USA. doi:10.1017/9781009157926.
- IPCC, 2023: *Summary for Policymakers*. In: *Climate Change 2023: Synthesis Report*. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland, pp. 1-34, doi: 10.59327/IPCC/AR6-9789291691647.001
- IEA (International Energy Agency). 2021. *Net Zero by 2050: A Roadmap for the Global Energy Sector*. Paris.
- Lopion, P., Markewitz, P., Robinius, M. and Stolten, D., 2018. A review of current challenges and trends in energy systems modeling. *Renewable and sustainable energy reviews*, 96, pp.156-166.
- Kakodkar, R., He, G., Demirhan, C.D., Arbabzadeh, M., Baratsas, S.G., Avraamidou, S., Mallapragada, D., Miller, I., Allen, R.C., Gençer, E. and Pistikopoulos, E.N., 2022. A review of analytical and optimization methodologies for transitions in multi-scale energy systems. *Renewable and Sustainable Energy Reviews*, 160, p.112277.
- Kamijo T. "Update of MHI CO2 capture technology" PCCC7 Plenary Presentation at PCCC7. Pittsburgh, PA. 2023.
- Li F.G. Actors behaving badly: Exploring the modelling of non-optimal behaviour in energy transitions. *Energy Strategy Rev*, 15 (2017), pp. 57-71.
- Liu, Z., Deng, Z., Davis, S. et al. Monitoring global carbon emissions in 2022. *Nature Reviews Earth & Environment* 4, 205–206 (2023). <https://doi.org/10.1038/s43017-023-00406-z>
- Louwen, A., Schreiber, S. and Junginger, M., 2020. Implementation of experience curves in energy-system models. In *Technological Learning in the Transition to a Low-Carbon Energy System* (pp. 33-47). Academic Press.

19. Miller, D.C., Siirola, J.D., Agarwal, D., Burgard, A.P., Lee, A., Eslick, J.C., Nicholson, B., Laird, C., Biegler, L.T., Bhattacharyya, D. and Sahinidis, N.V., 2018. Next generation multi-scale process systems engineering framework. In *Computer Aided Chemical Engineering* (Vol. 44, pp. 2209-2214). Elsevier.
20. Nemet, G.F., 2012. Inter-technology knowledge spillovers for energy technologies. *Energy Economics*, 34(5), pp.1259-1270.
21. Nikas, A., Lieu, J., Sorman, A., Gambhir, A., Turhan, E., Baptista, B.V. and Doukas, H., 2020. The desirability of transitions in demand: Incorporating behavioural and societal transformations into energy modelling. *Energy Research & Social Science*, 70, p.101780.
22. B.C. O'Neill, E. Kriegler, K.L. Ebi, E. Kemp-Benedict, K. Riahi, D.S. Rothman, M. Levy. The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. *Global Environ. Change*, 42 (2017), pp. 169-180.
23. Prina, M.G., Manzolini, G., Moser, D., Nastasi, B. and Sparber, W., 2020. Classification and challenges of bottom-up energy system models-A review. *Renewable and Sustainable Energy Reviews*, 129, p.109917.
24. Stephan, A., Bening, C.R., Schmidt, T.S., Schwarz, M. and Hoffmann, V.H., 2019. The role of inter-sectoral knowledge spillovers in technological innovations: The case of lithium-ion batteries. *Technological Forecasting and Social Change*, 148, p.119718.
25. Straus, J., Ouassou, J.A., Wolfgang, O. and Reigstad, G.A., 2021. Introducing global learning in regional energy system models. *Energy Strategy Reviews*, 38, p.100763.
26. Thakker, V. and Bakshi, B.R., 2022. Multi-scale sustainable engineering: Integrated design of reaction networks, life cycles, and economic sectors. *Computers & Chemical Engineering*, 156, p.107578.
27. Thambimuthu K. "Opening Remarks from IEAGHG for PCCC-7). Sept 25th 2023. Pittsburgh, PA.
28. Reuters. "Growing tensions between Asian palm oil producers and the European Union."  
<https://www.reuters.com/markets/commodities/growing-tensions-between-asian-palm-oil-producers-european-union-2023-01-13/>



© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>

# Sustainable Process Systems Engineering for Chemicals within Planetary Boundaries

Gonzalo Guillén-Gosalbez\*

Institute of Chemical and Bioengineering, Department of Chemistry and Applied Biosciences, ETH Zurich, Vladimir-Prelog-Weg 1, 8093, Zurich, Switzerland

\* Corresponding Author: [gonzalo.guillen.gosalbez@chem.ethz.ch](mailto:gonzalo.guillen.gosalbez@chem.ethz.ch)

## ABSTRACT

The planetary boundaries (PBs) define ecological limits that are critical to preserve the stability of the Earth. Six of them have already been exceeded, which calls for urgent action to optimize industrial systems capable of operating within the safe operating space that they define for humanity. Here we discuss the challenges and opportunities of including PBs in a range of application domains in Process Systems Engineering, focusing on chemicals and fuels production and the use of mathematical programming coupled with life cycle assessment to support sustainable decision-making.

**Keywords:** Environment, Renewable and Sustainable Energy

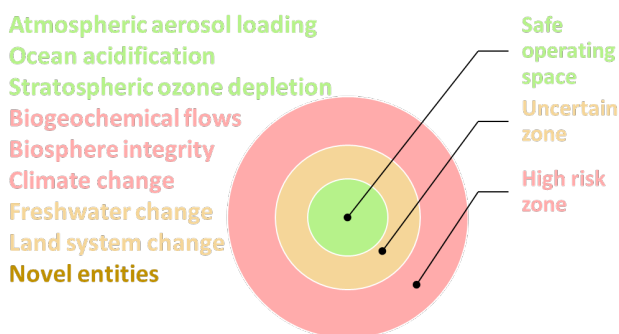
## INTRODUCTION

### The role of the planetary boundaries in the transition to sustainable chemicals

The chemical sector currently faces the challenge of curbing its carbon emissions to meet the climate goals. This will require replacing fossil carbon with renewable carbon as main feedstock, either in the form of captured CO<sub>2</sub> from the air, waste (e.g., polymer waste) or biomass. Unfortunately, emerging renewable-carbon based technologies often lead to inherent trade-offs when compared to the fossil business-as-usual counterpart, not only in terms of cost vs. environmental impacts, but also between environmental categories. For example, chemicals produced *via* carbon capture and utilization (CCU) are often expensive due to the high cost of electrolytic hydrogen, while at the same time could worsen human toxicity impacts owing to the large amounts of renewable power required to activate the CO<sub>2</sub> molecule [1].

This occurrence of burden-shifting (one impact, e.g., global warming, improves at the expense of worsening others) should be carefully investigated to ensure a truly sustainable transition to a defossilized chemical sector. However, while the environmental assessment methods that exist today are useful for comparing alternative technologies, they provide very limited insights into their impact in absolute terms. Because of this, the

broad implications of the large-scale deployment of technologies, specifically concerning their potential global environmental collateral damage, remain unclear, which can lead to spurious conclusions and wrong advice.

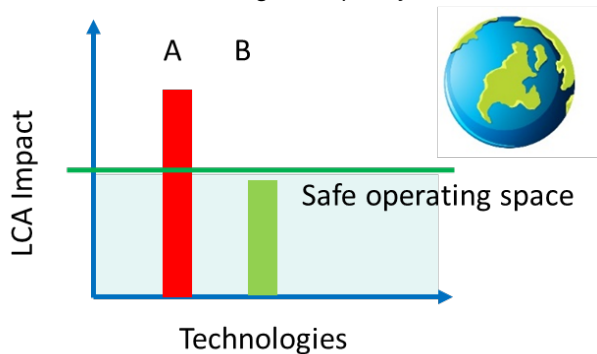


**Figure 1.** Planetary boundaries defined on nine Earth-system processes. Earth-system processes are coloured according to the current transgression level (note that the upper end of the zone of increasing risk has not yet been quantitatively defined for the novel entities PB [3]).

The planetary boundaries (PBs) provide an excellent framework to evaluate the global damage on the planet of emerging routes and the severity of burden-shifting when attempting to combat climate change. The PBs, originally proposed by Rockström et al. [2], define limits on nine Earth-system processes (Fig. 1), all key for

regulating the stability of the planet. For each Earth-system process, thresholds were proposed that define a safe operating space (SOS) for humanity for many centuries to come. Exceeding such limits could trigger critical events that could challenge the resilience of the planet and shift its equilibrium state to a new one with unknown consequences. According to the latest update, six PBs (out of nine) have been already transgressed [3], which calls for urgent action to design and operate chemical systems within the ecological limits of the Earth.

The PBs were not originally intended to be directly applied to industrial systems, yet recent work linking emissions data to impacts on the PBs control variables enabled their application to chemical processes. Standard life cycle assessments (LCAs), the current prevalent approach to evaluate chemical technologies environmentally, lack absolute thresholds to interpret impact values. This shortcoming limits their application to relative comparisons, as already said. Specifically, given two alternative technologies, we could conclude with a standard LCA that one is less environmentally impactful than the other in a given category, but whether they are truly sustainable in absolute terms would remain unclear. The PBs explicitly address this specific question by comparing impact values with a given threshold derived from the carrying capacity of the planet Earth (Fig. 2). If the threshold is exceeded, the system is deemed unsustainable, and would be considered sustainable otherwise. Hence, coupling LCA with PBs allows to explicitly evaluate whether industrial systems operate sustainably within the Earth's ecological capacity.



**Figure 2.** Application of PBs to the assessment of industrial systems. A threshold based on the PBs is established to interpret the impact values quantified *via* LCA.

## PLANETARY BOUNDARIES IN PROCESS SYSTEMS ENGINEERING (PSE)

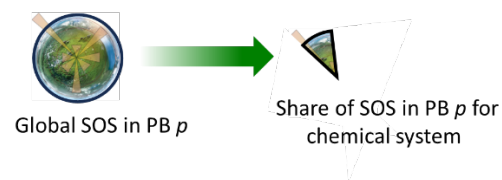
### Planetary boundaries application to industrial systems

The PBs assessment of technologies comprises two steps (Fig. 3). First, a share of the SOS is assigned to the

system studied following some downscaling (sharing) principles. The SOS should be shared among all economic activities jointly, so shares of the total budget need to be defined and allocated to the specific system investigated prior to the assessment. This is a controversial step on which there is no consensus yet and that can result in different environmental budgets depending on the principle applied. Once the threshold is established for a given chemical system, a standard LCA is conducted. However, in the impact phase, impacts on the control variables of the PBs (rather than standard LCIA metrics) are determined using recently developed planetary damage models [4-5]. In the last LCA step, the results are interpreted, and conclusions and recommendations are provided considering the planet-wide impact of the studied industrial system.

The added value of performing an absolute LCA based on the PBs is that it allows classifying systems as sustainable or unsustainable in each Earth-system process. This, in turn, allows interpreting environmental trade-offs by quantifying the extent to which processes contribute to trespassing global boundaries, not only those linked to climate change.

Step 1: Calculate shares of the safe operating space ( $SoSOS_p$ ).



Step 2: Calculate the transgression levels.

$$\text{Transgression}_p = \frac{\text{IMP}_p}{\text{SoSOS}_p} \forall p$$

$$\left\{ \begin{array}{l} < 1 \rightarrow \text{Sustainable} \\ > 1 \rightarrow \text{Unsustainable} \end{array} \right\}$$

**Figure 3.** Application of PBs to the assessment of industrial systems. First, a share of the SOS is determined. Then, the transgression level is calculated by comparing the impact values with such a share.

PBs were originally defined to monitor the environmental state of the planet Earth. Here I argue that they could also be incorporated into the design and operation of sustainable industrial systems as additional constraints/criteria, to enforce that they ultimately comply with global environmental guardrails that are essential for ensuring sustainable development. This paradigm shift towards PBs-based decision-making opens the door for a myriad of applications, within the chemical sector and beyond, for which computer-aided tools quantifying explicitly the PBs impact of engineering decisions across scales could be developed. While, in principle, any industrial process could be evaluated using the PBs, PBs studies are (arguably) better suited to large-scale systems



with large potential impacts that could destabilize Earth-system processes. For example, PBs assessments of bulk chemicals, such as ammonia, might be more relevant as their current fossil production routes emit large amounts of CO<sub>2</sub> contributing strongly to the transgression of the climate change boundary, regarded as a core planetary boundary. On the other hand, PBs assessments of chemicals produced at smaller scales might be less appealing, as the implications for the planet's stability will likely be less critical. Note, however, that the control variable for the novel entities PB is, in principle, independent of the quantities produced (i.e., percentage of synthetics released into the Earth system) [3], so the above might not hold fully true for this case.

Overall, the PBs framework provides valuable insights out of reach for standard LCAs, and here I suggest that they are adopted in sustainability problems to complement other environmental metrics, ultimately guiding research and policymaking more sensibly.

### Application domains in Process Systems Engineering

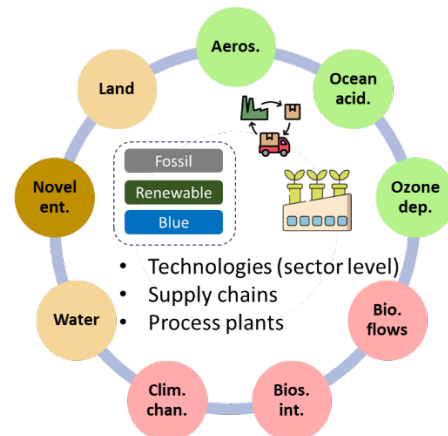
The PBs offer a comprehensive framework to design sustainable chemical systems operating within the SOS. In essence, this could be accomplished by leveraging the existing life cycle optimization framework [6], widely applied in PSE [1, 7-8], in conjunction with the recently proposed PBs characterization factors [4-5]. Following the life cycle PBs-based optimization approach, an optimization model shall be formulated where life cycle assessment principles are explicitly included *via* linear constraints linking mass and energy flows with emissions data and their PBs impact, as shown in the general compact formulation below.

$$\begin{aligned}
 \min \quad & f(x, y) \\
 \text{s. t.} \quad & h(x, y) = 0 \\
 & g(x, y) \leq 0 \\
 & \sum_i CF_{ip} LCI_i \leq SoSOS_p \quad \forall p \\
 & x \in \mathbb{R}^n, y \in \{0,1\}^m
 \end{aligned}$$

Here,  $f(x, y)$  represents the objective function to be optimized, often related to economic performance, whose value is determined from the continuous variables  $x$  (e.g., temperatures, pressures, capacities of supply chain (SC) nodes, etc.), and binary variables  $y$  (e.g., selection of unit operations in a process flowsheet or SC entities in a network problem). Equality and inequality constraints describe the system studied and may include, in addition to the formulas required to compute the objective function value, equations linked to mass and energy balances and thermodynamic constraints in process synthesis problems, or capacity limitations and mass balances in supply chain problems. An inequality constraint

can be added to the model to impose a maximum impact on the PBs. To this end, we link the life cycle inventories entries  $i$  (i.e., life cycle feedstock requirements, emissions and waste calculated for the functional unit taken as a basis in the LCA calculations) to the impact on the control variables of the PBs using tailored characterization factors  $CF_{ip}$  defined for each Earth-system process  $p$ ; moreover, we enforce that the resulting impact should not surpass the share of the SOS allocated to that particular system. Such a PBs constraint might be relaxed, allowing the system to transgress the PBs share, while penalizing such transgression in the objective function using slack variables and penalty coefficients.

The optimization model can take distinct forms depending on the scope of the analysis, e.g., MILP/LP in sectoral and supply chain problems or an MINLP in process synthesis or molecular design problems. We can then capitalize on the rich optimization theory and existing software tools to identify solutions with minimum PBs impact (for further details on mathematical programming applied to sustainable process systems engineering, see [9-12]) Moreover, standard LCAs, extensively employed in the environmental assessments of a wide range of chemical technologies, can be easily enlarged in scope to evaluate the PBs impact, thereby providing additional insights. These conventional LCAs have become very popular in chemical engineering, with increasing applications in the evaluation of low TRL chemical technologies [13].



**Figure 4.** Examples of application domains across scales, including network models optimization (at the sector level), supply chain optimization, and process design.

Focusing on PSE application domains (Fig. 4), the areas of sectoral analysis (network modeling), supply chain optimization and process synthesis seem to offer the largest potential for applying the PBs framework. In network models, a superstructure of technologies is postulated to find the best portfolio to cover the demand of given products, often at minimum cost. Superstructures were for example defined for biomass routes to chemicals and fuels [14], CCU technologies [1] and chemical



recycling pathways [15]. Here the problem is often formulated as an LP, using simplified equations based on linear yields to model technologies. Such simplified network models are typically employed to generate insights into the best routes to meet some demand (top-down) or the best technologies to transform given feedstock into valuable products (bottom-up). The spatial scope can be quite broad, particularly when considering global demands and capacity limitations based on the global availability of resources. Network models are very well suited to identify technologies that are truly sustainable in absolute terms. Here, a share of the global SOS can be allocated to the demanded products to define environmental constraints that should not be violated, as we did in recent work where we studied how to produce plastics sustainably within the PBs [16].

At a higher level of detail, supply chain (SC) models optimize the network topology and associated operations, including the location and capacities of nodes (e.g., plants, warehouses, distribution centers) and the production rates and transport flows between the SC entities. The inclusion of PBs in standard SC formulations is of particular relevance when considering technological decisions, which often lead to critical trade-offs. For example, we recently included PBs in the optimal design of hydrogen supply chains, finding solutions that decrease substantially the pressure exerted on the Earth-system processes by replacing grey with green H<sub>2</sub> [17].

PBs could also be considered in process design problems to find the optimal topology and operating conditions to optimize economic performance while respecting global environmental limits. One option here is to formulate the design problem in global terms, focusing on finding the optimal design of a plant of standard size that could be installed across the world to cover the global demand of given chemicals. Because the technology (chemical pathway) is already fixed here, the potential for environmental gains is typically more modest. This is because once a chemical route is selected, processes implementing such a pathway often operate near the stoichiometric amounts of the required reactants, which often constitute the main environmental hotspot contributing the most to total impacts. Notably, although heat integration can still play a role, impacts (particularly in bulk chemicals) are still mostly dictated by the raw materials' environmental footprint. Consequently, implementing alternative unit operations and optimized operating conditions is unlikely to reduce the impact sharply unless the provenance of the reactants is modified (e.g., grey vs. green H<sub>2</sub> in carbon capture and utilization routes). However, given the large production volumes of some chemicals, particularly bulk chemicals, even marginal improvements are worthy to pursue. Recently, we introduced PBs in the design of a methanol flowsheet, letting the model select the hydrogen source depending on the

objective function optimized. We found that substantial reductions in global impacts can be attained by resorting to renewable carbon feedstock [18].

Lastly, PBs can also be easily incorporated into standard LCAs. For example, using absolute sustainability LCAs based on the PBs, we showed that almost all widespread fossil chemicals transgress at least one PB due to their large CO<sub>2</sub> emissions [19]. We also demonstrated with LCAs based on PBs that current fossil platform chemicals require around one quarter of a planet to operate, also owing to their large carbon footprint, and investigated ways to defossilize chemicals production sustainably by shifting to renewable carbon sources [5].

## CONCLUDING REMARKS

The chemical sector should transition to defossilized technologies to close the carbon loop and operate sustainably within the Earth's ecological capacity. This will require assessing and minimizing the unintended detrimental effects of emerging routes on environmental categories beyond climate change. In this context, the planetary boundaries offer an excellent framework to perform holistic assessments and identify solutions operating within the safe operating space for humanity. Here I argue that due to its systems thinking and powerful computer-aided tools, Process Systems Engineering is in a unique position to embrace this new methodology in current assessments and optimizations of sustainable industrial systems, from molecules, through process to supply chains, sectors and the planet level.

By incorporating the impact on the PBs explicitly in decision-support tools, the broad environmental implications of chemical systems will become clearer. This will allow us to identify the most promising technologies consistent with the planet's ecological capacity early on, guiding experimental research and policy-making more effectively.

## ACKNOWLEDGEMENTS

This publication was created as part of NCCR Catalysis (grant no. 180544), the National Centre of Competence in Research funded by the Swiss National Science Foundation.

## REFERENCES

1. Ioannou, I, Galán-Martín, A, Pérez-Ramírez, J, and Guillén-Gosálbez, G. Trade-offs between Sustainable Development Goals in carbon capture and utilisation. *Energy Environ. Sci.* 16: 113–124 (2023). DOI: 10.1039/D2EE01153K
2. Rockström J et al. A safe operating space for humanity. *Nature* 461: 472–475 (2009). DOI:

- 10.1038/461472a.
3. Richardson K et al. Earth beyond six of nine planetary boundaries. *Sci. Adv.* 9 (37):1–16 (2023). DOI: 10.1126/sciadv.adh2458.
  4. Ryberg, MW, Owsianiak, M, Richardson, K, Hauschild, MZ. Development of a life-cycle impact assessment methodology linked to the Planetary Boundaries framework. *Ecol. Indic.* 88: 250–262 (2018). DOI: 10.1016/j.ecolind.2017.12.065.
  5. Galán-Martín, A, Tulus, V, Díaz, I, Pozo, C, Pérez-Ramírez, J, Guillén-Gosálbez, G. Sustainability footprints of a renewable carbon transition for the petrochemical sector within planetary boundaries. *One Earth* 4 (4): 565–583. DOI: 10.1016/j.oneear.2021.04.001.
  6. Azapagic, A, Clift, R. The application of life cycle assessment to process optimisation. *Comput. Chem. Eng.* 23 (10):1509–1526 (1999). DOI: 10.1016/S0098-1354(99)00308-7.
  7. Hugo A, Pistikopoulos, EN. Environmentally conscious long-range planning and design of supply chain networks. *J. Clean. Prod.* 13 (15): 1471–1491 (2005). DOI: 10.1016/j.jclepro.2005.04.011.
  8. Gebreslassie, BH, Waymire, R, You, F. Sustainable design and synthesis of algae-based biorefinery for simultaneous hydrocarbon biofuel production and carbon sequestration. *AIChE J.* 59 (5):1599–1621 (2013). DOI: 10.1002/aic.14075.
  9. Guillén-Gosálbez, G, You, F, Galán-Martín, A, Pozo, C, Grossmann, IE. Process systems engineering thinking and tools applied to sustainability problems: current landscape and future opportunities. *Curr. Opin. Chem. Eng.* 26:170–179 (2019). DOI: 10.1016/j.coche.2019.11.002.
  10. Bakshi, BR. Methods and tools for sustainable process design. *Curr. Opin. Chem. Eng.* 6:69–74 (2014). DOI: 10.1016/j.coche.2014.09.005.
  11. Barbosa-Póvoa, AP, da Silva, C, Carvalho, A. Opportunities and challenges in sustainable supply chain: An operations research perspective. *Eur. J. Oper. Res.* 268 (2):399–431 (2018). DOI: 10.1016/j.ejor.2017.10.036.
  12. Nikolopoulou, A, Ierapetritou, MG. Optimal design of sustainable chemical processes and supply chains: A review. *Comput. Chem. Eng.* 44 (14): 94–103 (2012). DOI: 10.1016/j.compchemeng.2012.05.006.
  13. Ioannou, I, D'Angelo, SC, Galán-Martín, A, Pozo, C, Pérez-Ramírez, J, Guillén-Gosálbez, G. Process modelling and life cycle assessment coupled with experimental work to shape the future sustainable production of chemicals and fuels. *React. Chem. Eng.* 6:1179–1194 (2021). DOI: 10.1039/d0re00451k.
  14. Kim, J, Sen, SM, Maravelias, CT. An optimization-based assessment framework for biomass-to-fuel conversion strategies. *Energy. Environ. Sci.* 6:1093–1104 (2013). DOI: 10.1039/c3ee24243a.
  15. Meys, R et al. Achieving net-zero greenhouse gas emission plastics by a circular carbon economy. *Science* 374 (6563):71–76 (2021). DOI: 10.1126/science.abg9853.
  16. Bachmann, M et al. Towards circular plastics within planetary boundaries. *Nat. Sustain.* 6:599–610 (2023). DOI: 10.1038/s41893-022-01054-9.
  17. Ehrenstein, M, Galán-Martín, A, Tulus, V, Guillén-Gosálbez, G. Optimising fuel supply chains within planetary boundaries: A case study of hydrogen for road transport in the UK. *Appl. Energy* 276:115486 (2020). DOI: 10.1016/j.apenergy.2020.115486.
  18. Vázquez, D., Guillén-Gosálbez, G. Process design within planetary boundaries: Application to CO<sub>2</sub> based methanol production. *Chem. Eng. Sci.* 246:116891 (2021). DOI: 10.1016/j.ces.2021.116891.
  19. Tulus, V, Pérez-Ramírez, J, Guillén-Gosálbez, G. Planetary metrics for the absolute environmental sustainability assessment of chemicals. *Green Chem.* 23:9881–9893 (2021). DOI: 10.1039/d1gc02623b.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# CO<sub>2</sub> Mitigation in Chemical Processes: Role of Process Recycle Optimization

Diane Hildebrandt<sup>a\*</sup>, James Fox<sup>b</sup>, Neil Stacey<sup>c</sup>, Baraka C. Sempuga<sup>d</sup>

<sup>a</sup> Department of Chemical and Biochemical Engineering, Rutgers, The State University of New Jersey, New Jersey, USA

<sup>b</sup> Helical Energy, London, UK

<sup>c</sup> University of the Witwatersrand, Department of Chemical Engineering, Johannesburg, Gauteng, South Africa

<sup>d</sup> University of South Africa, Institute for Catalysis and Energy Solutions, Johannesburg, Gauteng, South Africa

\* Corresponding Author: [dh985@soe.rutgers.edu](mailto:dh985@soe.rutgers.edu)

## ABSTRACT

In designing low-carbon processes, the unintended emission of CO<sub>2</sub> remains a significant concern due to its global environmental impact. This paper explores carbon management within chemical processes, specifically examining the correlation between the process material balance (PMB) and CO<sub>2</sub> emissions to understand and identify the potential for reducing these emissions. We interrogate the foundational issue of carbon discharge by analyzing the interplay among mass, energy, and entropy balances, which collectively influence the PMB. We introduce the concept of the Target Material Balance (TMB), which represents the material balance of a process corresponding to minimum CO<sub>2</sub> emissions within the given constraints. We could ask what decisions in the design and operation of processes result in higher CO<sub>2</sub> emissions than the TMB. We will focus on the interaction between reactions and recycles and how the arrangement of recycles in processes can inadvertently change the PMB, thereby increasing CO<sub>2</sub> emissions substantially above that of the TMB. We will demonstrate how recycle streams can be introduced to modify the PMB to approximate the TMB more closely. We will finally illustrate these concepts using simple examples, which demonstrate that a carbon-emitting process, such as reforming, can have the TMBs adjusted to create a design that not only reduces carbon emissions but eliminates them entirely.

**Keywords:** Carbon Dioxide, Process Synthesis, Energy, Optimization, Methane Reforming, Process Material Balance, Target Material Balance, Minimizing CO<sub>2</sub> Emissions, Work Analysis, Entropy Analysis

## INTRODUCTION

Process Systems Engineering (PSE) is a connecting force in chemical engineering, providing both a scientific foundation and computational resources to tackle current and future challenges in areas like energy, environmental science, the 'industry of tomorrow,' and sustainability [1]. The methodologies developed within PSE can be classified into heuristic, insight-based, and mathematical optimization approaches [2]. Heuristic methods leverage the designer's experience and intuition, while insight-based approaches depend on deep understanding and analytical insights.

Heat exchanger network synthesis, commonly referred to as "pinch analysis," is a heuristic method developed by Linnhoff and team in the late 1970s [3][4]. The

approach offers a systematic method for reducing energy consumption in industrial processes by optimizing heat recovery systems.

Mathematical optimization techniques systematically enumerate various potential unit operations, alternate system configurations, process integration structures, operating modes, and other crucial aspects within a superstructure representation. Sorin et al. [5] use a reducible superstructure to minimize a system's exergy. However, the effectiveness of a superstructure-based solution is contingent on the comprehensiveness and richness of the proposed superstructure and process units, with no assurance of identifying the global optimum of the system.

The second law of thermodynamics, or entropy, is seldom explicitly employed in process design [6]. Instead, entropy's application is primarily implicit, such as

in determining phase and reaction equilibria, rather than being directly utilized in the process design or synthesis. Echoing this sentiment, Leites et al. (2003) [7] identify a "lack of understanding" of how second law methodologies can enhance process reversibility as a principal factor behind suboptimal designs, which in turn leads to elevated energy consumption. A review of the methodologies for conducting exergy and energy analyses of thermal power plants notes that most approaches either adopt a superstructure methodology or focus on the analysis and optimization of a specific proposed flowsheet [8]. This highlights a gap in leveraging thermodynamic principles, specifically entropy, more effectively and explicitly in pursuing energy-efficient process designs.

The second law of thermodynamics is implicitly used in process design, for example, by minimizing Gibbs Free Energy to predict chemical and phase equilibrium [9] [10], analyze the temperature dependence of reactions in novel chemistries [11], or predict the effect of feed composition on unwanted side reactions [12].

These methods fall short in offering guidance on the explicit application of second-law analysis for ascertaining the performance limits of a reversible process. The scope of the pinch approach has been extended to encompass targeting for both work and heat exchangers [13]. A superstructure methodology is utilized in order to identify the optimal network configuration.

Process Targeting refers to any methodical procedure for determining the fundamental performance limits of a proposed process pathway based on its inputs and outputs. It enables designers to compare process pathways purely based on thermodynamics, unconstrained by the specifics of design and equipment. This conceptual phase offers the broadest optimization space of any stage of design; the optimization spaces of all subsequent stages of design are just subsets of the optimization space defined during Process Targeting.

By omitting this phase of design, the chosen superstructure or flowsheet may operate within a constrained optimization space that may not contain the global optimal performance and, hence, may impose unnecessary performance limits. The majority of a process's economic and environmental impacts are determined during the conceptual stage of design [14], so it is crucial that the superstructure or flowsheet chosen contains sufficient richness and appropriate interconnections so that its performance is contained in full optimization space.

There is also an argument to be made for Process Targeting based on the scientific principle of Falsificationism – a hypothesis regarding process performance cannot be falsified by a plant design failing to meet the stipulations of the hypothesis because that failure could be attributable to specific design decisions or equipment limitations. Conversely, if the fundamental performance

limits defined by Process Targeting do not meet the stipulations of a hypothesis, then that hypothesis is sufficiently falsified.

Hence, initiating the design procedure with Process Targeting is important for two main reasons – it is more compatible with the defined scientific method than other design approaches that start with equipment and/or a pre-defined flowsheet, and it allows for consideration of a broader optimization space for finding the best-performing processes.

This manuscript seeks to demonstrate this design approach by applying it to a straightforward chemical system in the earliest stages of design, by defining a Target Material Balance (TMB) which serves as a performance target against which subsequent Process Material Balances (PMBs) can be compared at each stage of design. At each stage of design, the available optimization space narrows down to a selected subset of the space defined by previous design decisions and, correspondingly, the limits of the PMB gradually become more constrained with each subsequent design decision.

The TMB serves as a goal, indicating an optimum point in terms of a selected objective function [15], and decisions made at each stage of design can be evaluated in terms of how the new constraints it imposes impacts the PMB, that is moves it further from the TMB.

This approach is intended to adhere to the knowledge-generation standards defined by the Falsificationist scientific method and to thereby circumvent design decisions that impose unnecessary constraints during the early design stages.

The proposed approach [15] [16] [17] uses the deceptively simple concepts of material, energy and entropy balances during the conceptual phase of design, as the first and second laws of thermodynamics dictates the outer limits of the optimization space and therefore, the bounds of the PMB.

Heuristics to improve the reversibility of a process have been suggested, including reducing driving forces and the design of the reactor among others [7]. However, the nuanced impact of choices related to recycles on the PMB tends to be overlooked in the initial design phases. These decisions may inadvertently set the bounds of process efficiency and CO<sub>2</sub> emissions, their impact being locked into subsequent design stages, ultimately influencing the profitability and performance of the operating plant. Recognizing this intricate interplay underscores the need for a more comprehensive integration of mass, energy and entropy/work considerations in the early design stages.

## THEORY

A process can be defined by the set of material balances and associated extents [18]. Consider a set of N

species, where a species  $j$  is denoted by  $A_j$  and  $1 \leq j \leq N$ . We specify that  $j$  must either be a feed to or a product (final or intermediate) of the process. We can represent the set of species as vector  $\underline{\mathbf{A}}^T = \{A_1, A_2, \dots, A_N\}$  where  $\underline{\mathbf{A}}^T$  represents the transpose of vector  $\underline{\mathbf{A}}$ . Let  $n_j$  be the number of moles of species  $A_j$ . We can represent the composition of a process stream as vector  $\underline{\mathbf{n}}^T = \{n_1, n_2, \dots, n_N\}$ .

We denote the number of independent material balances (IMBs) that describe the relationship between the elements of  $\underline{\mathbf{A}}$  as  $S$ . IMBs have also been called independent chemical reactions, but in this paper, we will refer to IMBs, as the IMBs may not reflect or coincide to the actual individual reactions occurring in the process. Different procedures to determine the IMBs have been reported by Yin, (1989) [6].

The stoichiometric matrix  $\mathbf{v}$ , with elements  $v_{ij}$  where  $1 \leq i \leq N$  and  $1 \leq j \leq S$ , is defined such that the  $S$  IMBs can be written as  $\mathbf{v}\underline{\mathbf{A}} = \underline{\mathbf{0}}$ . We define the extent vector  $\underline{\boldsymbol{\varepsilon}}$  of the  $S$  IMBs as  $\underline{\boldsymbol{\varepsilon}}^T = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_S\}$ . The relationship between the molar composition of the reaction mixture  $\underline{\mathbf{n}}_{out}$ , the initial or feed composition  $\underline{\mathbf{n}}_{in}$  and the stoichiometric matrix  $\mathbf{v}$  is given by:

$$\underline{\mathbf{n}}_{out} = \underline{\mathbf{n}}_{in} + \mathbf{v}^T \underline{\boldsymbol{\varepsilon}} \quad (1)$$

where the range of  $\underline{\boldsymbol{\varepsilon}}$  is such that all  $n_i \geq 0$

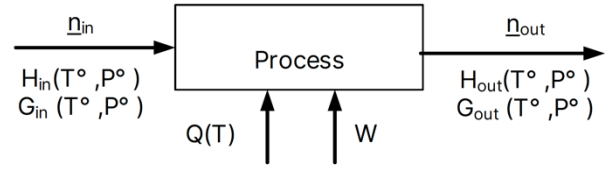
We begin by defining properties that will be used to analyze heat and work flows in a process. The molar enthalpy of component  $A_i$  in a mixture at temperature  $T$  and pressure  $P$  is denoted  $\hat{H}_i(T, P)$ . The molar enthalpies of formation of  $\underline{\mathbf{A}}$  can be combined into vector  $\underline{\hat{H}}(T, P)^T = \{\hat{H}_1(T, P), \hat{H}_2(T, P), \dots, \hat{H}_N(T, P)\}$ . Similarly, the molar Gibbs Free energy of formation vector of  $\underline{\mathbf{A}}$  in a mixture at  $T$  and  $P$  is defined as  $\underline{\hat{G}}(T, P)^T = \{\hat{G}_1(T, P), \hat{G}_2(T, P), \dots, \hat{G}_N(T, P)\}$ , where  $\hat{G}_i(T, P)$  is the molar Gibbs Free energy of formation component  $A_i$  at  $T$  and  $P$ .

The enthalpy of a stream of composition  $\underline{\mathbf{n}}$  at temperature  $T$  and pressure  $P$  is defined by:  $H(T, P) = \underline{\mathbf{n}}^T \cdot \underline{\hat{H}}(T, P)$ . The Gibbs Free Energy of a stream at temperature  $T$  and pressure  $P$  is similarly defined by:  $G(T, P) = \underline{\mathbf{n}}^T \cdot \underline{\hat{G}}(T, P)$ . Consider a process, as shown in Fig. 1, where the inlet and outlet stream are at ambient temperature  $T^\circ$  and ambient pressure  $P^\circ$ , and where a quantity  $Q$  of heat at  $T^\circ$  and work  $W$  are added to the process, as shown in Fig. 1.

An energy balance across the process gives:

$$\Delta H = H_{out}(T^\circ, P^\circ) - H_{in}(T^\circ, P^\circ) = Q(T^\circ) + W \quad (2)$$

Similarly, an entropy balance, combined with eq. (2), assuming reversibility gives, work or Gibbs Free Energy balance:



**Figure 1.** Schematic diagram of mass, heat, and work flows in a generalized process. Note the process is not assumed to be either isothermal or isobaric

$$\Delta G = G_{out}(T^\circ, P^\circ) - G_{in}(T^\circ, P^\circ) = W_{rev} \quad (3)$$

For irreversible processes, the actual work added to the process,  $W$ , is such that  $W > W_{rev} = \Delta G$ . (When  $\Delta G < 0$ ,  $|W| < \Delta G$ ) Thus, more work than the reversible limit (target) is added to processes where  $\Delta G < 0$ , and less work is produced than the reversible limit when  $\Delta G > 0$ .

For the overall process, including utility streams,  $\Delta G \leq 0$  and typically  $\Delta H \leq 0$  in that the process must be designed so as to supply heat and work requirements. Thus, processes need to satisfy the following constraints.

$$\underline{\boldsymbol{\varepsilon}} = (\mathbf{v}^T)^{-1}(\underline{\mathbf{n}}_{out} - \underline{\mathbf{n}}_{in}) \quad (4)$$

$$\Delta H = \mathbf{v}^T \underline{\boldsymbol{\varepsilon}} \cdot \underline{\hat{H}}(T, P) \leq 0 \quad (5)$$

$$\Delta G = \mathbf{v}^T \underline{\boldsymbol{\varepsilon}} \cdot \underline{\hat{G}}(T, P) \leq 0 \quad (6)$$

The PMB is given by  $\mathbf{v}\underline{\mathbf{A}} \cdot \underline{\boldsymbol{\varepsilon}} = 0$  and is thus set by the choice of extents of the IMBs. Furthermore, the PMB also sets  $\Delta H$  and  $\Delta G$  for the process and ensures that eq.(5) and (6) are satisfied. We define the TMB as  $\mathbf{v}\underline{\mathbf{A}} \cdot \underline{\boldsymbol{\varepsilon}}^* = 0$ , where the equality holds for either equation (5) and (6), ie  $\Delta H=0$  and  $\Delta G < 0$  in a heat-limited process or  $\Delta H < 0$  and  $\Delta G=0$  in a work-limited process. The TMB describes the limit or target for the PMB.

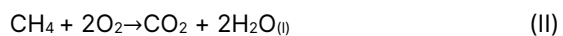
We will explore the relationship between the TMB, PMB and flowsheet by using the concept of the process target on a simple, well studied and industrially important process, namely production of methanol from natural gas

### Example: Analyzing the overall process for conversion of methane to methanol

Consider reforming methane to produce 1mol/s of methanol. We define  $\underline{\mathbf{A}}^T = \{\text{CH}_4, \text{O}_2, \text{CH}_3\text{OH}, \text{H}_2\text{O}, \text{CO}_2\}$ . There are only two IMBs in this case, and the set of independent equations will be defined as:



$$\Delta H_I = -164.3 \text{ kJ/mol}; \Delta G_I = -115.6 \text{ kJ/mol}$$



$$\Delta H_{II} = -890.3 \text{ kJ/mol}; \Delta G_{II} = -817.8 \text{ kJ/mol}$$

Defining the extent of the reactions as  $\varepsilon_I$  and  $\varepsilon_{II}$ , respectively, we see that the requirement of a production rate of 1 mol/s of methanol, sets the value of  $\varepsilon_I=1$ . Thus,



we require:

$$\text{From eq(4) : } \varepsilon_I \Delta H_I + \varepsilon_{II} \Delta H_{II} = -164.3 - 890.3 * \varepsilon_{II} \leq 0$$

$$\text{From eq(5): } \varepsilon_I \Delta G_I + \varepsilon_{II} \Delta G_{II} = -115.6 - 817.8 * \varepsilon_{II} \leq 0$$

We see that the above equations are satisfied for  $\varepsilon_I=1$  and  $\varepsilon_{II}=0$ . As  $\Delta H$  and  $\Delta G$  are negative for IMB (I), the process would be exothermic and would be irreversible if work was not recovered from the process. Combusting more  $\text{CH}_4$  via IMB (II) will make  $\Delta H$  and  $\Delta G$  even more negative, consequently increasing  $\text{CO}_2$  emissions.

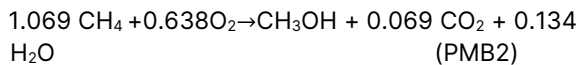
The PMB for  $\varepsilon_I=1$  and  $\varepsilon_{II}=0$  is:



with  $\Delta H=-164.3$  kJ/mol and  $\Delta G=-115.6$  kJ/mol.

Thus, in principle, 1 mole of methane is converted to 1 mole of methanol and no  $\text{CO}_2$  or water is produced as a waste product. The process will also reject heat, and if no work recovery is in place, it will lose 115.6 kJ/mol of work potential.

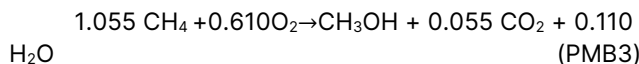
According to Monjur et al. (2021) [20], the PMB for a conventional natural gas to methanol process is:



with  $\Delta H=-225.7$  kJ/mol and  $\Delta G=-172.0$  kJ/mol

PMB2 corresponds to extent  $\varepsilon_I=1$  and  $\varepsilon_{II}=0.069$ . Notice that a conventional process produces both  $\text{CO}_2$  and water as waste products and simultaneously rejects more heat and work than PMB1.

Monjur et al. optimizes the natural gas to methanol flow sheet and finds PMB3 with  $\varepsilon_I=1$  and  $\varepsilon_{II}=0.055$  giving:



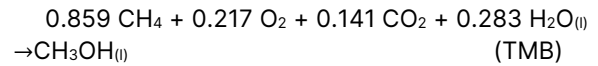
with  $\Delta H=-213.2$  kJ/mol and  $\Delta G=-160.5$  kJ/mol

We see that this improved design reduces the amount of  $\text{CO}_2$  produced and that both  $\Delta H$  and  $\Delta G$  across the process are also less negative compared to PMB2. Thus, the improved process described by PMB3 would be less exothermic and have less lost work than PMB2. This process still, however, does not match the performance of the original PMB 1, where no  $\text{CO}_2$  was emitted.

We could then ask how we might improve PMB3 even further? Without having a process target, one cannot decide if design decisions were made that impose unnecessary constraints on the process described by PMB 3. One could, for example, falsify the proposition that PMB 3 is optimal by proposing that the  $\text{CO}_2$  and  $\text{H}_2\text{O}$  in the product be recycled back into the process. If this were done, then PMB 3 would be transformed to PMB 1. Notice that introducing the recycles would result in reduced flowrates of  $\text{CH}_4$  to the process and simultaneously reduce the heat and work losses. This seems to be a win-win for increasing plant profitability and reducing

environmental impact.

PMB1 is better than PMB2 or PMB3 in terms of feed utilization, less lost work, and lower  $\text{CO}_2$  emissions. However, is PMB1 the "best" PMB? The "best" PMB is the TMB which requires that either  $\Delta H$  or  $\Delta G$  be zero, and PMB1 does not satisfy this. The process target for this system is work-limited and the TMB for  $\Delta G=0$  corresponds to  $\varepsilon_I^*=1$  and  $\varepsilon_{II}^*=-0.141$ ; giving:



with  $\Delta H=-35.5$  kJ/mol and  $\Delta G=0$  kJ/mol

The TMB corresponds to the process target – no process based on IMB's (i) and (ii) can perform better than this. Furthermore, the TMB can be used to compare processes via the PMB to identify if there are opportunities for improvement in process performance. Any improvements will result in improved  $\text{CH}_4$  utilizing, and simultaneously reduce heat and work losses.

Examining the TMB, we see that both  $\text{CO}_2$ , and  $\text{H}_2\text{O}$  are consumed and, by doing so, the reversible process would convert 0.859 moles of methane to 1 mole of methanol. It might appear, because of the negative value of  $\varepsilon_{II}^*$ , that the TMB requires a unit operation where methane is "un-combusted" or equivalently where  $\text{CO}_2$  and  $\text{H}_2$  are converted to  $\text{CH}_4$ . However, (I) and (II) are IMBs and may not correspond to actual reactions occurring in the process. Although PMB1 was better than either PMB2 or PMB3, it is still very far from the TMB.

In summary, the TMB describes the PMB for a reversible process, and so we would not expect a real process to match the TMB. However, the TMB gives information about how much room there is for improvement in the PMB. It can be used as a guide against which design decisions are made, so that as we take irreversibilities in the process into account by modifying the process and thus the PMB, we can benchmark the impact of the decision on the PMB by comparing it to the TMB.

This approach can also be used on sections of a plant to define the TMB across sections of the plant is and, thus, gives insights on how to integrate the mass, heat, and workflows between sections so as to approach the TMB of the overall process.

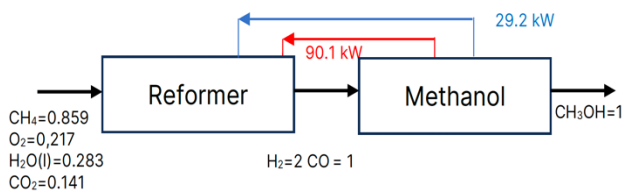
### Example: Analyzing the process subsections in the conversion of methane to methanol

Methane to methanol plants typically consist of two subsections, namely a section which generates the synthesis gas, a mixture of  $\text{CO}$ ,  $\text{CO}_2$  and  $\text{H}_2$ , which we will call the reformer section followed by a methanol synthesis section where the syngas is converted to methanol as shown in Fig. 2.

We will again use a production rate 1mol/s of methanol as a basis. Looking further into the process details,



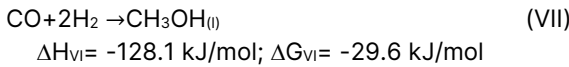
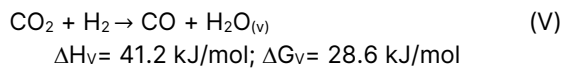
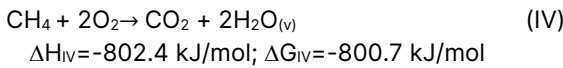
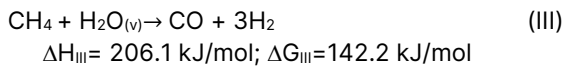
we need to include the components of syngas in the definition of  $\underline{A}$ . Furthermore, water will exist in both the liquid and vapor phases (and thus have different G and H), and so we will include these as two separate species.



**Figure 2.** The methane to methanol process broken in a reformer and methanol synthesis section. Molar flowrates (mol/s) corresponding to the TMB are indicated in black, heat flows in red and workflows in blue

We therefore define  $\underline{A}^T = \{CH_4, O_2, CH_3OH(l), H_2O(l), H_2O(v), CO_2, CO, H_2\}$ .

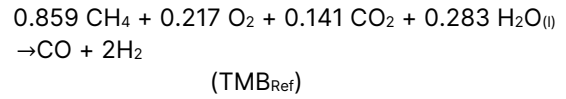
There are 5 IMBs in this case, and we will define the IMBs in terms of reactions or unit operations that may be occurring in the process, namely:



We can calculate the extent of these IMBs that correspond to the overall TMB. In order to achieve the TMB we require  $\underline{\epsilon}^{*T} = \{\epsilon_{III}^*, \epsilon_{IV}^*, \epsilon_V^*, \epsilon_{VI}^*, \epsilon_{VII}^*\} = \{0.75, 0.109, 0.25, 0.283, 1\}$ . For illustrative purposes, we will specify the composition of the syngas, leaving the reformer section as having ratio  $CO:H_2 = 1:2$ . We will further assume that the liquid feed water is evaporated in the reformer section. Thus, IMBs III, IV, V, and VI will occur in the reformer section, while only IMB VII will occur in the methanol section because of the specified composition of the intermediate syngas is in the stoichiometric ratio defined by IMB (VII). This allows us to determine the TMB of the individual sections of the plant that will be consistent with the overall TMB. Conditions (5) and (6) can be lifted for process subsections, with the understanding that a plant subsection with  $\Delta H > 0$  implies that heat integration needs to be done between this plant subsection and the other, which will be exothermic, namely with  $\Delta H < 0$ . Similarly, a plant subsection with  $\Delta G > 0$  implies that work integration

needs to be done between this plant subsection and another with  $\Delta G > 0$ .

The TMB for the reformer section,  $TMB_{Ref}$ , can be calculated using  $\underline{\epsilon}^{*T} = \{\epsilon_{III}^*, \epsilon_{IV}^*, \epsilon_V^*, \epsilon_{VI}^*, \epsilon_{VII}^*\} = \{0.75, 0.109, 0.25, 0.283, 0\}$  giving:



with  $\Delta H = 90.14 \text{ kJ/mol}$  and  $\Delta G = 29.23 \text{ kJ/mol}$

Similarly, the TMB for the methanol synthesis section,  $TMB_{Meth}$ , can be calculated using  $\underline{\epsilon}^{*T} = \{\epsilon_{III}, \epsilon_{IV}, \epsilon_V, \epsilon_{VI}, \epsilon_{VII}\} = (0, 0, 0, 0, 1)$ , giving:



with  $\Delta H = -128.1 \text{ kJ/mol}$  and  $\Delta G = -29.23 \text{ kJ/mol}$

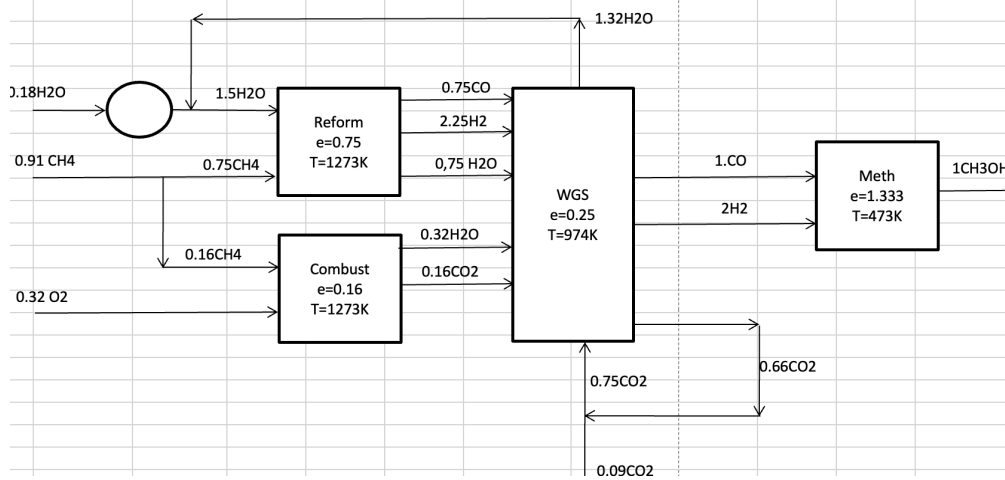
We see that both work and heat will need to be transferred from the methanol synthesis to the reformer section, as indicated in Fig. 2. We will consider the reformer section further to show how we can use the  $TMB_{Ref}$  to develop a flowsheet and related PMB.

### Example: Development of the Reformer Flowsheet using the $TMB_{Ref}$

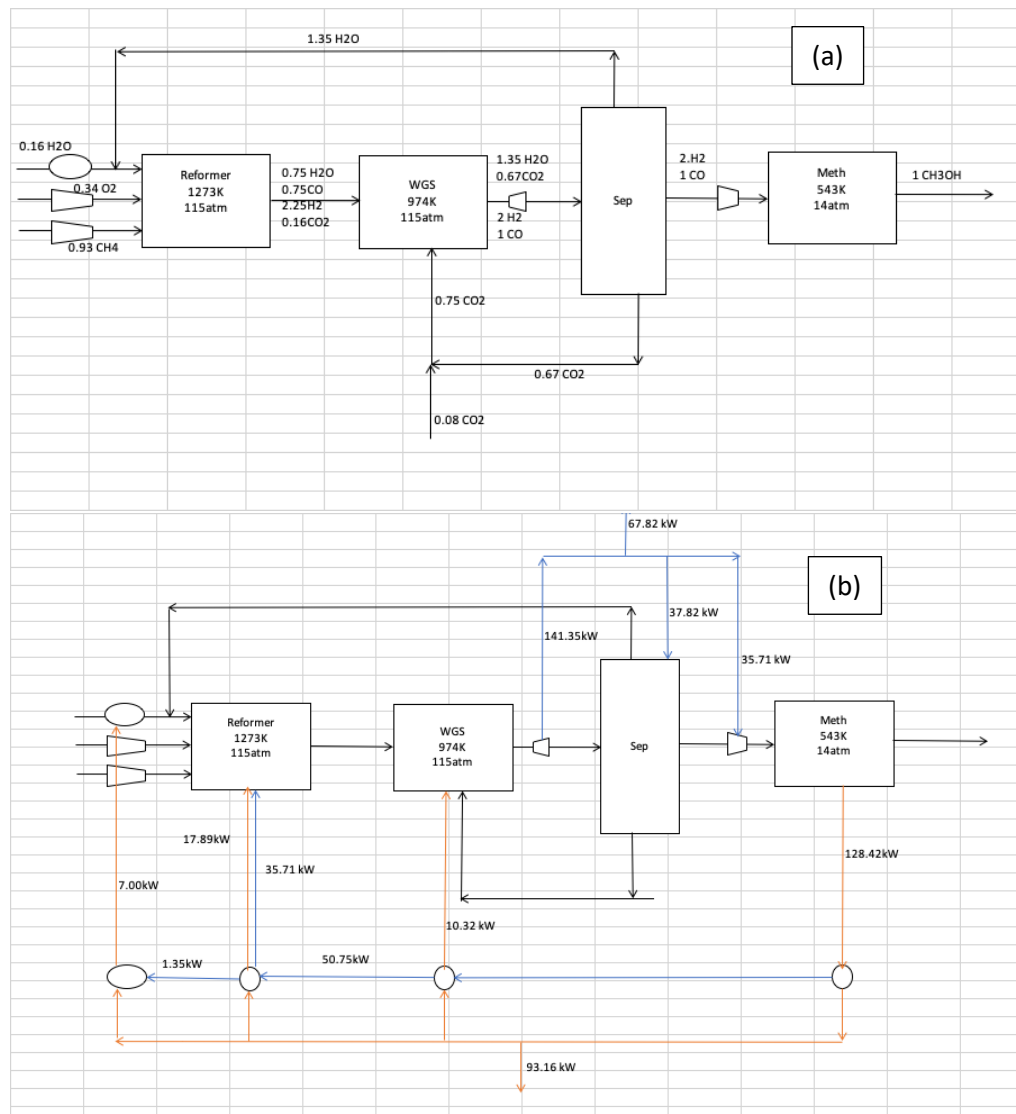
The development of the reformer flowsheet is iterative in terms of setting up the flows and recycles and the heat and work integration. This paper will focus on setting up the mass flows (Fig. 3) and will show an example where heat and mass integration has been incorporated (Fig. 4).

The reformer section consists of a combustion process where  $CH_4$  is combusted to provide the work to drive the reforming reaction, a steam reforming reaction and a water gas shift (WGS) reaction to shift the gas to the correct  $CO:H_2$  ratio of 1:2. The extents of reforming, WGS and combustion are initially set by the  $TMB_{Ref}$ .

There are two exothermic, work producing reactions, namely combustion and methanol synthesis. The amount of work that can be recovered from a process is given by  $\Delta H(1 - T^0/T)$  [16], where  $T$  is the temperature of the heat ( $\Delta H$ ) added or rejected by the reaction. The combustion process is particularly irreversible, as the Carnot temperature for combustion of methane is very large ( $\sim 147000K$ ). Thus, the amount of  $O_2$  fed to the process for combustion must be increased so that the recovered work from both the combustion and the methanol synthesis subprocess is sufficient to supply the overall process work. This is done by increasing the extent of combustion ( $\epsilon_{IV} > \epsilon_{IV}^*$ ) and, consequently, increasing the amount of oxygen and methane fed to the combustion reaction. This design decision introduces process irreversibility which results in the PMB moving away from the target TMB. We describe how the new PMB is calculated below.



**Figure 3:** Example of the molar flows around the reformer section (mol/s) corresponding to PMB 4.



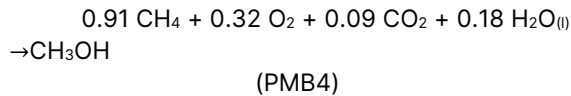
**Figure 4:** (a) Mass flows and (b) heat flows in red and work flows in blue for a process producing 1 mol/s methanol which has been heat and work integrated.

Further design constraints can be specified; for example, we set the reforming temperature at 1273K and specify the ratio of H<sub>2</sub>O:CH<sub>4</sub>=2:1 entering the reformer in order to suppress carbon depositing; as we know the flow rate of methane entering the reformer ( $\varepsilon_{III}$ ) this then sets the water flowrate entering the reformer ( $2\varepsilon_{III}$ ), and consequently the flowrate of water in the stream leaving the reformer as well as the water recycle flowrate. Setting the reforming temperature sets  $\varepsilon_{IV}$  as we require that the work rejected by the combustion when the heat is rejected at 1274K, and methanol synthesis reactions must supply the work for the reforming process, i.e.:

$$\varepsilon_{III}^* \Delta G_{III} = \varepsilon_{VII}^* \Delta G_{VII} + \varepsilon_{IV} \Delta H_{IV} \left(1 - T^0 / 1274\right) \quad (7)$$

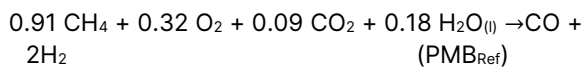
Solving eq.(7) gives that  $\varepsilon_{IV} = 0.16$ , thus setting the flowrate of O<sub>2</sub> to 0.32 mol/s, and the flowrate of methanol entering the combustion process as 0.16 mol/s, which ensures that sufficient work is supplied to the process once the irreversibility of work transfer from combustion is taken into account, as shown in Fig. 3.

In order for the WGS reaction to operate at its Carnot Temperature  $T_c$ , we require that the WGS equilibrium constant,  $K_{WGS}$ , be 1 (i.e.  $\Delta G(T_c, p) = 0$ ), which corresponds to an operating temperature of  $T_c = 974$  K. As we know the flow rate of CO, H<sub>2</sub> and H<sub>2</sub>O leaving the WGS reactor, we can then determine the amount of CO<sub>2</sub> in the product stream of the WGS reactor so as to meet the requirement that ( $K_{WGS} = 1$ ); this then allows us to subsequently calculate the CO<sub>2</sub> recycle flowrate. The resulting flowsheet is shown in Figure 3. The PMB for the overall process in Fig. 3 is given by:



with  $\Delta H = -83.65$  kJ/mol and  $\Delta G = -42.38$  kJ/mol

The PMB for the reformer section is described by:



with  $\Delta H = 44.4$  kJ/mol and  $\Delta G = -12.8$  kJ/mol

The PMB<sub>IV</sub> consumes less CO<sub>2</sub> than the TMB; however, it is considerably better than the PMB's 1, 2 or 3 given earlier. This is mainly due to the use of recycles to ensure that we approach the TMB more closely than in the previous processes. The process flowsheet can be iteratively developed to ensure that the workflows to each section are such that  $\Delta G \leq 0$  for each reaction system. Work can, for example, be *added* to a process by operating at high pressure when there is a decrease in moles of gas across the process. The exact operating pressure can be calculated to ensure that the correct amount of work is added via this route. The flow sheet can be developed, taking temperature constraints and

work flows to ensure that  $\Delta G \leq 0$  in each reactor, and an example of the flow sheet that is mass, heat and work integrated is given in Figure 4 [8]. This process not only consumes CO<sub>2</sub>, but also produces 67 kW of work (based on a production rate of 1 mol/s methanol).

## CONCLUSION

The TMB represents the overall process material balance for either a reversible process (in the case of a G-limited process) or an adiabatic process (in the case of an H-limited process). The TMB also sets the limiting process heat and work flows. The TMB represents the limit of performance of the process; thus the overall material balance across a process, which we refer to as the PMB, may at best, in the limit approach the TMB. The further the PMB is from the TMB, the more room for improvement there may be. Any design decision that results in a sizable departure of the new PMB from the TMB therefore represents the introduction of irreversibility which constrains the performance limits for subsequent design steps.

The TMB can be calculated using only the Gibbs Free Energy and Enthalpies of formation of the feed and product species. The TMB in the case of a G-limited process corresponds to the situation where the chemical potential of the feed (s) is conserved in the product. It also represents the limiting amount of CO<sub>2</sub> that is emitted from a process.

A design procedure can, therefore, reliably and falsifiably search for optimal processes by first defining the TMB and comparing the subsequent PMBs resulting from subsequent design decisions against that benchmark. In this way, the TMB can be used as a tool for iterative design, which conserves as much of the feasible optimization space as possible through each stage of design.

The TMB can also be modified by the inclusion of specific constraints that may reduce the optimization space, such as heat transfer occurring at temperatures other than the process or reaction Carnot temperature. Thereafter, in the case of reversible reactions, recycles can be introduced to ensure that the desired extent of reaction is achieved. In this way one can work toward the desired TMB rather than accepting the PMB given by a simulation.

While the TMB represents a performance benchmark that might not be achievable in real processes, it is useful as a means of identifying, and hopefully avoiding, design decisions that prematurely constrain the optimization space for design which can in some case result in both carbon release and inefficient utilization of the feed. It is also a useful basis of comparison between processes of differing levels of technological maturity, where it can identify which processes have the most inherent potential regardless of current levels of optimization and equipment design, making them a useful tool for long-range

planning and R&D investment decisions.

## REFERENCES

1. Pistikopoulos EN, Barbosa-Povoa A, Lee JH, Misener R, Mitsos A, Reklaitis GV, Venkatasubramanian V, You F, Gani R, Process systems engineering – The generation next? *Comp.&Chem.Eng.* 147: 107252 (2021)
2. Andiappan, V. State-Of-The-Art Review of Mathematical Optimisation Approaches for Synthesis of Energy Systems. *Process. Integr. Optim. Sustain.* 1: 165–188 (2017).
3. Linnhoff B, Ahmad S, Cost optimum heat exchanger networks—1. Minimum energy and capital using simple models for capital cost, *Comp.&Chem.Eng.* 14(7):729-750 (1990)
4. Ahmad S, B. Linnhoff B, Smith R, Cost optimum heat exchanger networks—2. targets and design for detailed capital cost models, *Comp.&Chem.Eng.* 14(7):751-767 (1990)
5. Sorin M, Hammache A, Diallo O, Exergy based approach for process synthesis, *Energy*, 25(2):105-129 (2000)
6. Gool W, Exergy Analysis of Industrial, *Energy*, 17 (8): 791-803 (1992)
7. Leites IL, Sama DA, Lior N, The theory and practice of energy saving in the chemical industry: some methods for reducing thermodynamic irreversibility in chemical technology processes, *Energy*, 29(1):55-97 (2003)
8. Kaushik SC, Siva Reddy V, S.K. Tyagi SK, Energy and exergy analyses of thermal power plants: A review, *Renew. Sustain. Energy Rev.* 15(4): 1857-1872 (2011)
9. Hansen N, van Gunsteren WF, Practical Aspects of Free-Energy Calculations: *A Review J. Chem. Theory Comput.* 10 (7): 2632-2647 (2014)
10. Liu Q, Proust C, Gomez F, Luart D, Len C, The prediction multi-phase, multi reactant equilibria by minimizing the Gibbs energy of the system: Review of available techniques and proposal of a new method based on a Monte Carlo technique *Chem. Eng. Sci.* 216: 115433 (2020)
11. Marques JG, Costa AL, Pereira C. Gibbs free energy ( $\Delta G$ ) analysis for the NaOH (sodium-oxygen-hydrogen) thermochemical water splitting cycle. *Int. J. Hydrogen Energy.* 44(29):14536-14549 (2019)
12. Demidov DV, Mishin IV, Mikhailov MN, Gibbs free energy minimization as a way to optimize the combined steam and carbon dioxide reforming of methane, *Int. J. Hydrogen Energy.* 36(10): 5941-5950 (2011)
13. Lin Q, Liao Z, Sun J, Jiang B, Wang J, Yang Y, Targeting and Design of Work and Heat Exchange Networks, *Ind. Eng. Chem. Res.* 59 (27): 12471-12486 (2020)
14. Uerdingen E. Retrofit Design of Continuous Chemical Processes for the Improvement of Production Cost-Efficiency. ETH Zurich (2002).
15. Patel B, Hildebrandt D, Glasser D, Hausberger B, Synthesis and integration of chemical processes from a mass, energy and entropy perspective. *Ind.Eng.Chem.Res* 46(25):8756-8766 (2007).
16. Sempuga BC, Hildebrandt D, Patel B, Glasser D, Work to chemical processes: the relationship between heat, temperature, pressure and process complexity. *Ind.Eng.Chem.Res* 50(14):8603-8619 (2011).
17. Fox JA, Hildebrandt D, Glasser D, Patel B, Process flowsheet synthesis: reaching targets for idealized coal gasification. *AIChE J* 60(9):3258-3266 (2014).
18. Shinnar R, Feng CA, Structure of complex catalytic reactions: thermodynamic constraints in kinetic modelling and catalyst evaluation. *Ind.Eng.Chem.Fun* 24(2):153-170 (1985).
19. Yin F, A simpler method for finding independent reactions. *Chem.Eng.Comm* 83 (1):117-127 (1989).
20. Monjur MS, Demirel SE, Li J, Hassan MMF, Spice\_mars: A process synthesis framework for membrane-assisted reactive separations. *Ind.Eng.Chem.Res* 60(4):7635-7655 (2021).
21. Fox JA, Hildebrandt D, Glasser D, Patel B, A graphical approach to process synthesis and its application to steam reforming. *AIChE J* 59(10):3714-3729 (2013).

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Towards a Sustainable and Defossilized/Decarbonized Chemical and Process Industry

Mariano Martín<sup>a\*</sup>

<sup>a</sup> Universidad de Salamanca, Department of Chemical and textile Engineering, Salamanca, Salamanca, Spain

\* Corresponding Author: mariano.m3@usal.es

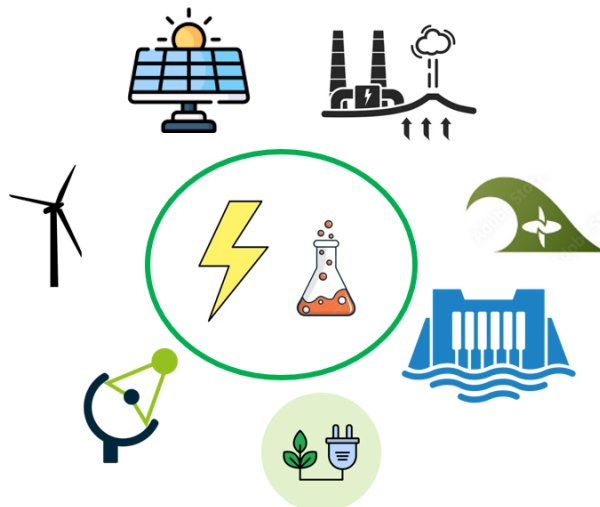
## ABSTRACT

This work presents an overview of the path towards the use of renewable and nonconventional resources for a sustainable chemical and process industry. The aim is not only to lead the way to meet the sustainable development goals but also to maintain the style and quality of life achieved by the technologies and products developed within this sector. Alternative raw materials are to be used and processed differently while a new paradigm for utilities is to be established. The development of technologies and their deployment faces several barriers that we as process engineers can help overcome by providing insight into the alternatives, the thresholds to achieve to become competitive, and strategic analyses.

**Keywords:** Process Design, Renewable and Sustainable Energy, Modelling, Process Synthesis, Energy Storage

## INTRODUCTION

The current chemical and process industry stands at a crossroads. Over the last decades its products have provided a lifestyle and wellbeing to society non precedented. Even though there are thousands of chemicals, most of them can be directly related to eight building blocks (ammonia, methanol, ethylene, propylene, benzene, toluene, and mixed xylenes) that are typically produced from fossil resources. Energy intensive chemical production processes consumed 14% of global oil and 9% of global gas and released 13% of global industrial direct CO<sub>2</sub> emissions in 2020,<sup>1,2</sup> it represents the largest energy consumer and the third largest direct CO<sub>2</sub> emitter.<sup>3</sup> Adding to production the use of the product, the chemical industry accounts for 45% of global greenhouse gas (GHG) emissions. Therefore, reducing the use of resources this sector, it is possible to cut the global emissions by 39% (22.8 Bt).<sup>4</sup> These processes can be divided into different business such as chemicals (including consumer products and pharma), food and beverage, petroleum refining, iron and steel, cement, among others that required utilities of different grades.<sup>5</sup> Thus, the transition towards a new process industry starts from the raw materials but it also must include the utilities required to process them into the final products, such as thermal and electrical energy or water, see Figure 1.<sup>6</sup>



**Figure 1.** More sustainable production system

In the path to meet the sustainable development goals both aspects are to be addressed simultaneously. However, the problem has been addressed by pieces by different research groups. The efforts have been placed either on evaluating alternative sources for the chemicals or the production of utilities. A wider and more systematic view<sup>7</sup> is required for the integration of resources and

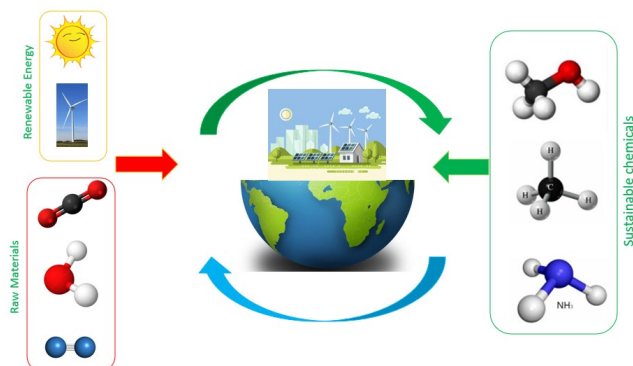


technologies, and the selection of the best use of the natural resources towards sustainability.

In this perspective, we present the different efforts and challenges at process scale and at strategic level as well as the barriers<sup>8</sup> new technologies face that include the social perspective.<sup>9</sup> Once the information is on the table, the opportunities that process system engineering has to offer to help overcome those barriers and contribute to the sustainable transformation of the chemical industry are presented. Chemical engineering has contributed to the past industrial transitions, there is no reason to think why it is not possible to do that again.

## STATE OF THE ART ON THE TRANSFORMATION TECHNOLOGIES

In the path towards a sustainable chemical and process industry there have been several efforts in parallel, raw materials, utilities and the process design itself, but that are intrinsically related. Over the last years there has been a trend to evaluate the possibility of substituting the production of basic chemicals from biomass and waste, specifically the major building blocks, see Figure 2.



**Figure 2.** Use of CO<sub>2</sub> as raw material

### Renewable Raw Materials

It is expected that by 2050 the chemical system will reduce the use of fossil feedstock down to 10–42%.<sup>1</sup> While the early works focused on the use of biomass and wastes to produce biofuels such as ethanol or biodiesel<sup>1</sup>, the extension of the application to substitute fossil fuels has been limited so far. Beyond ethanol and glycerol, as biodiesel major byproduct,<sup>10</sup> that can be used as a source of other chemicals including monomers, (i.e. ethylene, butadiene), the production of platform chemicals from biomass (i.e. Dimethyl furfural, xylitol)<sup>11</sup>, as well as specialty chemicals, (i.e. limonene, phenols)<sup>12,13</sup> has been the follow-up effort. So far two main barriers have been identified, their economics, the higher production cost<sup>1</sup> and the tight margins of the bulk chemicals<sup>14</sup> together with the fact that the petrochemical industry and the related

business are well established. The additional processing steps required in waste processing add cost to the final product, even if the raw material is cheap or even free,<sup>15,16</sup> so that the use of residues is first and foremost a waste management strategy before a true starting point for a circular economy. Only in the case of high added value products it is interesting to use biomass as a resource and biorefineries that produced them as principal product, together with others including biofuels and/or utilities are the most interesting processes.<sup>12,13</sup> The similarities between the crude oil refineries and thermochemical biorefineries have provided the opportunity to retrofit conventional facilities into biobased ones.<sup>17</sup>

The discussion extends and holds true for CO<sub>2</sub> as raw material.<sup>18,19</sup> The increase in the CO<sub>2</sub> concentration in the atmosphere is one of the major concerns because of its effect on global warming. While it is possible to capture it from point sources such as industries (cement, steel) and power plants, direct air capture (DAC) has become an interesting technology to reduce the CO<sub>2</sub> already in the atmosphere. However, recent works show that the energy consumption to capture and further use of the CO<sub>2</sub> can represent a high share leading to a high level of related emissions due to the capture step unless less carbon intense technologies and resources are used to produce PV panels and wind turbines.<sup>20</sup> Utilization of CO<sub>2</sub> rather than just sequestration<sup>21</sup> can provide a way to create a circular economy around it, once the levels in the atmosphere are back into acceptable ones. The use of renewable energy to process CO<sub>2</sub> is one of the first cases of integrating alternative energy sources as utilities within major chemical processes. As a result, CO<sub>2</sub> has been the base for e-fuels including methane, methanol, Dimethyl ether (DME), Fischer Tropsch-fuels among others. The reduction of CO<sub>2</sub> via hydrogenation<sup>14</sup> or electroreduction<sup>22,23</sup> are possible technologies that have been evaluated towards the production of methanol or ethylene and all the way to polymers. However, some are still at low TRL. In addition, most of them can be produced also from biomass, several studies compare both alternatives.<sup>14</sup> Comparing CO<sub>2</sub> and biomass as raw material for the same product, biomass-based processes require more processing stages to prepare the syngas, but in general they are more mature technologies while the need for renewable based electricity results in higher investment costs in Solar panels and wind turbines. In addition, the need to overdesign of the units to operate over a year or the storage of hydrogen represents an additional burden. Biomass based products show lower production costs<sup>14,24</sup> but biomass availability is limited, and it has a wider spectrum of final products represents a decision on the best use is yet to be taken. But it has to be made not at process level, but a more strategic one which calls for a multiscale approach.

### Sustainable Process Design

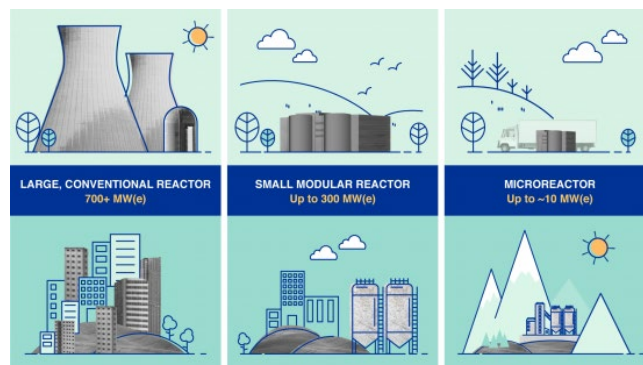


Process design towards efficiency in the use of energy and raw materials is and has been the first step towards sustainability. Selection of processing paths and technologies is an important synthesis problem to be solved. The catch is that novel technologies present a lack of information on their operation and/or the uncertainty in their performance. Modelling and simulation are key for process synthesis and represent a challenge when dealing with novel technologies, but first principles can help and provide the first approach while artificial intelligence (AI) is becoming a powerful tool also. Apart from the economic feasibility, the distributed availability of the resources also jeopardizes the facilities using alternative resources. No longer scale-up and economies of scale play an important role but scale down and the effect that it has on the selection of technologies to build a process.<sup>25</sup> Distributed production and modularization are part of the new design approaches for the exploitation of biomass, solar and wind energy.<sup>26</sup> This is also true for the case of small nuclear reactors<sup>27</sup>, see Figure 3, which also adds the social acceptance as a variable to the deployment of non CO<sub>2</sub> based technologies. To help in the path, circular process design, such as the effort in plastic recycling,<sup>30,31</sup> creating a circular economy around CO<sub>2</sub>, process integration to avoid the use of external chemicals<sup>2</sup> and intensification<sup>32</sup> are additional tools to be developed and implemented.

## Renewable Utilities

It is also important to highlight another line of work that typically progresses in parallel with the production of chemicals. The transformation processes require energy and cooling. Even if the raw material is a residue or a renewable resource, the utilities required must also be provided avoiding fossil-based CO<sub>2</sub> emissions. This has been a common weakness in the analysis of the process. Either the utilities are decarbonized and/or new technologies are to be implemented. However, even new processes using unconventional raw materials have been built based on the same principles, so that the same utilities are used.<sup>28,29</sup>

Utilities refer to electricity and heat representing 20 % and 80% respectively of the total energy consumption in the production of chemicals.<sup>29</sup> Therefore, utilities decarbonization means heat decarbonization,<sup>28,35,36</sup> all the way from steam decarbonization,<sup>37</sup> either using biofuels<sup>20</sup> or hydrogen for high temperature requirements<sup>38</sup>, to the electrification of the units, refrigeration cycles<sup>39</sup> including heat pumps cycle lay out and fluid selection,<sup>40</sup> or the use of Solar based ones such as gasification or reforming or to provide energy for endothermic reactions in general.<sup>41</sup> In addition, recovering heat from waste streams to produce power<sup>33</sup>, and water network<sup>34</sup> design require additional analysis but is proven to reduce the need for external resources.



**Figure 3.** Scale down of technologies. effect on the facility and on the surroundings<sup>27</sup>

As in the case presented for the transformation stages, the use of nonconventional utilities requires unit and process design. In addition, novel process approaches such as intensification, plasma technology, ultrasound microwaves and others have reduced the costs and the generation of wastes.<sup>29</sup> Some of them are related to the use of utilities differently. The design of such novel units is still in their infancy and involves the concept of process and product design and requires considering additional principles beyond traditional chemical engineering since solar and electrical heating are involved. Transport phenomena, unit operations and reactor engineering need to team up with electrical engineering and electro chemistry and can play an important role for the new technology to become competitive. The role of minerals and their availability to build these units can represent another limitation on the extent of the expansion of the penetration of renewable energy.<sup>42</sup> So far only pilot plant scale studies are available. The MeOH is among the first examples of defossilization considering both the raw materials and the utilities involved.<sup>24</sup>

CO<sub>2</sub> utilization as well as utilities decarbonization is linked to the use of renewable energy and electricity and green hydrogen. Carbon capture and utilization (CCUS) involves the production of chemicals out of it, including methane, DME, methanol or urea creating a circular economy for CO<sub>2</sub>. But it is also possible to fully avoid the use of carbon-based chemicals developing non-carbon-based power plants<sup>44,45</sup> via the temporal storage of hydrogen ammonia, MgH<sub>2</sub><sup>44</sup> or Liquid organic hydrogen Carriers (LOCHs). In all these cases, the integration of solar and wind within the chemical transformation is key and affects process design.

## Challenges of Process Operation Integrating Variable Resources

Management of highly volatile resources such as wind and the Sun calls for process and resource integration<sup>46</sup> as well as smart storage, where chemicals can play an important role<sup>47</sup> to reduce the need to build batteries beyond the minerals availability.<sup>48</sup> The operation of

processes that rely on variable resources, variable because of their composition, i.e. biomass, or availability, solar and wind, results in demand site management problems that have been addressed at small scale<sup>45</sup> due to their complexity. In addition, most of the work in heat integration of power to heat for process industries integration has used the pinch technology, but the use of mathematical optimization has proven a powerful tool that has been used traditionally in the petrochemical industry and for the penetration of renewables but the full electrification results in problems that are mathematically really complex due to the variability of the resources and novel tools are being developed<sup>49,50</sup>

### Assessment of Process Sustainability

One question arises here, when a new process is more sustainable than the one is trying to substitute. Recent studies have posed that question and presented interesting results for several examples.<sup>51,52</sup> Further analysis is required to cover the entire spectrum. One on the current limitations is the tools to quantify the environmental impacts. LCA analysis present limitations when novel technologies are evaluated due to the lack data to characterize them. Similarly, the technoeconomic analysis of processes involving novel biomass pretreatments, energy collection technologies as well as new units that substitute the current state of the art to provide utilities<sup>53</sup> to the process represent an additional challenge.

## DEPLOYMENT OF TECHNOLOGIES

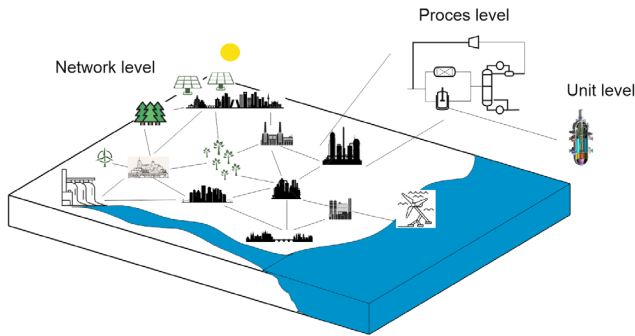
Process level provides the feasibility analysis of the technology and the economic and environmental evaluation that can establish the need for improvements until it can become competitive. But for the production of really high added value products, i.e. polymers or active ingredients for food or the pharma industry,<sup>12,13</sup> additional incentives are needed to substitute/modify production processes that require several years amortization. By enforcing emission levels and providing incentives<sup>54</sup> is how Europe,<sup>55</sup> or the US have been moving towards net zero emissions. However, the transition, the deployment of a decarbonized chemical industry, is to be carried out at a larger scale, at strategic scale. It is at this scale that the incentives need to be designed.

Multiscale approaches have been typically presented within the process community to be able to present strategic decisions to the decision makers, either stakeholders or the governments, see Figure 4.<sup>17</sup> These models require real data, that sometimes are not easy to find, the proper scale-up of the process so that the comparison among alternatives is consistent,<sup>56</sup> an issue that is sometimes overlooked, and the problem, if formulated, is highly complex to solve. Several cases of study are available for biofuels, power at country and continental

level<sup>57,58,59</sup> showing certain limitations as how far we can get. European size examples for electricity and fuels alone are the work of many groups and special solution techniques, variable aggregation, reduced area and time discretization are to be implemented.<sup>57</sup>

However, these studies use the resources towards a particular target. The actual issue is when the raw materials are to be shared to meet the current demand of all products, how far is it possible to get using waste biomass, residues, and renewable resources in general<sup>1</sup>. The ultimate question is what is the best use of the limited resource? what to produce out of it? Therefore, the decision boils down to the availability of natural resources required<sup>60</sup>, including water<sup>61</sup>, area, or mineral among others and the social, environmental, and economic advantages of producing it via that path. Area, water, and nutrients are needed to grow biomass, area is required to install PV panels and wind turbines as well as heliostat fields. To build these units as well as others such as electrolyzers, or bateries<sup>62</sup> we need minerals. Preliminary studies show that it is not possible to cover the first generation of electrified systems, the one that will allow start recycling.<sup>60</sup> On the one hand the actual availability of area, harvesting sites and soil properties have lately been analyzed via GIS, originally for food production<sup>63</sup> but also as a potential for energy crops<sup>64</sup>. To reduce the use of area several strategies arise such as the use of the residues from food production, beyond the needed for animal feed, as well as novel developments in PV panels that are transparent and allow for biomass growing in the same field<sup>65</sup>. On the other hand, planetary boundaries have been presented as a way to evaluate how far it is possible to go with the natural resources,<sup>60,66</sup> but show several issues on how we allocate the possible growth. Such a problem is basically the combination of some of the ones presented for the different cases but has not been addressed so far.

Strategic studies also allow evaluating the social issues that the deployment of the new system can bring to regions that host the new facilities<sup>44</sup>. How to measure circularity and the social impact of the transformation of industry require further analysis and metrics beyond the ones available such as job based<sup>5</sup> indexes and marginalization.<sup>67</sup> The challenge may not be to decarbonize industry but to get to zero net emissions and to create an economy that makes the best use to the resources substituting those that are harmful for the system creating circular economies for those wastes to avoid accumulation in the system.



**Figure 4.** Deployment of technologies

## OVERCOMING THE BARRIERS: INCENTIVES AND SOCIAL IMPACT

The transition to a new system presents opportunities in terms of generation of wealth and jobs due to the installation of transformation facilities. But at the same time, the transition is closing down business, i.e. coal mining and thermal power plants.<sup>44</sup> As such, the decision on where to install them poses a number of social impacts. The first barrier to address is *the social or behavioral one*. Acceptance of a technology has been related to the psychology in product selection, that has been a field of study over the last decades. There is a high activation energy to change. This is similar to customer loyalty to a brand. To address these issues, international agencies campaign to present the need and benefits, such as the United Nations with the SDGs<sup>68</sup> and the effort to convince the population that another production and consumption system is possible. These organizations require systematic studies to provide with reliable data to be able to come to those advertisements. Acceptance of a technology may have an additional positive outcome. On the one hand, society can push their governments to change if the people are convinced. On the other hand, it can be a barrier. One example is nuclear power. Recent efforts focus on small reactors to avoid the society concerns on large facilities close to the cities, see Figure 3.<sup>27</sup> Psychological factors also affect the investors since they need to envision benefits.<sup>54</sup> Apart from this barrier, technical barriers are in place.<sup>69</sup> The *technical barriers* are the ones presented along the first sections of the paper. It is something we as engineers are directly involved in. The third barrier is the so-called *Organizational barrier*. It depends on the companies that are running the business. It is somehow related to the first one in the sense that there is a psychological aspect behind. They need to see the advantage of the new technology. It can also be linked to the incentives that they are going to get.<sup>54,70</sup> Being the first operating a technology provides expertise, at a cost. Technology lock-in only appears when it is widely used<sup>71</sup>. Governments, agencies and other associations, (i.e. EU, UN) play a role in that. The creation of incentives, to

reduce taxes or impose ones to certain technologies, help in the development of the technologies and reduce bureaucracy,<sup>54,72</sup> and present targets such as the ones in emissions or temperature increase, has been the usual one. From the systematic studies of the entire system, it is possible to design and define such incentives since they identify the bottlenecks of the technologies themselves as well as their deployment. The competitors also represent a barrier. New technologies come to take a piece of the cake and represent a threat to a running business. However, some of the companies are already diversifying their portfolios to embrace alternative resources.<sup>54</sup> Finally, *political barriers* are somehow linked to the ideology that not all the times is linked to scientific facts and have an effect on the organizational and the social but also on the development of the technologies due to the incentives Governments can assign and the campaign in favor or against it.

We as process system engineers can play a role in most of them by providing tools, and results that can support the selection of new technologies, the breakeven points (the threshold effect) for one technology to be deployed as well as scenario based analysis for the decision makers to be aware of economical, environmental and social issues as well as defining the incentives<sup>73</sup> required for a particular transition to occur or for a technology to enter the market

## CONCLUSIONS

In this perspective we aimed at presenting the current trends to pave the way towards a more sustainable production system, but in particular, to transform the chemical and process industry into a more sustainable operation. The barriers new technologies are to face come from society itself as well as the stakeholders. Process system engineering, with its systematic analysis of technologies and multi scale studies, are in a pivotal role to provide insights for the decision makers to select the best use of resources considering not only economics and environmental metrics, but also the social impact that the deployment of the new technologies can have.

## ACKNOWLEDGEMENTS

The author acknowledges H2MetAmo project from the Planes complementarios of the Plan de Recuperación, Transformación y resiliencia

## REFERENCES

1. Meng, F et al. (2023) Planet-compatible pathways for transitioning the chemical industry PNAS 120 (8) e2218294120
2. Gabrielli, P., Rosa, L., Gazzani, M., Meys, R., Bardow,

- A., Mazzotti, M., Sansavini, G. (2023) Net-zero emissions chemical industry in a world of limited resources, *One Earth*, <https://doi.org/10.1016/j.oneear.2023.05.006>
3. Chung, C., Kim, J., Socacoo, B.K., Griffiths, S., Bazilian, M., Yang, M. (2023) Decarbonizing the chemical industry: A systematic review of sociotechnical systems, technological innovations, and policy options. *Energy Res. Soc. Sci.* 96, 102955
  4. <https://www.weforum.org/agenda/2023/03/chemicals-industry-low-carbon-economy/>
  5. Rightor, E., Whitlock, A., Elliott, R.N. (2020) Beneficial electrification in industry. *ACEEE*.
  6. USDOE(2022) *Industria decarbonization Roadmap*. USDOE DOE/EE 2635
  7. Martín, M., Grossmann I.E. (2018) Optimal integration of renewable based processes for fuels and power production: Spain case study. *Applied Energy* 213, 595-610
  8. Abdul Qadir, S., Al-Motairi, H., Tahir, F., Al-Fagih, L. (2021) Incentives and strategies for financing the renewable energy transition: A review, *Energy Reports*, 7, 3590-3606
  9. Chung, C., Kim, J., Socacoo, B.K., Griffiths, S., Bazilian, M., Yang, M. (2023) Decarbonizing the chemical industry: A systematic review of sociotechnical systems, technological innovations, and policy options. *Energy Res. Soc. Sci.* 96, 102955
  10. Almena, A., Bueno, L., Díez, M., Martín M (2018) Integrated biodiesel facilities: Review of transformation processes of glycerol based production of fuels and chemicals. *Clean Technol. Environ. Pol.* 20:1639-1661
  11. Galán, G.; Martín, M., Grossmann, I.E. (2021) "Integrated Renewable Production of Sorbitol and Xylitol from Switchgrass. *Ind. Eng. Chem Res.* 60,15, 5558-5573
  12. Criado, A., Martín, M., (2020) Integrated multiproduct facility for the production of chemicals, food and utilities from oranges. *Ind. Eng. Chem. Res.* 59, 16, 7722-7731
  13. Guerras, L.; Sengupta, D.; Martín, M., El-Halwagi, M. (2021) Multi-layer approach for product portfolio optimization: Waste to added value products. *Acs. Sust. Chem. Eng.* 9, 18, 6410-6426
  14. Martín, M (2017) Artificial vs natural reuse of CO<sub>2</sub> for DME production. *Are we getting any close? Engineering.* 3(2) 166-170
  15. Hernández, B., Martín, M. (2017) Optimal Integrated Plant for Production of Biodiesel from Waste *ACS Sust. Chem. Eng.*, 5(8), 6756-6767.
  16. Sánchez A, Martín M.\* (2018) Optimal renewable production of ammonia from water and air, *J. Clean. Prod.*, 178, 325-342
  17. Floudas, C.A., Niziolek, A.M., Onel, O., Matthews, L.R., 2016, *Multi-Scale Systems Engineering for Energy and the Environment: Challenges and Opportunities*. *AIChE J.* 62(3), 602-623
  18. Artz, J., Müller, T.E., Thenert, K., Kleinekorte, J., Meys, R., Sternberg, A., Bardow, A., Leitner, W., (2018) Sustainable Conversion of Carbon Dioxide: An Integrated Review of Catalysis and Life Cycle Assessment. *Chemical Reviews* 118 (2) 434-504
  19. Ioannou, I., Javaloyes-Antón, J., Caballero, J.A., Guillén-Gosálbez, G. (2023) Economic and Environmental Performance of an Integrated CO<sub>2</sub> Refinery. *ACS Sust. Chem. Eng.* 11 (5) 1949-1961
  20. Galán, G., Martín, M., Grossmann, I.E. (2023) Systematic comparison of natural and engineering methods of capturing CO<sub>2</sub> from the air and its utilization. *Sust Prod. Consumpt.* 37, 78-95
  21. Gabrielli, P., Rosa, L., Gazzani, M., Meys, R., Bardow, A., Mazzotti, M., Sansavini, G. (2023) Net-zero emissions chemical industry in a world of limited resources, *One Earth*, <https://doi.org/10.1016/j.oneear.2023.05.006>
  22. Kim, J., Henao, C.A., Johnson, T.A., Dedrick, D.E., Miller, J.E., Stechel, E.B., Maravelias, C.T., (2011) Methanol production from CO<sub>2</sub> using solar-thermal energy: process development and techno-economic analysis. *Energy Environ. Sci.*, 4, 3122-3132
  23. Sharp, S, González-Hernández, S., Chen, C., Sheehan, S.W., (2021) Alcohol production from Carbon dioxide: Methanol as a fuel and Chemical Feedstock. *Joule*, 5 (1), 59-76
  24. Chen, C., Lu, Y., Banares-Alcantara, R. (2019) Direct and indirect electrification of chemical industry using methanol production as a case study. *Appl. Energy.* 243, 71-90
  25. Sánchez, A., Martín, M. (2018) Scale up and Scale down issues of renewable Ammonia plants: Towards modular design. *Sust. Prod. Consumpt.*, 16, 176-192
  26. Baldea, M., Edgar, T.F., Stanley, B.L., Kiss, A.A. (2017) Modular manufacturing processes: Status, challenges, and opportunities *AIChE j.* 63 (10), 4262-4272
  27. <https://www.iaea.org/newscenter/news/what-are-small-modular-reactors-smrs>
  28. Martín, M. (2023) Heat decarbonization: Towards a sustainable utility system. *Joule*, 7,1, 15-17,
  29. Mallapragada D.S. et al (2023) Decarbonization of the chemical industry through electrification: barriers and opportunities. *Joule* 7, 23-41
  30. Li, H., et al. Expanding plastics recycling technologies: chemical aspects, technology status and challenges (2022) *GREEN CHEMISTRY* 24 (23), 9329-9329
  31. Hernández, B., Kots, P., Selvam, E., Vlachos, D.G.,



- lerapetritou, M.G., (2023) Techno-Economic and Life Cycle Analyses of Thermochemical Upcycling Technologies of Low-Density Polyethylene Waste ACS Sust. Chem Eng. 11 (18), 7170-7181
32. Li, Q., Finn, A.J., Doyle, S.J., Smith, R., Kiss, A.A. (2023) Synthesis and optimization of energy integrated advanced distillation sequences Separation and Purification Technology 315, 123717
  33. Anteportalatina-García, V.M., Martín, M (2022) Process synthesis for the valorisation of low-grade heat: Geothermal brines and Industrial waste streams. Renew. Energ. 198, 733-748
  34. Ahmetović, E., Grossmann, I.E., Kravanja, Z., Maréchal, F., Klemeš, J.J., Savulescu, L., Dong, H. (2023) Combined water and heat integration in the process industries Frontiers in Chemical Engineering 4, 1012754
  35. Madeddu, S., Ueckerdt, F., Pehl, M., Peterseim, J., Lord, M., Kumar, K.A., Krüger, C., Luderer, G. (2020) The CO<sub>2</sub> reduction potential for the European industry via direct electrification of heat supply (power- to heat). *Environ. Res. Lett.* **15** 124004
  36. Kim, J.K. (2022) Studies on the conceptual design of energy recovery and utility systems for electrified chemical processes. Renew. Sust. Energy. Revs. 167. 112718
  37. Pérez Uresti, S.I., Lima, R., Martín, M., Jiménez-Gutiérrez, A. (2023) On the design of renewable-based utility plants using time series clustering. Comp. Chem Eng. 170, 108124
  38. Gilbert, T., Menon, A., Dames, C., and Prasher, R. (2022) Heat Source and Application Dependent Levelized Cost of Decarbonized Heat. Joule 7(1) 128-149
  39. Chen, B., González-Ayala, J., Calvo Hernández, A., Luo, R., Yang, H., Guo, J. (2023) A novel electrochemical system with adiabatic pre-charging and pre-discharging processes for efficient refrigeration. Energ. Convers. Manage. 293, 117518
  40. Vermani, S (2022) thermodynamic modelling of high temperature heat pump systems. Msc Thesis TU Delft
  41. Martín, M., (2022) Challenges and Opportunities of Solar thermal energy towards a sustainable chemical industry Comp. Chem. Eng. 165, 107926
  42. Chang, I., Taghizadeh-Hesary, F., Mohsin, M. 820239 Role of mineral resources trade in renewable energy development. Renew. Sust. Ener. Revs. 181, 113321
  43. Davis, S.J. et al (2018) Net- zero emissions energy systems. Science, 360, eaas9793, 1419
  44. Heras, J., Martín, M., (2020) Social issues in the energy transition: Effect on the design of the new power system. Applied Energy. Applied Energy 278 (2020) 115654
  45. Sánchez, A., Castellano, E., Martín, M. (2023) Methanol and Ammonia as Emerging Green Fuels: Evaluation of a New Power Generation Paradigm. Renew. Sust. Revs. 175, 113195
  46. Rightor, E., Whitlock, A., Elliott, R.N. (2020) Beneficial electrification in industry. ACEEE.
  47. Sánchez, A., Martín, M., Zhang, Q. (2021) Optimal Design of Sustainable Power-to-Fuels Supply Chains for Seasonal Energy Storage Energy. 234, 121300
  48. Michaux, S., (2021) Assessment of the Extra Capacity Required of Alternative Energy Electrical Power Systems to Completely Replace Fossil Fuels GTK Open File Work Report 42/2021
  49. Hart, W.E., Laird, C.D., Watson, J.P., Woodruff, D.L., Hackebeil, G.A., Nicholson, B.L., Sirola, J.D., Pyomo optimization modelling in python. Springer
  50. Chen, Q., Johnson, E.S., Bernal, D.E., et al. (2022) Pyomo. GDP: an ecosystem for logic-based modeling and optimization development Optimization and Engineering 23 (1), 607-642
  51. Tulus, V., Pérez-Ramírez, J., Guillén-Gosálbez, G. (2021) Planetary metrics for the absolute environmental sustainability assessment of chemicals Green Chemistry 23 (24), 9881-9893
  52. Cao, G., Handler, R.M., Luyben, W.L., Xiao, Y., Chen, C-h., Baltrusaitis, J. (2022) CO<sub>2</sub> conversion to syngas via electrification of endothermal reactors: Process design and environmental impact analysis. Energ. Convers. Manag. 265, 115763
  53. Pérez Uresti, S., Martín, M., Jiménez Gutierrez, A (2019) Estimation of renewable-based steam costs. Applied Energy 250 (2019) 1120-1131
  54. Abdul Qadir, S., Al-Motairi, H., Tahir, F., Al-Fagih, L. (2021) Incentives and strategies for financing the renewable energy transition: A review, Energy Reports, 7, 3590-3606
  55. A.SPIRE Board of Directors (2020) Processes4Planet Transforming the European Process Industry for a sustainable society
  56. Martín, M., Taifouris, M., Galán, G. (2023) Lignocellulosic biorefineries: A multiscale approach for resource exploitation. Bioresourc. Technol. 385, 129397
  57. Potrč, S., Čuček, L., Martin, M., Kravanja, Z. (2021) Sustainable Renewable Energy Supply Networks Optimization – The Gradual Transition to a Renewable Energy System within the European Union by 2050. Renewable and Sustainable Energy Reviews 146, 111186
  58. Elia, J.A., Baliban, R.C., Floudas, C.A., Gurau, B., Weingarten, M.B., Klotz, S.D., 2013. Hardwood Biomass to Gasoline, Diesel, and Jet Fuel: 2. Supply Chain Optimization Framework for a Network of Thermochemical Refineries. Energy Fuels. 27(8),



4325–4352

59. Marvin, W.A., Schmidt, L.D., Benjaafar, S., Tiffany, D.G., Daoutidis, P. (2021) Economic optimization of a lignocellulosic biomass-to-ethanol supply chain. *Chemical Engineering Science* 67 (1), 68-79
60. Michaux, S., (2021) Assessment of the Extra Capacity Required of Alternative Energy Electrical Power Systems to Completely Replace Fossil Fuels GTK Open File Work Report 42/2021
61. Di Martino, M., Linked, P., pistikopoulos, E.N. (2023) a comprehensive classification of food-energy-water nexus optimization studies: state of the art. *J clean prod.* 420. 138293
62. Zhang, C., Zhao, X., Sacchi, R., You, F. (2023) Trade-off between critical metal requirement and transportation decarbonization in automotive electrification. *Nature Comm.* 14:1616
63. <https://geomarvel.com/analyze-food-security-with-gis/>
64. Fiorese, G., Guariso, G. (2010) A GIS-based approach to evaluate biomass potential from energy crops at regional scale. *Environ. Modell. Soft.* 25 (6) 702-711
65. Lee, K., Um, H.-D., Choi, D., park., J., Kim, N., Kim, H., Seo, K. (2020) The development of transparent photovoltaics. *Cell reports Phys. Sci.* 1 (8) 100143
66. Galan-Martin, A., Tulus, V., Diaz, I., Pozo, C., Perez-Ramirez, J., Guillen-Gosalbez, G. (2021) Sustainability footprints of a renewable carbon transition for the petrochemical sector within planetary boundaries *One Earth* 4 (4) 565-583
67. Matheson, F.I., Dunn, J.R., Smith, K.L., Moineddin, R., Glazier, RH. (2012) Development of the Canadian Marginalization Index: a new tool for the study of inequality. *Canadian Journal of Public Health/Revue Canadienne De Sante'e Publique*, S12–6
68. <https://sdgs.un.org/goals>
69. Chung, C., Kim, J., Socacoo, B.K., Griffiths, S., Bazilian, M., Yang, M. (2023) Decarbonizing the chemical industry: A systematic review of sociotechnical systems, technological innovations, and policy options. *Energy Res. Soc. Sci.* 96, 102955
70. Erickson, E.D., Tominac, P.A., Zavala, V.M. (2023) Biogas production in United States dairy farms incentivized by electricity policy changes *Nature Sustainability*, 1-9
71. Struben, J., Stermann, J. D. (2008) Transition challenges for alternative fuel vehicle and transportation systems *Environment and Planning B: Planning and Design*, 35, 1070 – 1097
72. Nicole, T., Jaehyung, A., Igor, V., Oleg, L., Lyubov, K., Sergey, B., Olga, K. (2022) Renewable energy incentives on the road to sustainable development during climate change: A review *Frontiers in Environmental Science*, 10,

10.3389/fenvs.2022.1016803

73. Martín-Hernández, E., Hu, Y., Zavala, V.M., Martín, M, Ruiz-Mercado, G (2022) Analysis of incentive policies for phosphorus recovery at livestock facilities in the Great Lakes area *Resources, Conservation and Recycling* 177, 105973

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Life Cycle and Sustainability Analyses for Designing Chemical Circular Economy

David Perez<sup>a</sup>, John D. Chea<sup>a</sup>, Jose D. Hernandez-Betancur<sup>b</sup>, and Gerardo J. Ruiz-Mercado<sup>c\*</sup>

<sup>a</sup> Oak Ridge Institute for Science and Education, hosted by Office of Research & Development, US Environmental Protection Agency, Cincinnati, OH 45268, USA.

<sup>b</sup> Faculty of Mines, Universidad Nacional de Colombia, Medellin 050041, Colombia and Department of Chemical Engineering, Universidad de Salamanca, Salamanca 37008, Spain

<sup>c</sup> Office of Research & Development, US Environmental Protection Agency, Cincinnati, OH, 45268, USA and Chemical Engineering Graduate Program, Universidad del Atlántico, Puerto Colombia 080007, Colombia

\* Corresponding Author: ruiz-mercado.gerardo@epa.gov

## ABSTRACT

Sustainability and circular economy enclose initiatives to achieve economic systems and industrial value chains by improving resource use, productivity, reuse, recycling, pollution prevention, and minimizing disposed material. However, shifting from the traditional linear economic production system to a circular economy is challenging. One of the most significant hurdles is the absence of sustainable end-of-life (EoL)/manufacturing loops for recycling and recovering material while minimizing negative impacts on human health and the environment. Overcoming these challenges is critical in returning materials to upstream life cycle stage facilities such as manufacturing. Chemical flow analysis (CFA), sustainability evaluation, and process systems engineering (PSE) can supply chemical products and processes performances from environmental, economic, material efficiency, energy footprint, and technology perspectives. These holistic evaluation techniques can improve productivity, source material reduction, reuse, recycling, and prevent and minimize releases and disposal rates. Therefore, this contribution offers a computational framework that covers CFA, sustainability assessment, and risk evaluation for quantifying the benefits and challenges of chemical circular economy routes versus conventional linear systems. Finally, this contribution shows promising techniques and challenges for employing CFA, sustainability evaluation, and PSE as multicriteria decision-making tools for designing a closed-loop chemical management infrastructure and transforming the US chemical industry sector from linear to circular.

**Keywords:** Supply Chain, Life Cycle Analysis, Modelling, Machine Learning, Wastewater

## INTRODUCTION

### Tracking Chemicals and Releases at Industrial End-of-Life Activities

The Frank R. Lautenberg Chemical Safety for the 21st Century Act amended Toxic Substances Control Act (TSCA) instructs the US Environmental Protection Agency (USEPA) to conduct risk evaluations of existing high-priority chemicals to find whether a chemical substance in the US market may pose an unreasonable risk of harming the environment or human health across its life cycle stages (manufacturing, use, end-of-life (EoL))

[1]. Therefore, chemical risk evaluation supports alternative assessments of selecting chemicals and materials with safer profiles and ensures sustainable circular economy paths (recycling, reusing, and recovering processes). However, the list of registered chemicals used in US commerce continuously grows. Also, data gathering, reconciliation, and management for performing life cycle risk evaluation are time-consuming and challenging. Notably, evaluating the chemical risk and safety at its EoL stage is more difficult due to high uncertainty, less data availability, variability, material blends, and the absence of directing requirements for proper reporting and traceability.

Several efforts aim to create a data engineering framework for building an EoL database and tracking chemicals at their EoL stage. This framework performs a data engineering procedure to gather, clean, wrangle, and integrate information from publicly available USEPA databases such as the Toxics Release Inventory (TRI), the Facility Registry Service (FRS), and the Substance Registry Services (SRS). Furthermore, the framework supports data cleaning, filtering, filling data gaps, and transformation into machine-readable queries to address and enhance chemical risk evaluation of chemicals of concern.

As described in **Figure 1**, the first module of the EoL framework addresses EoL flows generated by industrial facilities (manufacturing and industrial use life-cycle stages) and transferred for further EoL management at recycling, energy recovery, treatment, and disposal facilities (RETDFs). Also, a case study of methylene chloride (dichloromethane) showed the practicality of the developed framework for tracking and analyzing EoL chemical flows from the generator industry sector of a chemical (e.g., inorganic chemical manufacturing), conditions of use (e.g., reactant, formulation component, products), the EoL facility industry sector (e.g., cement manufacturing), EoL activity (e.g., recycling, incineration, landfill), and the environmental compartment (e.g., net recycling, onsite soil releases, etc.), and estimating potential releases at EoL activities [1].

Moreover, Hernandez-Betancur et al. (2021b)

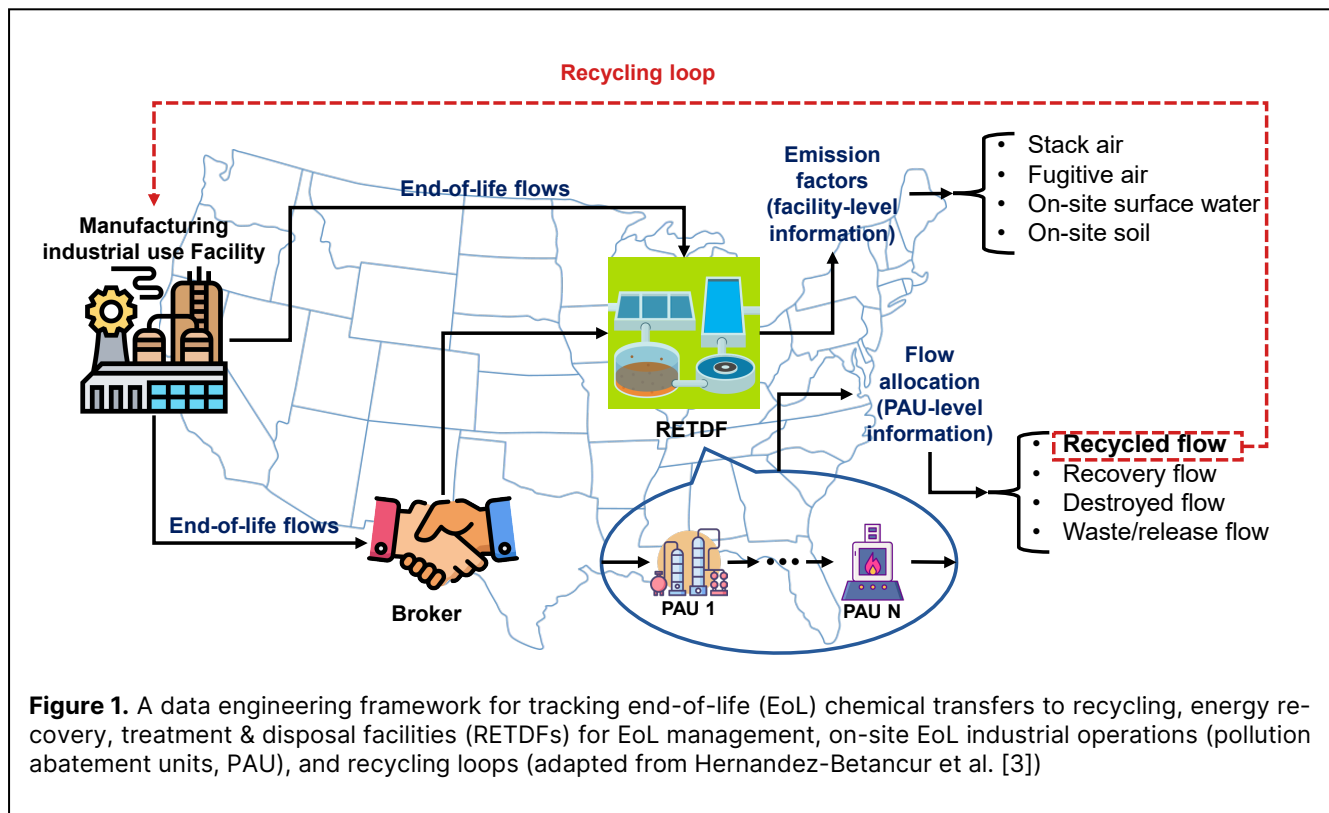
enhanced the data engineering framework by using EoL pollution abatement units (PAU)-level information, chemical process design, and equipment performance data for calculating emission factors and distributing releases [2]. On-site EoL activities consist of PAUs like distillation, aerobic treatment, neutralization, incinerator, etc. The enhanced framework also offers structure design alternatives, efficiencies, chemical releases, exposure media, operating expenses, and capital expenditures for the PAU operations.

Although the developed data engineering framework estimates EoL industrial releases and distribution, it must supply potential circular economy paths and destination of recycled/recovered chemical flows transferred to upstream chemical life cycle stage facilities (e.g., manufacturing).

## METHODOLOGY

### Data Analytics Techniques Supporting Safer Circular Economy Paths

The presence of chemicals causing significant adverse human health and environmental effects during EoL stages is a challenge for implementing sustainable materials management efforts towards a circular life-cycle path of chemicals. Therefore, conducting chemical risk evaluation and exposure assessment of potential EoL scenarios can aid the understanding of the chemical EoL management chain for its safer use in a circular life-cycle



environment.

Thus, we enhanced the data engineering framework by implementing probabilistic graphical models like Markov random field to mimic the relationship between the chemical EoL management components like the EoL flow generators, RETDFs, and manufacturing facilities [3]. The updated framework finds potential safer circular economy scenarios for recycled chemicals. These circular economy scenarios consider whether the recycled chemical might end up in industrial, commercial, consumer uses, and processing operations. Additionally, the framework can find if the recycling activities target the chemical under analysis instead of managing other chemical substances, such as the EoL multicomponent material flow transfers. The novel approach also considers regulatory constraints on closing the recycling loop operations and supplies an estimation of environmental releases related to occupational exposure.

## RESULTS AND DISCUSSION

Completing life cycle and sustainability analyses requires adapting different strategies based on the varying levels of information available. In this computational framework, data-driven model systems may be used to accurately track and predict material pathways throughout the life cycle when process, chemical, and material exchange information are available in larger quantities. Conversely, chemical flow analysis (CFA) may be used to qualitatively map the allocation of chemicals and quantify the relative mass flow distribution across EoL pathways and technologies. The output of these material tracking methods can establish a basis for sustainability analysis because the process efficiency and material releases can be calculated. With sufficient process-level information, PSE models can enhance the granularity of the chemical release approximation, such as releases from individual process units. Subsequently, the operation details can be acquired and used to identify the potential occupational exposure scenarios and areas of improvement throughout a chemical life cycle analysis to achieve a circular economy.

### Data-Driven Model Systems to Estimate Releases from Novel Chemical End-of-Life Scenarios

Hernandez-Betancur et al. (2023a) enhanced the EoL database by integrating cross-year and country data and considering the change of the regulatory foundations across different geographical locations and reporting years [4]. Thus, the data engineering framework is scalable and adaptable to develop machine learning (ML) models for predicting potential EoL activities, chemical

flow allocation, environmental releases, and exposure routes for chemicals outside the comprehensive EoL database.

Hernandez-Betancur et al. (2023b) utilized the EoL database created by Hernandez-Betancur et al. (2023a) to investigate tree-based ML algorithms to construct an EoL release modeling approach influenced by Quantitative Structure-Activity Relationship (QSAR) models often used in cheminformatics and toxicology research [5]. In addition, the developed ML models use chemical descriptors, EoL chemical transfer quantities, industry sectors, environmental policies, industry-sector value-added, chemical structure, and chemical prices to predict potential EoL activities and the chemical transfer distribution among the EoL activities. Therefore, these ML models aid early stage chemical process/product designers in deciding if chemical substances at their potential EoL stage meet environmental regulations or need redesign before entering the market.

Although these data-driven developments based on top-down facility information aid the tracking and estimation of EoL chemical releases and finding potential occupational exposure routes, decision-makers would need to incorporate PSE methods and tools to design and evaluate chemical processes and implement enhancements to minimize or end negative environmental and economic impacts at process equipment level toward a more holistic sustainable performance (bottom-up approach).

### Process Systems Engineering to Estimate Releases and Evaluate Sustainability from Chemical End-of-Life Scenarios

Chemical process synthesis, design, modeling, and simulation developments aid stakeholders in performing CFA, occupational exposure, energy footprint, release estimation, and life cycle inventory generation during EoL chemical stages. For example, Chea et al. (2023) performed a qualitative and quantitative tracking of plastic chemical additives to estimate their potential environmental releases, life cycle inventories, occupational exposure, greenhouse gas emissions, energy footprint, and environmental impacts within the plastic EoL stage and activities [6]. Also, this work shows a sensitivity analysis of future chemical additive release scenarios involving chemical recycling, increasing recycling rates, and chemical additive extraction. To complement these PSE applications, Agbleze et al. (2023) showed recent efforts to estimate chemical releases and evaluate sustainability indicators at the chemical manufacturing stage of existing chemicals, new manufacturing processes, and dynamic operating conditions [7]. Also, this contribution describes how to apply the Gauging Reaction Effectiveness for the ENvironmental Sustainability of Chemistries with a multi-Objective Process Evaluator (GREENSCOPE) sustainability assessment tool to evaluate the sustainability

of chemicals synthesized from a laboratory-scale reaction or manufactured in a plant-scale reaction [8], [9]. GREENSCOPE quantifies and assesses the sustainability of a given process or supply from a multi-stakeholder perspective, including material efficiency, environmental impacts, energy requirements, and cost through mathematical indicators.

Finally, after describing recent PSE applications on EoL material flow analysis for supporting chemical risk assessment, sustainability evaluation, and safer circular economy paths, the subsequent sections describe crucial PSE applications for managing the EoL stage of novel product manufacturing techniques like additive manufacturing (AM) and quantifying and tracking chemical releases in unconventional EoL paths like publicly owned treatment works (POTWs). These efforts aid decision-makers in finding new potential chemical exposure routes and updating regulatory requirements due to ineffective EoL chemical management methods.

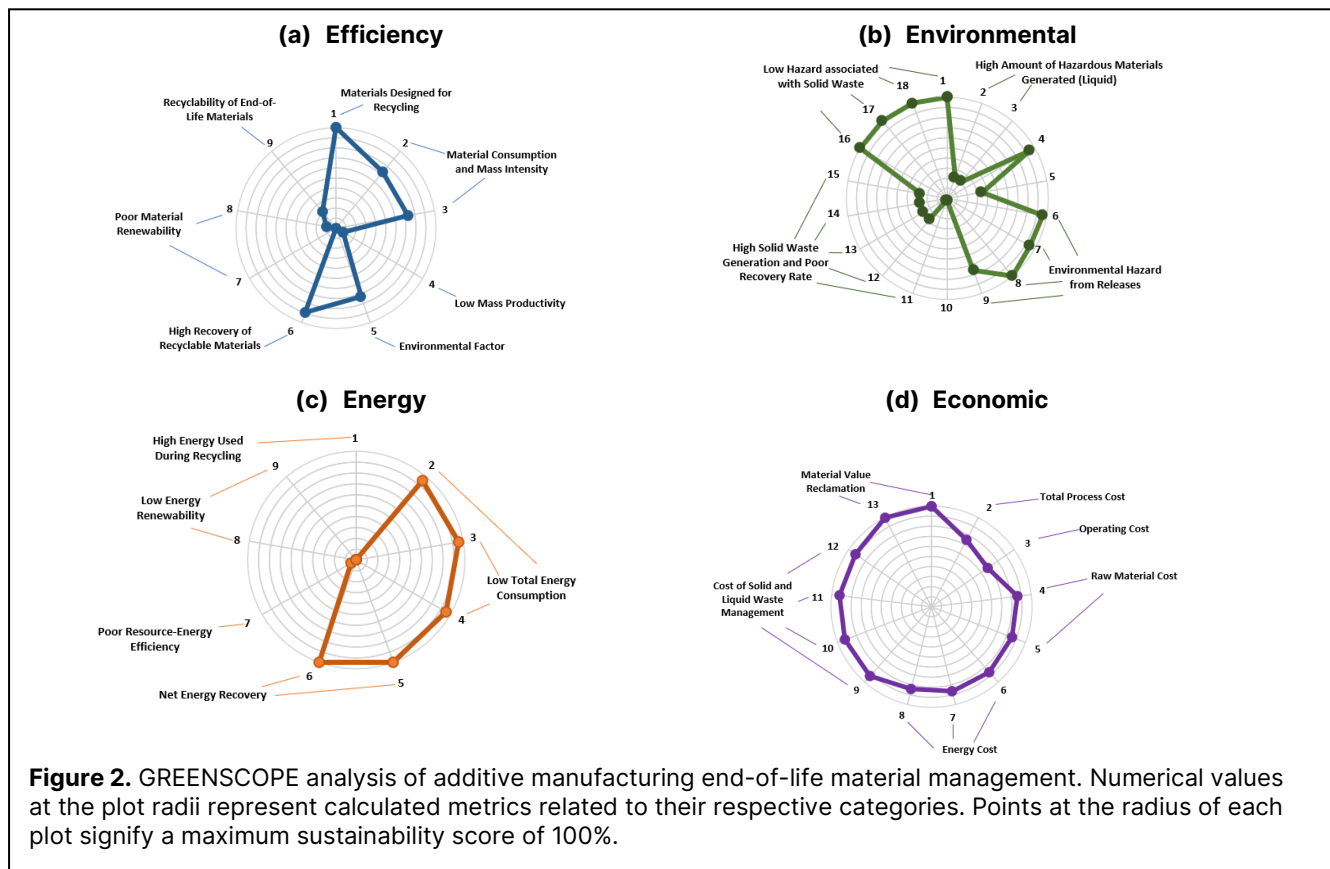
### Shifting Additive Manufacturing End-of-Life Management Toward a Circular Economy

AM techniques have been widely used to create customized products and parts from a multitude of raw material types with high efficiency [10], [11]. However, most attention toward AM challenges is centered on the manufacturing stage of the material life cycle, emphasizing releases and occupational hazards [12]–[14].

Achieving sustainability and the state of a circular economy within a given supply chain requires that the fate of all materials used is accounted for throughout the life cycle.

Presently, AM practices can result in potential releases to the environmental compartments, namely land, water, and air, and creating new exposure pathways in the EoL stage if unmitigated. The variety of materials used during the manufacturing stage creates high uncertainties on the fate of the materials. Solid materials for EoL management ultimately combine with conventional municipal solid waste (MSW) management infrastructure at the forefront of AM material management, while liquid EoL materials are treated as hazardous waste. Releases from the manufacturing stage can go unreported because improper EoL chemical management methods or cleaning practices were used or the total releases at individual operation sites are lower than the reporting thresholds.

The GREENSCOPE tool was used to find specific challenges and process modification opportunities within the EoL stage of the AM life cycle. The calculated indicators were compared against best- and worst-case scenario values to compute a sustainability percent score ranging from 0 to 100%, in which a 100% score signifies the best sustainability scenario. **Figure 2** illustrates the sustainability analysis, summarizing concerns on the low recycling rate of AM EoL materials and the inherently low



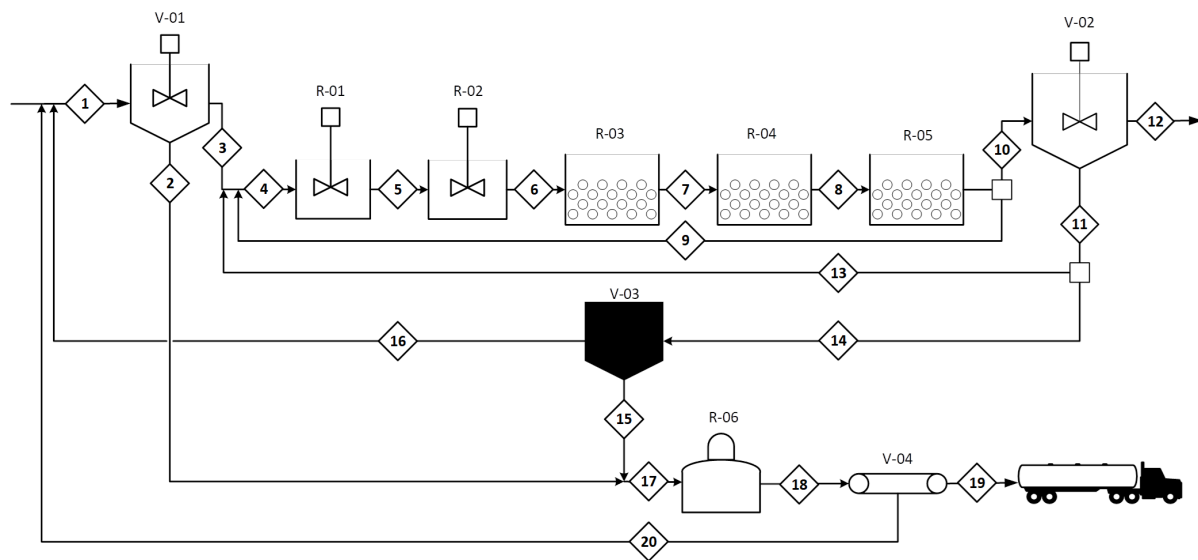


renewability. The overall AM life cycle also generates a high amount of hazardous EoL materials and their releases to the environment. The total energy consumption is relatively low, but the renewability of the energy source requires improvement. Alternatively, the various cost aspects of AM EoL management processes are close to optimality. However, variation can be expected because the management infrastructure varies between regions, in addition to the variety in material types expected from releases and the manufacturing stage.

### Tracking and Estimating Industrial Wastewater Transfers to Publicly Owned Treatment Works

The influx of industrial wastewater with conventional and emerging chemical contaminants to POTWs poses potential exposure risks, leading to environmental impacts and human health risks caused by polluted water discharges, land depositions, and air releases. Quantifying and tracking industrial chemical releases to POTWs are necessary to find unaddressed potential exposure routes. Therefore, this research effort proposes PSE approaches to tracking and estimating chemicals of concern across POTWs that foster bottom-up process modeling and chemical allocation to simulate fate and transport comprehensively.

The work introduces a POTW design, modeling, and simulation model that integrates detailed representation



**Figure 3.** A process flow diagram of a generic publicly owned treatment works (POTWs) vessels(V-0X) and reactors(R-0X). Process units include a primary clarifier (V-01), a secondary clarification unit (V-02), a sludge thickening unit (V-03), a dewatering unit (V-04), aerated and un-aerated activated sludge unit (R-01;R-05), and anaerobic digestion equipment (R-06)

**Table 1.** Methylene Chloride mass flows across streams #1 to #20 describing fate and transport

Stream	1	2	3	4	5	6	7	8	9	10
Flow Rate (m <sup>3</sup> /d)	21096	148	20948	103201	103203	103203	103203	103203	61922	41281
Mass Flow-CH <sub>2</sub> Cl <sub>2</sub> (kg/d) in Various Phases										
Liquid	104	0.7	102	112	111	110	74	49	20	13
Solid	3	1	1	24	24	24	16	10	4	3

Stream	11	12	13	14	15	16	17	18	19	20
Flow Rate (m <sup>3</sup> /d)	20641	20641	20331	310	31	279	179	179	10	169
Mass Flow-CH <sub>2</sub> Cl <sub>2</sub> (kg/d) in Various Phases										
Liquid	6.5	6.3	6.4	0.1	.01	0.1	0.6	0.4	.02	0.3
Solid	3	.003	3	.04	.04	.001	1.2	0.2	0.2	.004

of individual process units from primary clarification to advanced biological treatment, delineating settling characteristics, biological transformations, nutrient removal, and sludge management. Process units include a primary clarifier, a secondary clarification unit, an aerated and un-aerated activated sludge unit, an anaerobic digestion equipment, and an influent generator model that creates influent profiles mimicking POTW influent variations as illustrated in **Figure 3**. The model captures interactions between chemicals of concern and wastewater treatment operations, allowing for the estimation of chemical partitioning across unit operations such as sedimentation, biological treatment, adsorption, volatilization, and transport mechanisms. The selection of volatile organic compounds (VOC) as the first chemical class for model integration stems from their distinct behavior within the proposed treatment processes, characterized by their volatility, sorption, and biological uptake tendencies.

**Table 1** shows the mass flow changes and notable reductions to solid and liquid phase methylene chloride ( $\text{CH}_2\text{Cl}_2$ ) from influent to effluent, showcasing removal rates at each process unit. The substantial reduction in  $\text{CH}_2\text{Cl}_2$  concentrations, primarily attributed to biological degradation mechanisms within secondary treatment (stream # 4-10), along with the lack of adsorption capacity seen throughout the plant, aligns with real-world scenarios [15].

Our proposed PSE-based approach, employing bottom-up methodologies, offers a valuable tool for assessing the efficiency of different treatment strategies and showing potential areas for improvement in EoL chemical management. The usability and value of this model extend to its potential applications in environmental risk assessment, regulatory compliance, and the development of more effective treatment methodologies.

## CONCLUSIONS

This work describes a computational framework based on data engineering, data-driven modeling and PSE models to predict EoL chemical releases and potential occupational exposure scenarios. This contribution also shows a comprehensive EoL database created from publicly accessible domestic and international regulatory systems. Moreover, this contribution describes the application of GREENSCOPE to inform the development of more sustainable chemicals by reducing their environmental and human health impacts, implementing pollution prevention, and source reduction while increasing economic and social benefits. This multi-stakeholder analysis can be extended toward potential trade-off analysis, such as trading costs, the highest-performing category, investing in recovery processes, better waste treatment and management procedures, and selecting renewable materials sources for manufacturing. The gradual adoption of material management and selection guidelines and process improvement opportunities can allow EoL

chemical generators to take on a more active role in aiding the shift of linear economy practices toward the state of a circular economy.

Future steps are to understand the near-field exposure associated with circular economy paths and compare the composition and functionality between recycled and brand-new products to avoid unintended recycling of toxic substances at post-recycling/recovery activities, including chemicals not currently subject to USEPA regulatory programs. To resolve this problem, we can use data engineering and data-driven approaches to effectively combine facility, PAU, and equipment level data. This integration would enable the detection of potential EoL exposure scenarios, calculation of releases, and monitoring of upstream recycling scenarios for chemicals of concern. The CFA will utilize PAU/equipment-level data to estimate releases inside chemical facilities and allocate chemicals into environmental compartments. Facility-level data will provide insights on regulatory affairs, as well as the behavior of the supply chain and interconnectivity of chemicals in the EoL stage of the industry. Future research should concentrate on integrating isolated environmental regulatory data and developing modeling pipelines and architecture. These can range from graph representation learning for supply chain simulation to hybrid modeling for combining top-down (chemical facility data and models) and bottom-up (operating units, equipment) approaches.

## ACKNOWLEDGEMENTS

This research was supported in part by an appointment to the US Environmental Protection Agency (EPA) Research Participation Program administered by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between the US Department of Energy (DOE) and the US EPA. ORISE is managed by ORAU under DOE contract number DE-SC0014664. **Disclaimer:** The views expressed in this article are those of the authors and do not necessarily represent the views or policies of the EPA. Any mention of trade names, products, or services does not imply an endorsement by ORAU/ORISE, the US Government, or the EPA. The EPA does not endorse any commercial products, services, or enterprises.

## REFERENCES

1. JD Hernandez-Betancur, GJ Ruiz-Mercado, JP Abraham, M Martin, WW Ingwersen, RL Smith. Data engineering for tracking chemicals and releases at industrial end-of-life activities. *J Hazard Mater.* 405:124270 (2021). DOI: 10.1016/j.jhazmat.2020.124270.
2. JD Hernandez-Betancur, M Martin, GJ Ruiz-Mercado. A data engineering framework for on-site

- end-of-life industrial operations. *J Clean Prod* 327: 129514 (2021). DOI: 10.1016/j.jclepro.2021.129514.
3. JD Hernandez-Betancur, M Martin, GJ Ruiz-Mercado. A data engineering approach for sustainable chemical end-of-life management. *Resour Conserv Recycl* 178:106040 (2022). DOI: 10.1016/j.resconrec.2021.106040.
  4. JD Hernandez-Betancur, GJ Ruiz-Mercado, M Martin. Tracking end-of-life stage of chemicals: A scalable data-centric and chemical-centric approach. *Resour Conserv Recycl* 196: 107031 (2023a). DOI: 10.1016/j.resconrec.2023.107031.
  5. JD Hernandez-Betancur, GJ Ruiz-Mercado, M Martin. Predicting Chemical End-of-Life Scenarios Using Structure-Based Classification Models. *ACS Sustain Chem Eng* 11:3594–3602 (2023b). DOI: 10.1021/acssuschemeng.2c05662.
  6. JD Chea, GJ Ruiz-Mercado, RL Smith, DE Meyer, M Gonzalez, WM Barrett. Material Flow Analysis and Occupational Exposure Assessment in Additive Manufacturing End-of-Life Material Management. *Environ Sci Technol* XXXX:XXX:XXX (2024). DOI: doi.org/10.1021/acs.est.4c01562.
  7. S Agbleze, S Li, ET Quintanar-Orozco, GJ Ruiz-Mercado, FV Lima. Current Progress in Sustainability Evaluation, Pollution Prevention, and Source Reduction Using GREENSCOPE. In: Sustainability Engineering. Ed: 1st Editor Eric Tan CRC Press: Taylor & Francis Group (2023).
  8. GJ Ruiz-Mercado, RL Smith, MA Gonzalez. Sustainability Indicators for Chemical Processes: I. Taxonomy. *Ind Eng Chem Res* 51:2309–2328 (2012). DOI: 10.1021/ie102116e.
  9. GJ Ruiz-Mercado, RL Smith, MA Gonzalez. Sustainability Indicators for Chemical Processes: II. Data Needs. *Ind Eng Chem Res* 51: 2329–2353 (2012). DOI: 10.1021/ie200755k.
  10. R Chaudhary, P Fabbri, E Leoni, F Mazzanti, R Akbari, C Antonini. Additive manufacturing by digital light processing: a review. *Prog Addit Manuf*, 8:331–351 (2022). DOI: 10.1007/s40964-022-00336-0.
  11. TD Ngo, A Kashani, G Imbalzano, KTQ Nguyen, D Hui. Additive manufacturing (3D printing): A review of materials, methods, applications and challenges. *Compos B Eng*, 143:172–196 (2018). DOI: 10.1016/j.compositesb.2018.02.012.
  12. J Bours, B Adzima, S Gladwin, J Cabral, S Mau, Addressing Hazardous Implications of Additive Manufacturing: Complementing Life Cycle Assessment with a Framework for Evaluating Direct Human Health and Environmental Impacts: Hazard Implications of 3D Printing Materials. *J Ind Ecol*, 21:S25–S36 (2017). DOI: 10.1111/jiec.12587.
  13. R Chen, H Yin, IS Cole, S Shen, X Zhou, Y Wang, S Tang. Exposure, assessment and health hazards of particulate matter in metal additive manufacturing: A review. *Chemosphere* 259:127452 (2020). DOI: 10.1016/j.chemosphere.2020.127452.
  14. AB Stefaniak, S Du Preez, J Du Plessis. Additive Manufacturing for Occupational Hygiene: A Comprehensive Review of Processes, Emissions, & Exposures. *J Toxicol Environ Health B Crit Rev* 24: 173–222 (2021). DOI: 10.1080/10937404.2021.1936319.
  15. WJ Parker, DJ Thompson, JP Bell. Fate of Volatile Organic Compounds in Municipal Activated Sludge Plants. *Water Environ Res* 65:58–65 (1993).

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Towards the Development of Digital Twin for Pharmaceutical Manufacturing

Katherine Raudenbush<sup>a</sup>, Nikola Malinov<sup>a</sup>, Jayanth V. Reddy<sup>a</sup>, Chaoying Ding<sup>a</sup>, Huayu Tian<sup>a</sup>, and Marianthi Ierapetritou<sup>a\*</sup>

<sup>a</sup> University of Delaware, Department of Chemical & Biomolecular Engineering, Newark, DE, USA

\* Corresponding Author: [mgi@udel.edu](mailto:mgi@udel.edu).

---

## ABSTRACT

Pharma 4.0 has continued to advance as the industry develops advances in process analytical technologies, automation, and digit-ization. Digital twins which transform on-line process measurements into meaningful outputs in real-time are being developed to seize the opportunity made possible with this shift. Digital twins can be used for improved process optimization on a range of scales, from determining optimal metabolite concentrations in upstream bioreactors to considering economic and environmental impacts of process decisions. In this paper, we explore the current uses of digital twins in solid-based pharmaceutical space and the bio-pharmaceutical manufacturing. Applications cover scale up of upstream processes, product quality control, and consideration of continuous systems. We also describe the intersection of digital twins in flow sheet modeling, sensitivity analysis and optimization, and design space evaluation. Finally, areas requiring further improvement for industry adoption are addressed.

---

**Keywords:** Digital twin, Process Modeling, Pharmaceutical manufacturing, Biopharmaceutical manufacturing

## INTRODUCTION

The pharmaceutical industry has been on the rise in the U.S., dominating the global market with roughly 50% of the sales revenue being generated in the U.S. Roughly 20% of the total revenue generated is directed to research and development (R&D), to bring new drugs to market and improve the current manufacturing processes. In the last two decades, research efforts have focused on improving the predictive ability of the current manufacturing practices to increase production rates, reduce resource losses and improve the overall effectiveness of manufacturing. With advances in digitization, Internet of Things (IoT), and automation, the pharmaceutical industry is transitioning towards Industry 4.0 [1]. Known as Pharma 4.0 in the pharmaceutical industry, this transition aims to improve efficiency, productivity, responsiveness and flexibility, with varying market demands of drug products and biologics [2]. A crucial component of this transition involves the development of a digital twin framework (a digital replica of the physical manufacturing system), which has been shown to be quite effective in risk assessment and regulation-

compliant manufacturing [1].

Process systems engineering tools have been used in the development of digital twins to ensure regulation-compliant manufacturing of drug products [3]. Predictive models developed for digital replicas prove valuable in correlating equipment, process, and material parameters with the critical quality attributes (CQAs) of pharmaceutical products. This ensures that the drug products meet quality standards set by regulations, making production more efficient and flexible. Additionally, these models assist in identifying and addressing potential risks [4].

The US FDA approved the 100<sup>th</sup> mAb in 2021 with 870 more in clinical development as of 2021 [5]. Process development of these therapeutics involves several stages including cell line development, media formulation, scale up, and screening product quality [6]. These therapeutics are produced to treat various types of diseases, some of which are chronic and affect a large patient population. This leads to a drastic increase in demand for therapeutics which require large scale, robust, economically feasible manufacturing processes to produce [7]. A large number of novel, blockbuster biotherapeutics are falling off patent in this decade. This is

leading to the rise of a biosimilar market. The number of biosimilars approved has been increasing every year [8]. Biosimilar products can lead to significantly lower costs compared with the novel therapeutic [9]. Hence, the market is shifting to the most affordable biosimilar. Efficient process development enables affordability of biosimilars. The US FDA has set guidelines for Quality by Design to improve the understanding between process conditions and product quality [10]. Understanding these relationships could require large amounts of experimental data by using statistical design of experiments. Application of process models can improve process understanding, cheaply evaluate large numbers of process conditions which help to reduce the design space, improve process control, and simulate the effect of process perturbations and critical material attributes on critical quality attributes. Robust process models can help monitor the process without the need for a large number of expensive analytical measurements required in traditional process analytical techniques [11]. Hence, the development of reliable process models used within digital twins is crucial for effective process monitoring and for guiding key experiments to reduce the overall experimental expenditure. Reducing the time required to perform expensive experiments can speed up process development timelines which is crucial for the staying on the cutting edge in the fast-paced biopharma industry and succeeding in the biosimilars boom that is to come.

## MODELING COMPONENTS OF A DIGITAL TWIN

A digital twin is a digital informational construct of a physical system, created as an entity on its own and linked with the physical system [12]. It consists of a physical component, a virtual component, and automated data communications between the physical and virtual components [13]. The following sections highlight the modeling components of a digital twin for solid-based pharmaceutical and biopharmaceutical manufacturing processes [14].

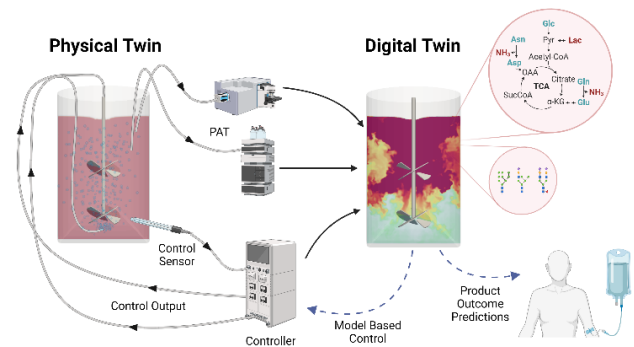
### Solid-based pharmaceutical manufacturing

In solid-based pharmaceutical manufacturing, various models have been developed to better understand the different steps involved [15]. This is not a comprehensive list but is offered as examples of such models. Loss-in-weight (LIW) feeders are used to feed ingredients in the continuous manufacturing processes. Data-driven models [16], first-order delay differential equation [17], semi-empirical models [18] and Discrete Element Models (DEM) [19] have been developed for this step. Continuous blenders blend various powdered materials to create a consistent powder mixture. DEM [20], Population Balance Model (PBM) [21], hybrid PBM-DEM [22],

compartment models [23] and residence time distribution models [24] have been utilized to simulate the blender unit. Roller compactors produce compacted ribbon products from powder materials under high pressure exerted by rotating rolls. These modeling effort includes steady state models [25] and dynamic models [26]. Data driven models, PBM and DEM are widely used for wet granulation [27, 28], which is used to achieve particle enlargement with the addition of liquid binders. For the tablet press unit, Kawakita equation, Heckel equation and their variations have been widely used [15].

### Biopharmaceutical manufacturing

In the biopharmaceutical space, digital twins have been developed to model many unit-based operations. These include bioreactor scale-up using computational fluid dynamics (CFD), product quality modeling, metabolic modeling, modeling of bioreactor physical parameters, and modeling of perfusion systems. The models of each operation may be combined to provide a digital twin of the upstream manufacturing process, illustrated in Figure 1. On-line measurements are taken continuously thanks to the advancements in process analytical technologies (PAT). The digital twin can be run alongside the physical twin, informed by continuous PAT measurements, which can inform model-based control of the bioreactor. For off-line investigations, the upstream digital twin can be integrated into flow sheet modeling, sensitivity analysis, design space identification, and process optimization, with an overarching goal of improving patient outcomes.



**Figure 1.** Upstream bioreactor cell culture performance is modeled in a digital twin representation through computational fluid dynamics, metabolic modeling, and glycosylation modelling.

### Bioreactor scale-up

A well-designed digital twin can simplify the expansion of initial processes and improve the chances of successful production runs in large-scale bioreactors from the first attempt. The inherent challenges associated with the prohibitively expensive and often impractical nature of experimental testing at large scales underscore



the critical role played by small-scale experiments and scale-up tools. These tools aid in establishing scale specific operational parameters and provide baseline expectations for yield and quality in large-scale bioreactors.

Digital twins enhance our ability to anticipate the impacts of scaling up, especially those that go beyond the capabilities of experimental measurements, such as local  $kLa$  distributions. They also provide a means to test and correct for unexpected variations observed upon process scale up.

A major concern during bioreactor scale up is the formation of heterogeneities: nutrient concentration gradients and gradients in important physical parameters, such as dissolved oxygen and carbon dioxide, temperature, pressure, and pH. Control of these parameters is key for consistent production and product quality [11]. Incorporating CFD modeling into a digital twin of the upstream process allows us to examine the behavior of throughout bioreactors at small and large scales. CFD can predict gradients across the volume of the bioreactor by capturing fluid flow, bubble dynamics, and heat and mass transfer. The prediction of gradients ahead of scale up proves invaluable for the assessment of the severity of scale up formed gradients [29]. Moreover, the predicted dynamics of cells flowing in and out of the gradients can inform scale up simulator experiments which mimic oscillations and measure impacts [30].

Creating effective digital twins for upstream biopharmaceutical manufacturing relies heavily on accurate models of metabolism and critical product quality attributes like glycosylation. There are several methods to couple CFD with metabolic modeling, though this remains a non-trivial task due to the wide range of time scales involved [31–34]. These models can be used to explore scenarios and ‘what if’ analyses, identify optimal operating conditions, and develop response plans to unexpected variations in bioreactor performance at the large scale [14]. Integration into flow sheet modeling of entire production systems can enable predictions of scale up effects on downstream processes and final product quality [35, 36]. Finally, comprehensive upstream digital twins provide a tool to better understand biological systems, an invaluable asset to biopharmaceutical research and process development.

### Product quality

N-linked glycosylation is an important Critical Quality Attribute (CQA) in the manufacturing of biotherapeutics. Changes to upstream process conditions can result in changes to the glycan fractions of the glycoprotein. This subsequently affects important properties of the product such as half-life, efficacy, immunogenicity, and aggregation [37]. Hence, it is especially important to study the upstream design space and monitor the N-linked glycosylation of the products. Measurement of this

CQA requires expensive and laborious analytics. Implementing industry 4.0 will involve development of a digital twin capable of predicting the N-linked glycosylation profiles of biologics as a function of the state (temperature, pH, media, etc.) of the bioreactor. Mechanistic models for N-linked glycosylation can be used to incorporate the effects of process parameters on N-linked glycosylation to develop a framework that improves process understanding and can simulate different operating conditions [11]. Models of N-linked glycosylation approximate the Golgi of CHO cells as continuously stirred tank reactors (CSTRs) [38] or plug flow reactors (PFRs) [39] have been developed which can provide mechanistic insights into glycosylation changes upon changes in critical process parameters (CPPs) such as bioreactor pH, temperature, and galactose addition on the N-linked glycosylation of monoclonal antibodies (mAbs). Site specific N-linked glycosylation is another layer of complexity in quality modeling. Our research has indicated that changes to bioreactor pH leads to changes in the fucosylation, galactosylation and sialylation on the mAb. Upon assessing site specific analysis, it was determined that fucosylation was only affected in the Fab region and not the Fc region. Previously developed glycosylation models are being used to explain the observed differences [38, 40]. Integrating mechanistic models of N-linked glycosylation into digital twins will expand the capabilities of digital twin predictions and enhance process simulation and decision-making capabilities.

### Cell Metabolism and Continuous Upstream Manufacturing

As the biopharmaceutical industry increases its rate of transition to continuous manufacturing, the resulting efforts to develop comprehensive digital twins will increase in parallel to the increased complexity and operational demands of continuous systems. Regarding upstream production of biologics, perfusion bioreactors and the associated cell retention modules constitute the most important unit operations, for which the required components of a digital twin as defined by Grieves and Vickers [12], are being individually developed in the literature.

Dynamic mechanistic models of host cell platform metabolism pose as strong candidates for the virtual component of a digital twin. Dynamic Metabolic Flux Analysis (DMFA) and Dynamic Flux Balance Analysis (DFBA) models are two such frameworks, which combine kinetic and stoichiometric constituents to predict intracellular metabolism as a function of bioreactor conditions with a minimal set of inputs [41–43]. These frameworks can explicitly link several of the key CPPs of perfusion related to bulk mass transfer, such as the perfusion rate, bleed rate, and media composition, to cell performance. Their mechanistic basis ensures physiologic constraints guaranteeing meaningful and interpretable predictions.

However, their curation takes time, sufficient process understanding, and requires high-quality process data. Their success is dependent on the availability of the required data, such as cell density, titer, and nutrient availability, which in turn are dependent on the existence of robust PAT for those measurements [14].

Extensive modelling work of the associated cell retention module, as either Alternating Tangential Flow Filtration (ATF) or Tangential Flow Filtration (TFF), has also been done [44-46]. The complexity of these models spans from straightforward mass-balanced based approaches, which assess product retention as a function of filter performance, to computationally expensive models leveraging CFD to interrogate mechanisms of filter fouling and the effect of flow patterns on cell viability. A critical step towards the development of robust digital twins for perfusion operation will require integration of bioreactor and cell retention modelling frameworks with tolerable computational costs.

## FLWSHEET MODELING

By integrating unit operation models with relevant information transfer, it becomes possible to construct flowsheet models for early-stage design, assessment, and decision-making processes.

In solid-based pharmaceutical manufacturing processes, flowsheet models have been developed for different production routes, including direct compaction route, dry granulation route, and wet granulation route [15] enabling systematic process analysis, including sensitivity analysis, design space identification, and optimization [47].

Flowsheet modelling has also been applied in the field of biomanufacturing. For example, Yang et al. [48] and Ding et al. [36] have developed process flowsheet to represent monoclonal antibody (mAb) production and conducted thorough analyses to compare batch and continuous operations. The primary objective was to assess the feasibility of transitioning from batch to continuous manufacturing from the economic and environmental perspective. Yang et al. [49] extended the application of flowsheet modeling to gene therapy, demonstrating the economic advantages of continuous operation in the production of recombinant adeno-associated virus (rAAV). Additionally, within the realm of plasma separation, the implementation of flowsheet modeling has enabled process scheduling and debottlenecking strategies to address potential bottlenecks, which could help enhance batch throughput and overall efficiency, and thus improve the overall process performance.

## SENSITIVITY ANALYSIS AND OPTIMIZATION

Sensitivity analysis enables the investigation of process risk based on the variability in outputs that may be caused by particular CPPs. Popular methods include Morris method and variance-based method, which have been investigated and applied to various solid based pharmaceutical manufacturing processes; CPPs that influence tablet properties have been identified such as dryer air temperature, drying time, and liquid to solid ratio [50, 51].

Utilizing a mechanistic model directly for optimizing pharmaceutical and biopharmaceutical manufacturing processes could be inefficient due to the inherent complexity, nonlinearity, and nonconvex nature of the optimization problem created by the model. A machine-learning enhanced optimization framework, specifically a surrogate-based feasibility-driven optimization algorithm, has been developed and explored to improve modelling efficiency without sacrificing model fidelity. This investigation considered various options for surrogate models, adaptive sampling acquisition functions, and strategies for distributing the sampling budget among feasibility characterization, objective function optimization, and global exploration [52-54]. This framework can be used for general deterministic or stochastic simulation-based optimization problems with black-box constraints, and it is exemplified in applications in both solid based and biopharmaceutical manufacturing [52-55]. Machine learning based optimization and surrogate modelling have the advantage of reduced computational time but also hold the potential to yield superior solutions compared to traditional optimization approaches.

## DESIGN SPACE EVALUATION

A design space is essential for defining a process's operational ranges to guarantee production rates and sufficient product quality and aid in understanding the impacts of process variables on these attributes as well. It can be interpreted as the region of the input domain where the corresponding predicted model response satisfies all the required constraints [56]. Surrogate-based feasibility analysis approaches have been proposed to solve problems with computationally expensive process models. Different surrogate models with corresponding infill criteria have been developed and tested on different example problems, including Gaussian process model [16, 57, 58], radial basis functions [59], artificial neural networks [60], support vector machines [54], and stochastic kriging [61]. The framework was applied and demonstrated to characterize the design space of direct compaction [47], roller compaction [57-59, 61], and wet granulation [60] processes in solid-based pharmaceutical manufacturing.

Models for N-linked glycosylation can be very computationally intensive as they involve solving NLPs or

PDEs [11] making it challenging to run these models for operations that require a large number of function evaluations. Surrogate-based feasibility analyses have been developed to determine design space boundaries for efficient bioreactor operations. Dynamic kriging was used to develop a surrogate model of N-linked glycosylation. This surrogate model was used to evaluate the design space of pH and temperature in an upstream bioreactor process to achieve a certain range of N-linked glycan quality attributes [62]. In the context of biopharma downstream processes, the design space was identified for a twin-column continuous Protein A chromatography column, and the impact of process variations on this design space was examined. Rather than relying on a detailed mechanistic model, this work utilized a framework of surrogate-based feasibility with adaptive sampling, aiming at striking a balance between accuracy and computational complexity [63].

## FUTURE DIRECTIONS - CHALLENGES

While the application of digital twins in the burgeoning pharmaceutical and biopharmaceutical industry has shown clear promise, there remain notable challenges. In pharmaceutical industry, one of the current challenges is enabling high accuracy of prediction (using tools such as discrete element modeling) with continuous data-transfer and model updates between predictive and physical plant. DEM being time-intensive necessitates the development of reduced order modeling tools that can retain the accuracy of the DEM models with reduced simulation times. Another challenge is integrating the impact of material properties in the digital twin. To this end, an accurate methodology for representation of material properties within the process models is imperative. This is quite challenging with the bulk material calibration efforts to obtain a unique representation of the powder material. Development of more detailed models are also needed, for example better describing the effects of electrostatics on powder processes and the subsequent impact on product quality.

Within the biopharmaceutical industry, as more sophisticated process unit operation models are developed in isolation, their integration into comprehensive flowsheet structures to enable a holistic understanding of bioprocesses presents a critical advancement. Balancing higher fidelity flowsheet models with tolerable computational costs and time will require continued innovative approaches. Successful refinement, validation, and application of flowsheet models towards sensitivity analyses and optimization will rely on the availability of the required PAT, such as recent spectroscopic-based methods, for in-line measurements of process parameters and CQAs to provide real-time analysis and corrective action.

## ACKNOWLEDGEMENTS

The authors would like to acknowledge financial support from US Food and Drug Administration (FDA) through grant U01FD006487, R01FD006588, 75F40120C00125, BAA 20-00123.

## REFERENCES

1. Arden, N.S., et al., Industry 4.0 for pharmaceutical manufacturing: Preparing for the smart factories of the future. *Int J Pharm*, 2021. **602**: p. 120554.
2. Chen, Y., et al., Digital Twins in Pharmaceutical and Biopharmaceutical Manufacturing: A Literature Review. *Processes*, 2020. **8**(9).
3. Narang, A.S. and A. Dubey, *Continuous Pharmaceutical Processing and Process Analytical Technology*. 2023: CRC Press.
4. Tian, G., et al., Quality risk assessment and mitigation of pharmaceutical continuous manufacturing using flowsheet modeling approach. *Computers & Chemical Engineering*, 2019. **129**.
5. Mullard, A., FDA approves 100th monoclonal antibody product. *Nature Reviews Drug Discovery*, 2021. **20**(7): p. 491-495.
6. Bielser, J.-M., et al., Perfusion mammalian cell culture for recombinant protein manufacturing – A critical review. *Biotechnology Advances*, 2018. **36**(4): p. 1328-1340.
7. Shukla, A.A. and J. Thömmes, Recent advances in large-scale production of monoclonal antibodies and related proteins. *Trends in Biotechnology*, 2010. **28**(5): p. 253-261.
8. Gherghescu, I. and M.B. Delgado-Charro, The Biosimilar Landscape: An Overview of Regulatory Approvals by the EMA and FDA. *Pharmaceutics*, 2020. **13**(1).
9. Pierpont, T.M., C.B. Limper, and K.L. Richards, Past, Present, and Future of Rituximab-The World's First Oncology Monoclonal Antibody Therapy. *Front Oncol*, 2018. **8**: p. 163.
10. Rathore, A.S. and H. Winkle, Quality by design for biopharmaceuticals. *Nature Biotechnology*, 2009. **27**(1): p. 26-34.
11. Reddy, J.V., et al., Cell-culture process optimization via model-based predictions of metabolism and protein glycosylation. *Biotechnology Advances*, 2023. **67**: p. 108179.
12. Grieves, M., J.J.T.p.o.c.s.N.f. Vickers, and approaches, Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. 2017: p. 85-113.
13. Kritzinger, W., et al., Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine*, 2018. **51**(11): p. 1016-1022.

14. Chen, Y., et al., Digital Twins in Pharmaceutical and Biopharmaceutical Manufacturing: A Literature Review. *Processes*, 2020. **8**(9): p. 1088.
15. Ierapetritou, M., F. Muzzio, and G. Reklaitis, Perspectives on the continuous manufacturing of powder-based pharmaceutical processes. *AIChE Journal*, 2016. **62**(6): p. 1846-1862.
16. Boukouvala, F., F.J. Muzzio, and M.G. Ierapetritou, Design Space of Pharmaceutical Processes Using Data-Driven-Based Methods. *Journal of Pharmaceutical Innovation*, 2010. **5**(3): p. 119-137.
17. Boukouvala, F., et al., An integrated approach for dynamic flowsheet modeling and sensitivity analysis of a continuous tablet manufacturing process. *Computers & Chemical Engineering*, 2012. **42**: p. 30-47.
18. Engisch, W.E. and F.J. Muzzio, Feedrate deviations caused by hopper refill of loss-in-weight feeders. *Powder Technology*, 2015. **283**: p. 389-400.
19. Bhalode, P. and M. Ierapetritou, Discrete element modeling for continuous powder feeding operation: Calibration and system analysis. *Int J Pharm*, 2020. **585**: p. 119427.
20. Gao, Y., M. Ierapetritou, and F. Muzzio, Periodic section modeling of convective continuous powder mixing processes. *AIChE Journal*, 2011. **58**(1): p. 69-78.
21. Sen, M., et al., Multi-dimensional population balance modeling and experimental validation of continuous powder mixing processes. *Chemical Engineering Science*, 2012. **80**: p. 349-360.
22. Sen, M., et al., Mathematical Development and Comparison of a Hybrid PBM-DEM Description of a Continuous Powder Mixing Process. *Journal of Powder Technology*, 2013. **2013**: p. 1-11.
23. Bhalode, P. and M. Ierapetritou, Hybrid multi-zonal compartment modeling for continuous powder blending processes. *International Journal of Pharmaceutics*, 2021. **602**: p. 120643.
24. Bhalode, P., et al., Using Residence time distribution in pharmaceutical solid dose manufacturing – a critical review. *International Journal of Pharmaceutics*, 2021: p. 121248.
25. Johanson, J.R., A Rolling Theory for Granular Solids. *Journal of Applied Mechanics*, 1965. **32**(4): p. 842-848.
26. Hsu, S.-H., G.V. Reklaitis, and V. Venkatasubramanian, Modeling and Control of Roller Compaction for Pharmaceutical Manufacturing. Part I: Process Dynamics and Control Framework. *Journal of Pharmaceutical Innovation*, 2010. **5**(1-2): p. 14-23.
27. Singh, M., et al., Challenges and opportunities in modelling wet granulation in pharmaceutical industry – A critical review. *Powder Technology*, 2022. **403**.
28. Zhao, J., G. Tian, and H. Qu, Pharmaceutical Application of Process Understanding and Optimization Techniques: A Review on the Continuous Twin-Screw Wet Granulation. *Biomedicines*, 2023. **11**(7).
29. Lara, A.R., et al., Living With Heterogeneities in Bioreactors: Understanding the Effects of Environmental Gradients on Cells. *Molecular Biotechnology*, 2006. **34**(3): p. 355-382.
30. Emmanuel Anane, I.M.K., Giles C. Wilson, Scale-down cultivation in mammalian cell bioreactors—The effect of bioreactor mixing time on the response of CHO cells to dissolved oxygen gradients. *Biochemical Engineering Journal*, 2021. **166**.
31. Farzan, P. and M.G. Ierapetritou, Integrated modeling to capture the interaction of physiology and fluid dynamics in biopharmaceutical bioreactors. *Computers & Chemical Engineering*, 2017. **97**: p. 271-282.
32. Farzan, P., B. Mistry, and M.G. Ierapetritou, Review of the important challenges and opportunities related to modeling of mammalian cell bioreactors. *AIChE Journal*, 2017. **63**(2): p. 398-408.
33. Wang, G., et al., Developing a Computational Framework To Advance Bioprocess Scale-Up. *Trends in Biotechnology*, 2020. **38**(8): p. 846-856.
34. Haringa, C., et al., Euler-Lagrange computational fluid dynamics for (bio)reactor scale down: An analysis of organism lifelines. *Engineering in Life Sciences*, 2016. **16**(7): p. 652-663.
35. Yang, O., S. Prabhu, and M. Ierapetritou, Comparison between Batch and Continuous Monoclonal Antibody Production and Economic Analysis. *Industrial & Engineering Chemistry Research*, 2019. **58**(15): p. 5851-5863.
36. Ding, C., et al., Process design of a fully integrated continuous biopharmaceutical process using economic and ecological impact assessment. *Biotechnology and Bioengineering*, 2022. **119**(12): p. 3567-3583.
37. Batra, J. and A.S. Rathore, Glycosylation of monoclonal antibody products: Current status and future prospects. *Biotechnology Progress*, 2016. **32**(5): p. 1091-1102.
38. Krambeck, F.J., et al., Model-based analysis of N-glycosylation in Chinese hamster ovary cells. *PLOS ONE*, 2017. **12**(5): p. e0175376.
39. Karst, D.J., et al., Modulation and modeling of monoclonal antibody N-linked glycosylation in mammalian cell perfusion reactors. *Biotechnology and bioengineering*, 2017. **114**(9): p. 1978-1990.
40. Jedrzejewski, P., et al., Towards Controlling the Glycoform: A Model Framework Linking

- Extracellular Metabolites to Antibody Glycosylation. *International Journal of Molecular Sciences*, 2014. **15**(3): p. 4492-4522.
41. Nolan, R.P. and K. Lee, Dynamic model of CHO cell metabolism. *Metabolic Engineering*, 2011. **13**(1): p. 108-124.
  42. Robitaille, J., J. Chen, and M. Jolicoeur, A Single Dynamic Metabolic Model Can Describe mAb Producing CHO Cell Batch and Fed-Batch Cultures on Different Culture Media. *PLOS ONE*, 2015. **10**(9): p. e0136815.
  43. Helgers, H., A. Schmidt, and J. Strube, Towards Autonomous Process Control—Digital Twin for CHO Cell-Based Antibody Manufacturing Using a Dynamic Metabolic Model. *Processes*, 2022. **10**(2): p. 316.
  44. Blaschczok, K., et al., Investigations on mechanical stress caused to CHO suspension cells by standard and single-use pumps. *Chemie Ingenieur Technik*, 2013. **85**(1-2): p. 144-152.
  45. Vickroy, B., K. Lorenz, and W. Kelly, Modeling Shear Damage to Suspended CHO Cells during Cross-Flow Filtration. *Biotechnology Progress*, 2007. **23**(1): p. 194-199.
  46. Radoniqi, F., et al., Computational fluid dynamic modeling of alternating tangential flow filtration for perfusion cell culture. *Biotechnology and Bioengineering*, 2018. **115**(11): p. 2751-2759.
  47. Wang, Z., M.S. Escotet-Espinoza, and M. Ierapetritou, Process analysis and optimization of continuous pharmaceutical manufacturing using flowsheet models. *Computers & Chemical Engineering*, 2017. **107**: p. 77-91.
  48. Yang, O., S. Prabhu, and M. Ierapetritou, Comparison between Batch and Continuous Monoclonal Antibody Production and Economic Analysis. *Ind. Eng. Chem. Res.*, 2019. **58**(15): p. 5851-5863.
  49. Yang, O., et al., Process design and comparison for batch and continuous manufacturing of recombinant adeno-associated virus. *Journal of Pharmaceutical Innovation*, 2023. **18**(1): p. 275-286.
  50. Metta, N., et al., Dynamic Flowsheet Model Development and Sensitivity Analysis of a Continuous Pharmaceutical Tablet Manufacturing Process Using the Wet Granulation Route. *Processes*, 2019. **7**(4).
  51. Chen, Y., et al., Optimization of key energy and performance metrics for drug product manufacturing. *International Journal of Pharmaceutics*, 2023. **631**: p. 122487.
  52. Boukouvala, F. and M.G. Ierapetritou, Derivative-free optimization for expensive constrained problems using a novel expected improvement objective function. *AIChE Journal*, 2014. **60**(7): p. 2462-2474.
  53. Wang, Z. and M. Ierapetritou, Constrained optimization of black-box stochastic systems using a novel feasibility enhanced Kriging-based method. *Computers & Chemical Engineering*, 2018. **118**: p. 210-223.
  54. Tian, H. and M.G. Ierapetritou, A Surrogate-Based Framework for Feasibility-Driven Optimization of Expensive Simulations. *AIChE Journal* (under review), 2023.
  55. Ding, C. and M. Ierapetritou, Machine learning-based optimization of a multi-step ion exchange chromatography for ternary protein separation. *Computers & Chemical Engineering*, 2024. **184**: p. 108642.
  56. Tian, H., et al., Feasibility/Flexibility-based optimization for process design and operations. *Computers & Chemical Engineering*, 2024. **180**.
  57. Boukouvala, F. and M.G. Ierapetritou, Feasibility analysis of black-box processes using an adaptive sampling Kriging-based method. *Computers & Chemical Engineering*, 2012. **36**: p. 358-368.
  58. Rogers, A. and M. Ierapetritou, Feasibility and flexibility analysis of black-box processes Part 1: Surrogate-based feasibility analysis. *Chemical Engineering Science*, 2015. **137**: p. 986-1004.
  59. Wang, Z. and M. Ierapetritou, A novel feasibility analysis method for black-box processes using a radial basis function adaptive sampling approach. *AIChE Journal*, 2017. **63**(2): p. 532-550.
  60. Metta, N., R. Ramachandran, and M. Ierapetritou, A novel adaptive sampling based methodology for feasible region identification of compute intensive models using artificial neural network. *AIChE Journal*, 2020. **67**(2).
  61. Wang, Z. and M. Ierapetritou, Surrogate-based feasibility analysis for black-box stochastic simulations with heteroscedastic noise. *Journal of Global Optimization*, 2018. **71**(4): p. 957-985.
  62. Yang, O. and M. Ierapetritou, mAb Production Modeling and Design Space Evaluation Including Glycosylation Process. *Processes*, 2021. **9**(2): p. 324.
  63. Ding, C. and M. Ierapetritou, A novel framework of surrogate-based feasibility analysis for establishing design space of twin-column continuous chromatography. *International Journal of Pharmaceutics*, 2021. **609**: p. 121161.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Towards 3-fold sustainability in biopharmaceutical process development and product distribution

Miriam Sarkis<sup>a,b</sup>, Steven Sachio<sup>a,b</sup>, Nilay Shah<sup>a,b</sup>, Cleo Kontoravdi<sup>a,b</sup>, Maria M. Papathanasiou<sup>a,b\*</sup>

<sup>a</sup> Sargent Centre for Process Systems Engineering, Imperial College London, SW7 2AZ, UK

<sup>b</sup> Department of Chemical Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ

\* Corresponding Author: [maria.papathanasiou11@imperial.ac.uk](mailto:maria.papathanasiou11@imperial.ac.uk).

## ABSTRACT

The (bio-)pharmaceutical industry is facing crossroads in an effort to ramp up its global capacity, while working to meet net-zero targets and to ensure continuous drug supply. Beyond geopolitical challenges faced worldwide, (bio-)pharmaceutical processes have been historically very complex to design, optimise and integrate in a global distribution network that is resilient and adaptable to changes. In this paper we offer a perspective of how Process Systems Engineering (PSE) tools can support and advance (bio-)pharma practices with an outlook towards 3-fold sustainability. The latter is considering three main pillars, namely social (drug supply), economical and environmental sustainability. We discuss PSE contributions that have revolutionised process design in this space, as well as the optimisation of distributions networks in pharmaceuticals. We do this by means of example cases: one on model-based unit operation design and a second one on sustainable supply chain networks in the space of advanced therapeutics. As such, this contribution offers a perspective on how PSE methodologies can offer a systematic way to integrate social, environmental, and economical sustainability throughout process design and product distribution.

**Keywords:** Process Design, Biosystems, Supply Chain, Machine Learning, Dynamic Modelling, Industry 4.0, Sustainability

## INTRODUCTION

As the world moves towards truly sustainable development, the process industries are re-evaluating their operations and consider the wider environmental impacts of their products. While the biopharmaceutical and life sciences sector is at the forefront of the economy (UK turnover of £81B and USA turnover of US\$285B) [1], its growth is linked to an increase in biohazardous waste, including disposable plasticware, media waste, and high volumes of purification buffers and resins. To establish profitable and resilient operations, (bio-) pharmaceutical manufacturers will need to revolutionise the current state-of-the-art. Radical changes are necessary to tackle challenges related to material-intensive research and development (R&D) and wasteful manufacturing operations. This goal has been embraced by (bio-) pharmaceutical companies, many of which have committed to ambitious net-zero targets that require a step-change to their

current modus operandi [2–6].

## (bio-) pharmaceutical challenges and future directions

Achieving resource-efficient R&D and manufacturing requires in-depth understanding of this industry that is governed by stringent regulatory constraints. Product quality and production rate are typically conflicting Key Performance Indicators (KPIs) with the former prioritized over the latter. The current, often purely experimental, approach to the identification of suitable operating conditions that satisfy quality KPIs is resource- and time-intensive and typically leads to suboptimal, inflexible processes. Within the biopharmaceutical sector, the adoption of platform processes means that upstream process operation follows fixed protocols that often overestimate nutrient requirements. The lack of end-to-end process design further exacerbates the above challenges. As the industry tends to work in silos, upstream and downstream

operations for the same manufacturing process are developed by different teams. Successes in increasing upstream product titres achieved over the past 3 decades have generally shifted the majority of operating costs downstream [7, 8]. Despite the Quality by Design initiative, development of quantitative process understanding often relies on statistical approaches, which are not generalizable and transferable across production processes. At the same time, the sparsity of available datasets or even unavailability of measurements, particularly in the case of cell-based biopharmaceuticals production, points to the need for mechanistic mathematical representations or hybrid approaches.

Against this background, the (bio-)pharmaceutical industry is currently transitioning to personalised therapeutics, which are manufactured for individuals or small cohorts of patients. An example of these are viral vectors for gene therapies [9]. Personalised therapeutics pose new challenges because they require on-demand manufacture and delivery to the clinic within constrained timeframes [10–12]. To meet patient needs, the industry needs to now consider drug product delivery to the patient within its operations envelope. In other words, manufacturing and supply chain design must be considered in tandem, with both economic and, primarily, patient-centric KPIs in mind.

### The concept of 3-fold sustainability in (bio-)pharma and the role of Process Systems Engineering (PSE)

The positive health impact of Life Sciences on the society has been, so far, outweighing the environmental footprint of the sector. This has been leading decisions around process and product development, ensuring that therapeutics meet the purity constraints and manufacturers demonstrate control over their process operation. Nonetheless, and despite scientific advances in this space, regulators report drug shortages due to batch failures as one of the most pressing challenges. At the same time, COVID-19 demonstrated that the resilience of the global (bio-)pharma network is susceptible to unforeseen events, raw material shortages, as well as rapid increase in the demand for advanced therapeutics. This is now jeopardising the overarching goal of the industry to always meet drug demand.

Drawing all objectives into the picture, one could summarise those as: (a) meeting the demand (social objective), (b) being economically efficient (economic objective) and (c) reducing the environmental footprint (environmental objective). The identification of the sweet spot that meets these three objectives requires holistic approaches that orchestrate R&D, manufacturing, and distribution decisions. In this new context, knowledge-transfer and wet-lab experimentation are no longer sufficient to advance the sector alone. The systematic use of

computational approaches to guide R&D and end-to-end process design all the way to the clinic has the potential to yield a step increase in efficiency addressing both economic and environmental sustainability targets. Specifically, the use of generalizable, first-principles mathematical models of cell metabolism [13], cell culture operation [14] and chromatographic separation [15] can complement experimental investigations to accelerate process development for the purpose of flexibility analysis, and [16] process optimisation [17]. Similarly, model-based tools that can integrate manufacturing uncertainties in the supply chain network can revolutionise the decision-making process when it comes to investment and capacity planning across the product lifecycle.

### PSE in (bio-)pharmaceutical process design and product distribution

The PSE community has a long-standing track record in developing cutting-edge methods and tools that can advance the way process and distribution networks are designed, optimised and operated. Specifically, in (bio-)pharmaceuticals, there have been contributions that investigate and propose methodologies for unit operation design, as well as end-to-end process flowsheet and optimisation, the fundamentals for many of which lie in seminal contributions in PSE [18, 19]. Indicatively, groups have proposed surrogate modelling [20–24], (adaptive) sampling [16, 22, 25, 26] and probabilistic approaches [27, 28] to map and identify a process design space. Other approaches consider variance-based methodologies to investigate the impact of design variables on the process performance and thereafter constrain the feasible space [29–31].

Similarly in operations, computer-aided tools can help assess trade-offs between KPIs during scale-up and supply chain development. Capacity planning under clinical trial and demand uncertainty pressures decision-makers to quantify cost benefits of early-stage scale up approaches, which are tied to higher initial capital investments and risk. In this space, stochastic programming and rolling horizon approaches are well-established tools in investment optimisation under uncertainty [32–35]. Network optimisation with respect to cost and environmental metrics has been explored by several contributions focusing on process industries and the pharmaceutical sectors [36–38]. These works integrate life cycle assessment (LCA) in the optimisation problem formulation, thereby quantifying environmental impacts of candidate solutions and constructing Pareto frontiers to explore cost-environmental trade-offs. In the specific context of biopharmaceuticals, pioneering work has studied the impact of single-use equipment in manufacturing of mAbs via LCA [39, 40]. There remain open questions regarding the identification of main sources of impact for emerging biopharmaceuticals and manufacturing platforms.

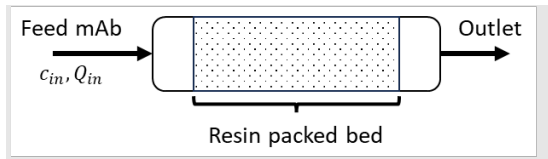
Optimisation-based PSE tools have an inherent potential to help assess trade-offs between cost and footprints and ensure therapy availability.

## EXAMPLE CASES

### Case study 1: Computer-aided design of chromatography

#### The system

Monoclonal antibodies have been the major growth driving force of the biopharmaceutical industry [41]. Their purification relies on a series of chromatographic separation steps and results in large requirements for buffers and chromatographic resins. When developing the downstream process for a new product, several commercially available resins are usually screened experimentally. Here, we demonstrate how computer-aided design space analysis can support resource-efficient resin screening as well as identifying suitable operating conditions that meet productivity and quality KPIs. Specifically, we consider the protein A affinity chromatography step (Fig. 1). The feed contains the product (mAb), as well as impurities that result from the upstream bioreactor that act as disturbances. The model describing this system is a Partial Differential and Algebraic Equation (PDAE) model and uses a general rate mass balance to describe the mass transport across the column length and radial axes [42]. The experimental validation of the model has been carried out for five different commercially available resins.



**Figure 1.** Illustrative schematic of the protein A column.

The performance of chromatographic units in bio-processing is typically assessed via two KPIs; namely yield ( $K_{YD}$  in %) (Eq.1) and productivity ( $K_{PD}$  in mg/ml/min) (Eq.2). Although the yield is constrained, productivity is usually monitored and aimed to be maximised. In this example, we introduce resin utilisation ( $K_{RU}$  in %) (Eq.3) as a third KPI. This is in an effort to design innovative separation processes that not only meet the product specifications, but also make better use of the materials towards more sustainable operation.

$$K_{YD} = 100 \frac{m_{fed} - m_{lost}}{m_{fed}} \quad (1)$$

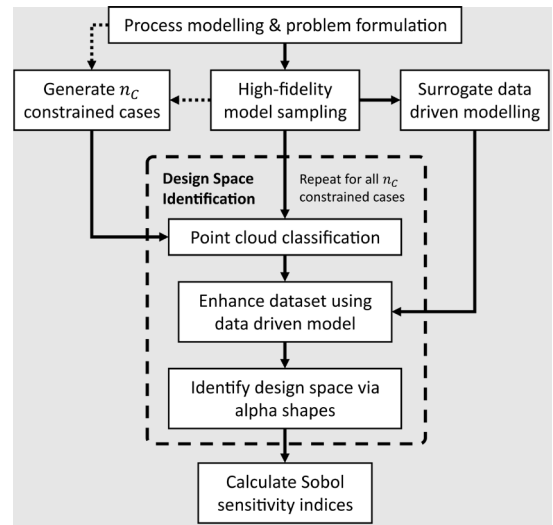
$$K_{PD} = \frac{m_{fed} - m_{lost}}{60V(1-\varepsilon_c)t} \quad (2)$$

$$K_{RU} = 100 \frac{m_{fed} - m_{lost}}{1000q_{max}V(1-\varepsilon_c)(1-\varepsilon_p)M} \quad (3)$$

where  $m_{fed}$  is the amount of mAb fed into the column,  $m_{lost}$  is the amount of mAb that leaves in the breakthrough,  $V$  volume of the column,  $\varepsilon_c$  and  $\varepsilon_b$  porosity of the column and the bed, respectively,  $t$  time,  $q_{max}$  maximum binding capacity, and  $M$  molar mass of the mAb. The productivity calculated is per volume of resin packed inside of the column, which enables comparison among the resins. For the purposes of this paper, the methodology and results are presented and discussed on the industrially relevant resin, MabSelect SuRe™.

#### Design Space Identification

To identify a feasible design space for this operation, we follow the framework in Fig. 2 presented by Satchio et al. [16].



**Figure 2.** Framework for model-based design space identification via machine learning.

#### Problem formulation

For any given process, the design space is identified as a set of points that satisfy user-defined constraints of the KPIs. For this, first, a design problem is formulated (Eq.4-6).

$$y = f(\theta) \quad (4)$$

$$\theta_L \leq \theta \leq \theta_U \quad (5)$$

$$g(y) \leq 0 \quad (6)$$

where  $y$  is the vector of monitored KPIs,  $f$  is the process model,  $\theta$  vector of the design decisions,  $\theta_L$  and  $\theta_U$  are the vector of lower and upper bound of the design decisions, respectively. While  $g$  represents the target KPI constraints that need to be satisfied. These can be upper and/or lower bound constraints on the monitored KPIs.

#### Data generation

Next, the design decisions are identified, and the high-fidelity model is sampled for the generation of a

point cloud that captures the process performance. Varying the design decisions within the bounds in Table 1, 4096 computational experiments are generated via Sobol sampling. Each combination of those decisions results into a different set of the process KPIs. Real-world processes are often highly nonlinear, therefore challenging computationally quasi-random sampling methods. For this, we integrate in the workflow, the development of an Artificial Neural Network (ANN) surrogate to be used as data interpolator. The objective of this is to increase the resolution of the search space, bypassing the computational complexity of sampling the high-fidelity model. This facilitates the identification of smooth boundaries and decreases the risk of void areas to be included in the design space.

**Table 1:** Design decisions considered and their respective bounds.

Design Decision	Lower bound	Upper bound
$c_{in}$ (mg/ml)	1.65	4.95
$Q_{in}$ (ml/min)	0.50	2.50
$T_{load}$ (min)	10	60

### Condition screening against the KPI constraints and design space identification

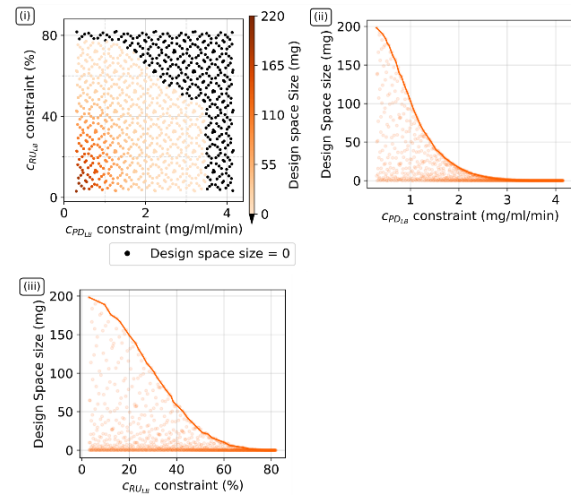
In this step, the point cloud is screened via the application of constraint combinations of the KPIs. For this problem case, process yield is strictly constrained ( $K_{YD} \geq 99\%$ ), while productivity and resin utilisation can vary between a lower and an upper bound ( $0.3 \leq K_{PD} \leq 4.2, 3 \leq K_{RU} \leq 82$ ). We generate 1024 combinations of constraints which are used to classify the original 4096 points for their feasibility to meet the constraints. Based on the density of the cloud in each case, the ANN may need to be employed for the generation of additional points such that the design space boundaries can be smoothly identified. For the design spaces generated in this case, an ANN has been used in most of the cases, generating 1-927 additional points, based on the density of each given space. The boundaries of the design spaces are then defined using alpha shapes [16].

### Resin performance and experimental validation

The 1024 identified design spaces are mapped against the two flexible constraints: productivity and resin utilisation (Fig. 3i). All the identified design spaces satisfy  $K_{YD} \geq 99\%$  yield constraint. The presented approach further allows quantification of the generated design space, whereby 0 mg is translated into absence of a feasible operating space under the given constraints. To display the density of each identified space, a colour-code is applied (Fig. 3i). On this occasion, black dots correspond to the absence of points that satisfy the given combination of constraints (0 mg space), while dark orange dots correspond to large design spaces (> 160 mg).

The size of the design space can be correlated to the operational flexibility of the process under the chosen constraints. The larger the design space (in mg), the greater the ability of the process to satisfy KPI constraints within the given bounds.

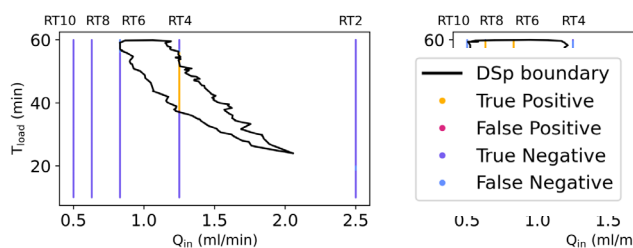
The presented approach enables manufacturers to gain insights on the process and material performance during the process development stages. For example, a trade-off between productivity and resin utilisation is observed (Fig. 3i), with the best achievable performance lying at  $K_{PD} \approx 2.8 \text{ mg/mlmin}$  and  $K_{RU} \approx 58\%$ . It is observed that the flexibility of the unit operation is inversely correlated to both productivity and resin utilisation (Fig. 3ii and 3iii). Amongst the two constraints, productivity results the most stringent one for the size of the design space.



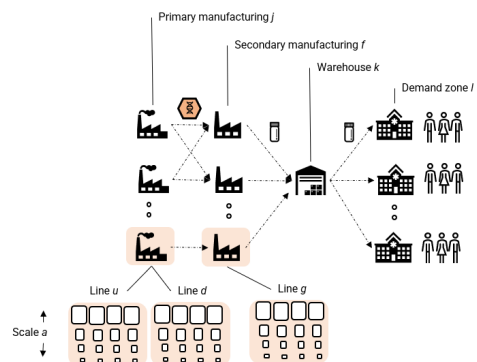
**Figure 3.** (i) 1024 design spaces generated as a function of the productivity and resin utilisation, (ii) design space-productivity constraint trade-off and (iii) design space-resin utilisation trade-off.

The generated design spaces have been validated using experimental data [42]. For the purposes of this paper, only the worst-performing scenario is displayed (Fig. 4). For this, the inlet product concentration ( $c_{in}$ ) is fixed at  $3.33 \text{ mg/mL}$  and the validation plot is presented in 2-D, where the design space is a function of the remaining two design decisions ( $Q_{in}$  and  $T_{load}$ ). It is observed that the identified design space is in good agreement with the experimental runs. RT6 is of particular interest and criticality, as the high-fidelity model and all the model-based analysis thereafter were blind to this experiment. Even in this case, the identified design space captures accurately the process performance and does not include any false positive points. This is of high importance as any consideration of false positive points within the design space boundaries can jeopardise the process performance and therefore the product specifications.

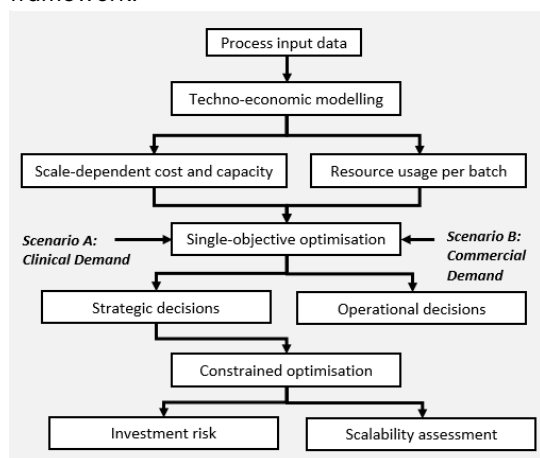




**Figure 4.** Experimental validation of the design space identified for the MabSelect SuRe™ resin.



**Figure 5.** Superstructure of supply chain optimisation framework.



**Figure 6.** Techno-economic modelling and optimisation framework.

### Remarks

The presented methodology harnesses the computational efficiency of model-based approaches to screen operating conditions and materials for their performance in an economically efficient manner. This also translates in overall better use of resources and more sustainable manufacturing platforms. The methodology can be used as an accompanying tool to accelerate process development in (bio-) pharmaceutical manufacturing and beyond. Importantly, the presented framework is adaptable and can be tailored to incorporate design decisions and KPIs of interest to the manufacturer. Critically, computational results of the example case compare favourably with wet-lab experimental data. This validation provides an

additional level of certainty that the design spaces generated with this approach can be trusted and do not run the risk of false positive operating points to be included in the design. Case study 1 is a demonstration of a PSE tool for flexible process design. This effectively assists manufacturers in the development of continuous and robust manufacturing platforms operating in a more resource efficient fashion.

## Case study 2: A sustainability assessment of advanced therapeutic supply chains

### The system

The second case study focuses on the integration of sustainability considerations in biopharmaceutical supply chains. We consider a multi-site capacity and distribution optimisation for viral vector supply chains, where nodes include upstream (USP), downstream (DSP) and fill-and-finish (F&F) (Fig.5).

### Data collection

The model [43] considers a generalized viral vector process that has been previously modeled in SuperPro Designer (Intelligen) for the calculation of batch sizes for each scale  $a$ , process times for USP, DSP, F&F, process bottleneck times and scale-dependent capital and operating costs. The techno-economic model also computes the resources ( $w$ ) consumed and emissions per batch. In this case we focus on consumption of water and electricity and CO<sub>2</sub> emissions for each process section and scale  $a$ . The analysis is conducted on different scales of primary and secondary manufacturing (50, 200, 1,000L and 2,000L bioreactor working volumes). Storage costs, capacity and electricity usage for 2 fridge types was recorded, namely MATOS PLUS Cloud 300 UF (Cloud) and MATOS PLUS Eco 300 UF (Eco) freezers. Data on distances between supply chain nodes, logistics costs, and CO<sub>2</sub> emissions per km traveled was also collected.

### The optimisation framework

The optimisation is formulated as a mixed-integer linear problem (MILP). Given the above set of scale-dependent manufacturing and logistics and a target demand, the optimisation determines network structures, selects manufacturing scale, production targets and computes associated costs and environmental footprint. The environmental footprint for each candidate design and scale assessed via the optimisation are calculated via formal LCA metrics, incorporating midpoint and endpoint categories. The environmental footprint is computed as the sum of the normalised mid-point impacts from water usage, energy usage and CO<sub>2</sub> emissions, leveraging on the normalisation factors presented in the Environmental Footprint (EF) 3.0 framework [44] (Table 2).

The problem can be solved as a single-objective optimisation with respect to each of these indicators. In this

case, we illustrate an optimisation for cost minimisation and monitor environmental footprint and production targets as output variables of the supply chain model. We generate candidate supply chain structures by considering a set of scenarios. Specifically, we illustrate how *risk-averse* decision-making leads to an optimisation for a lower demand target (Scenario A), where *risk-taking* approaches consider a larger demand target (Scenario B). The former staged-approach minimises capital investment in early stages of development under clinical trials uncertainty, whereas the latter case represents an early-scale up strategy.

**Table 2:** Impact categories and normalisation factors (NF).

Impact category	Units	NF
Climate change	kgCO <sub>2</sub> eq	$5.79 \times 10^{13}$
Water use	m <sup>3</sup> water eq	$7.91 \times 10^{13}$
Resource use, fossil-based	MJ	$4.5 \times 10^{14}$

### Scalability analysis & investment planning

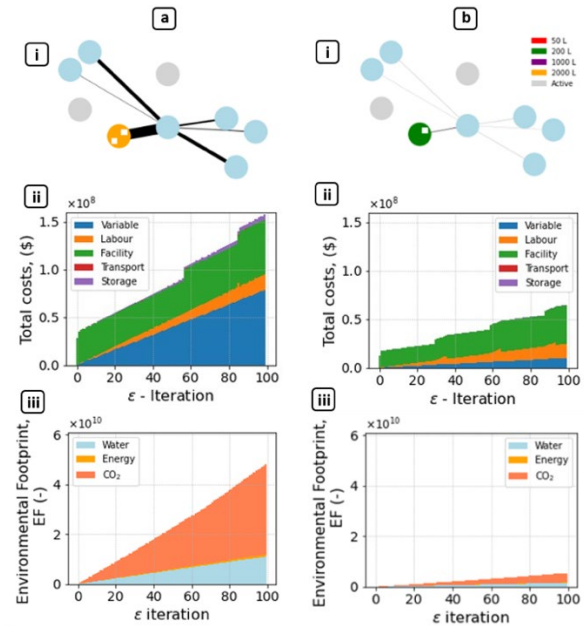
The performance of candidate network structures for Scenario A and B respectively can be assessed by constrained optimisation. Specifically, the network configuration and scales obtained via the solution of the single-objective cost minimisation can be fixed and tested for a range of demand realisations. This approach enables the assessment of the worst-case cost and emissions in the case of a demand decrease (i.e. clinical trial fails.), as a first attempt to integrate environmental sustainability in the pharmaceutical workflow. This information can also be used to identify the scalability of the candidate investment plan. The total environmental footprint and mid-point impacts of the supply chain operations are recorded.

### Cost-optimisation

The adoption of a risk-taking approach results in a centralised network consisting of 2 USP production lines, 1 DSP line and 1 F&F line. The scale selected is 2,000L, which is the largest available thereby maximising benefits from economies of scale given the target demand. At this cost-optimal point, total supply chain operating costs are computed as ~70M\$/y, with variable costs of manufacturing (~30%) and equipment expenses (~60%) being the main contributors.

Notably, storage costs from installing cheaper storage (Cloud) are minimal. This is the case for the range of demand realisations (Fig. 7a, ii). The total environmental footprint is mainly due to CO<sub>2</sub> (58-66%) emissions and water utilised during manufacturing (41-33%), whereas electricity usage in manufacturing and storage is

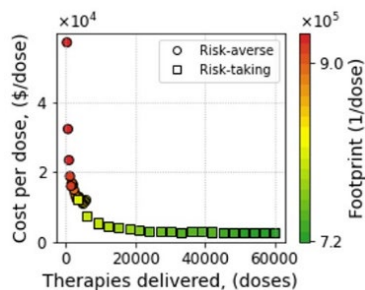
negligible (Fig. 7a, iii). A risk-averse approach results in a 200L scale being installed in USP, DSP and F&F, in a centralised manufacturing facility. The cost-optimisation results and scalability analysis highlight lower total costs and total footprint overall, with main cost drivers and environmental impacts remaining variable and facility-dependent expenses and water and CO<sub>2</sub> emissions for the range of demand realisations (Fig. 7b, iii).



**Figure 7.** Performance of candidate networks. Risk-taking (a): (i) nominal network structure, (ii) total costs, (iii) total emissions computed through scalability analysis. Risk-averse (b): (i) nominal network structure, (ii) total costs, (iii) total emissions computed through scalability analysis.  $\epsilon$  iterations correspond to the range of demand realisations from worst-case to best-case demand.

### Trade-off analysis

The *social* sustainability of the biopharmaceutical supply chain can be understood as the amount of therapy delivered. Quantifying the supply chain performance under demand uncertainty helps quantify the scalability of each investment decision. In this fashion, candidate investments can be compared with respect to cost and emissions per dose delivered and their respective scalability. The risk-averse investment results in larger emissions per dose compared to the risk-taking investment. Resource used in manufacturing are better used at larger scales in a similar fashion to economies of scale. In addition, risk-taking strategies and investing in a larger scale allows covering for up to 60,000 doses (Fig. 8). This is 10-fold larger than the risk-averse case which uses 200L manufacturing platforms.



**Figure 8.** Trade-off between financial risk and 3-fold sustainability.

The presented *what-if* analysis offers a first insight into a key trade-off between financial risk of investment and the 3-fold sustainability of the supply chain. As the therapy availability is maximised, the cost and environmental sustainability of supply chain operations is improved. This is seen by comparing Scenario A to B which result in 2 different operating scales, as well as maximising throughput of the facility once the asset is fixed. These trends highlight that a risk-taking approach in early stages of scale up can minimise emissions throughout clinical trials, although entail larger initial capital investments and risk if the clinical trial fails.

### Remarks

The proposed case study highlights that cost and environmental footprints for stainless steel facilities are not conflicting objectives. This may not be the case for other therapeutics and/or alternative equipment choices in manufacturing. We foresee more complex trade-offs by introducing in the framework the selection of equipment type for USP, DSP and F&F steps, comparing candidate manufacturing platforms relying on single-use and/or stainless-steel equipment. Each design decision would result in a range of environmental impacts from plastics disposal to water depletion and use of solvents for cleaning. In addition, the presented analysis focuses on 3 midpoint impacts within the LCA framework. The environmental footprint quantification can be augmented for a larger midpoint impact set as well as end-point aggregation. In this context, case study 2 illustrates a methodology to generate datasets for different candidate equipment technology and assess the sustainability of alternative manufacturing and network setups and its correlation to cost and service levels.

### CONCLUSIONS

As the (bio)pharmaceutical industry commits to net-zero targets, ramping up the global capacity in an economically and environmentally sustainable fashion amid a background of regulatory concerns related to drug shortages will push manufacturers to take bold decisions on

revolutionising their day-to-day operation. Adding the complex process dynamics of production systems results in a multifactorial, non-trivial problem that the sector is asked to solve in a timely manner.

In this setting, the legacy of the Process Systems Engineering (PSE) community in the development of cutting-edge methodologies, algorithms and tools becomes a vital enabler. PSE approaches are an excellent vehicle for the integration of social, economic, and environmental objectives when designing and optimising the next generation of (bio-) pharmaceutical processes and supply chains. Model-based approaches enable manufacturers to tackle challenges related to data unavailability, scale-up bottlenecks, raw material scarcity and uncertain demand profiles. Embedding such approaches in industrial workflows can shed light to novel process designs through better utilisation of the present assets or even by considering novel materials that offer improved performance. Another key advantage of PSE decision tools is the guidance towards better use of resources to reduce cost and impact per therapy/drug delivered.

To fully respond to today's needs, however, one needs to consider pathways whereby all available tools of the broad PSE portfolio are combined and used in a customised fashion. In that respect, high-fidelity process models remain a great resource that offers process insights and can help tackle challenges related to limited or complete lack of measurements. At the same time such models are highly transferrable across different products and modalities, following a suitable degree of re-parameterisation. On the other hand, Artificial Intelligence (AI) offers opportunities towards the reduction of the computational complexity of large-scale, nonlinear models, often encountered in this setting towards online deployment. Lastly, modelling manufacturing nodes dynamically and accounting for underlying uncertainties can enable the design of agile and responsive supply chain networks that meet regulatory, economic and patient related KPIs.

### ACKNOWLEDGEMENTS

The authors would like to acknowledge funding from: the UK Engineering & Physical Sciences Research Council (EPSRC) for the i-PREDICT: Integrated adaPtive pRocEss Deslgn and ConTrol (Grant reference: EP/W035006/1) and the Future Targeted Healthcare Manufacturing Hub hosted at University College London with UK university partners Grant Reference: EP/P006485/1). The authors would like to thank Prof. Blaž Likozar and Dr. Martin Grom (Department of Catalysis and Chemical Reaction Engineering, National Institute of Chemistry, Ljubljana) for expert input and data used in the design space example case. Input from Mr Alasdair Fyfe in the supply chain example case is also gratefully acknowledged.

## REFERENCES

1. Office for Life Sciences (2021) Life Science Competitiveness Indicators 2020.
2. The Paris Agreement | UNFCCC. <https://unfccc.int/process-and-meetings/the-paris-agreement>. Accessed 15 Dec 2023
3. Dukart H, Lanoue L, Rezende M, Rutten P Six new pharmaceutical industry trends | McKinsey. <https://www.mckinsey.com/capabilities/operations/our-insights/emerging-from-disruption-the-future-of-pharma-operations-strategy#>. Accessed 15 Dec 2023
4. ACS GCI Pharmaceutical Roundtable Portal » ACS GCI Pharmaceutical Roundtable Portal. <https://www.acsgcipr.org/>. Accessed 15 Dec 2023
5. Sustainability - MSD. <https://www.msd.com/company-overview/sustainability/>. Accessed 15 Dec 2023
6. GSK sets new environmental goals of net zero impact on climate and net positive impact on nature by 2030 | GSK. <https://www.gsk.com/en-gb/media/press-releases/gsk-sets-new-environmental-goals-of-net-zero-impact-on-climate-and-net-positive-impact-on-nature-by-2030/>. Accessed 15 Dec 2023
7. Alhuthali S, Kotidis P, Kontoravdi C (2021) Osmolality Effects on CHO Cell Growth, Cell Volume, Antibody Productivity and Glycosylation. *International Journal of Molecular Sciences* 2021, Vol 22, Page 3290 22:3290
8. Minimizing the Environmental Footprint of Bioprocesses - BioProcess International. <https://bioprocessintl.com/manufacturing/facility-design-engineering/minimizing-the-environmental-footprint-of-bioprocesses-303905/>. Accessed 15 Dec 2023
9. Bulcha JT, Wang Y, Ma H, Tai PWL, Gao G (2021) Viral vector platforms within the gene therapy landscape. *Signal Transduct Target Ther* 6:53
10. Levine BL, Miskin J, Wonnacott K, Keir C (2017) Global Manufacturing of CAR T Cell Therapy. *Mol Ther Methods Clin Dev* 4:92–101
11. Harrison RP, Zylberberg E, Ellison S, Levine BL (2019) Chimeric antigen receptor–T cell therapy manufacturing: modelling the effect of offshore production on aggregate cost of goods. *Cytotherapy* 21:224–233
12. Papathanasiou MM, Stamatis C, Lakelin M, Farid S, Titchener-Hooker N, Shah N (2020) Autologous CAR T-cell therapies supply chain: challenges and opportunities? *Cancer Gene Ther*. <https://doi.org/10.1038/s41417-019-0157-z>
13. Hefzi H, Ang KS, Hanscho M, et al (2016) A Consensus Genome-scale Reconstruction of Chinese Hamster Ovary Cell Metabolism. *Cell Syst* 3:434–443.e8
14. Nolan RP, Lee K (2011) Dynamic model of CHO cell metabolism. *Metab Eng* 13:108–124
15. Steinebach F, Ulmer N, Decker L, Aumann L, Morbidelli M (2017) Experimental design of a twin-column countercurrent gradient purification process. *J Chromatogr A* 1492:19–26
16. Sachio S, Kontoravdi C, Papathanasiou MM (2023) A model-based approach towards accelerated process development: A case study on chromatography. *Chemical Engineering Research and Design*. <https://doi.org/10.1016/j.chemd.2023.08.016>
17. Monteiro M, Fadda S, Kontoravdi C (2023) Towards advanced bioprocess optimization: A multiscale modelling approach. *Comput Struct Biotechnol J* 21:3639–3655
18. Grossmann IE, Calfa BA, Garcia-Herreros P (2014) Evolution of concepts and models for quantifying resiliency and flexibility of chemical processes. *Comput Chem Eng* 70:22–34
19. Bansal V, Perkins JD, Pistikopoulos EN (2002) Flexibility analysis and design using a parametric programming framework. *AIChE Journal* 48:2851–2868
20. Geremia M, Bezzo F, Ierapetritou MG (2023) A novel framework for the identification of complex feasible space. *Comput Chem Eng* 179:108427
21. Rogers A, Ierapetritou M (2015) Feasibility and flexibility analysis of black-box processes part 2: Surrogate-based flexibility analysis. *Chem Eng Sci* 137:1005–1013
22. Rogers A, Ierapetritou M (2015) Feasibility and flexibility analysis of black-box processes Part 1: Surrogate-based feasibility analysis. *Chem Eng Sci* 137:986–1004
23. Metta N, Ramachandran R, Ierapetritou M (2021) A novel adaptive sampling based methodology for feasible region identification of compute intensive models using artificial neural network. *AIChE Journal*. <https://doi.org/10.1002/AIC.17095>
24. Ding C, Ierapetritou M (2021) A novel framework of surrogate-based feasibility analysis for establishing design space of twin-column continuous chromatography. *Int J Pharm*. <https://doi.org/10.1016/j.ijpharm.2021.121161>
25. Kusumo KP, Gomoescu L, Paulen R, García Muñoz S, Pantelides CC, Shah N, Chachuat B (2020) Bayesian Approach to Probabilistic Design Space Characterization: A Nested Sampling Strategy. *Ind Eng Chem Res* 59:2396–2408
26. Zhao F, Paz Ochoa M, Grossmann IE, García-Muñoz S, Stamatis SD (2022) Novel formulations of



- flexibility index and design centering for design space definition. *Comput Chem Eng*. <https://doi.org/10.1016/J.COMPCHEMENG.2022.107969>
27. Close EJ, Salm JR, Bracewell DG, Sorensen E (2014) A model based approach for identifying robust operating conditions for industrial chromatography with process variability. *Chem Eng Sci* 116:284–295
  28. Kucherenko S, Giamalakis D, Shah N, García-Muñoz S (2020) Computationally efficient identification of probabilistic design spaces through application of metamodeling and adaptive sampling. *Comput Chem Eng* 132:106608
  29. Kotidis P, Demis P, Goey CH, Correa E, McIntosh C, Trepekli S, Shah N, Klymenko O V., Kontoravdi C (2019) Constrained global sensitivity analysis for bioprocess design space identification. *Comput Chem Eng* 125:558–568
  30. Nie L, Gao D, Jiang H, Gou J, Li L, Hu F, Guo T, Wang H, Qu H (2019) Development and Qualification of a Scale-Down Mammalian Cell Culture Model and Application in Design Space Development by Definitive Screening Design. *AAPS PharmSciTech*. <https://doi.org/10.1208/S12249-019-1451-7>
  31. Vogg S, Müller-Späth T, Morbidelli M (2020) Design space and robustness analysis of batch and counter-current frontal chromatography processes for the removal of antibody aggregates. *J Chromatogr A*. <https://doi.org/10.1016/j.chroma.2020.460943>
  32. Siganporia CC, Ghosh S, Daszkowski T, Papageorgiou LG, Farid SS (2014) Capacity planning for batch and perfusion bioprocesses across multiple biopharmaceutical facilities. *Biotechnol Prog* 30:594–606
  33. Jankauskas K, Papageorgiou LG, Farid SS (2019) Fast genetic algorithm approaches to solving discrete-time mixed integer linear programming problems of capacity planning and scheduling of biopharmaceutical manufacture. *Comput Chem Eng* 121:212–223
  34. Levis AA, Papageorgiou LG (2004) A hierarchical solution approach for multi-site capacity planning under uncertainty in the pharmaceutical industry. *Comput Chem Eng* 28:707–725
  35. Gatica G, Papageorgiou LG, Shah N (2003) Capacity Planning Under Uncertainty for the Pharmaceutical Industry. *Chemical Engineering Research and Design* 81:665–678
  36. Guillén-Gosálbez G, Grossmann IE (2009) Optimal design and planning of sustainable chemical supply chains under uncertainty. *AIChE Journal* 55:99–121
  37. Sabio N, Pozo C, Guillén-Gosálbez G, Jiménez L, Karuppiah R, Vasudevan V, Sawaya N, Farrell JT (2014) Multiobjective optimization under uncertainty of the economic and life-cycle environmental performance of industrial processes. *AIChE Journal* 60:2098–2121
  38. Elias Mota BA, Cerqueira de Sousa Gouveia Carvalho AI, Azevedo Rodrigues Gomes MI, Ferreira Dias Barbosa-Povoa AP (2020) Business strategy for sustainable development: Impact of life cycle inventory and life cycle impact assessment steps in supply chain design and planning. *Bus Strategy Environ* 29:87–117
  39. (2020) An environmental life cycle assessment comparison of single-use and conventional bioprocessing technology.
  40. Budzinski K, Constable D, D'Aquila D, Smith P, Madabhushi SR, Whiting A, Costelloe T, Collins M (2022) Streamlined life cycle assessment of single use technologies in biopharmaceutical manufacture. *N Biotechnol* 68:28–36
  41. Walsh G, Walsh E (2022) Biopharmaceutical benchmarks 2022. *Nature Biotechnology* 2022 40:12 40:1722–1760
  42. Grom M, Kozorog M, Caserman S, Pohar A, Likozar B (2018) Protein A affinity chromatography of Chinese hamster ovary (CHO) cell culture broths containing biopharmaceutical monoclonal antibody (mAb): Experiments and mechanistic transport, binding and equilibrium modeling. *J Chromatogr B Analyt Technol Biomed Life Sci* 1083:44–56
  43. Sarkis M, Tak K, Chachuat B, Shah N, Papathanasiou MM (2022) Towards Resilience in Next-Generation Vaccines and Therapeutics Supply Chains. *Computer Aided Chemical Engineering*. <https://doi.org/10.1016/B978-0-323-95879-0.50156-9>
  44. Global normalisation factors for the environmental footprint and Life Cycle Assessment - Publications Office of the EU. <https://op.europa.eu/en/publication-detail/-/publication/3ec9e2cb-f1cc-11e7-9749-01aa75ed71a1/language-en>. Accessed 15 Dec 2023

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Mining Chemical Process Information from Literature for Generative Process Design: A Perspective

Artur M. Schweidtmann<sup>a,\*</sup>

<sup>a</sup> Process Intelligence Research, Delft University of Technology, Department of Chemical Engineering, Delft, The Netherlands

\* Corresponding Author: [a.schweidtmann@tudelft.nl](mailto:a.schweidtmann@tudelft.nl).

---

## ABSTRACT

Artificial intelligence (AI) and particularly generative AI led to recent breakthroughs, e.g., in generating text and images. There is also a potential of these technologies in chemical engineering, but the lack of structured big domain-relevant data hinders advancements. I envision an open Chemical Engineering Knowledge Graph (ChemEngKG) that provides big open and linked chemical process information. In this article, I present the concept of “flowsheet mining” as the first step towards the ChemEngKG. Flowsheet mining extracts process information from flowsheets and process descriptions found in scientific literature and patents. The proposed technology requires the integration of data mining, computer vision, natural language processing, and semantic web technologies. I present the concept of flowsheet mining, discuss previous literature, and show future potentials. I believe the availability of big data will enable breakthroughs in process design through artificial intelligence.

**Keywords:** Artificial Intelligence, knowledge graph, data mining, computer vision, natural language processing

---

## INTRODUCTION

The transformation of the chemical process industry to renewable energy and feedstock supply requires the design of highly integrated, flexible, and efficient plants [1]. In the current setting, the development of chemical processes is a challenging task, which is mostly performed by manual simulation or optimization approaches that rely on hierarchical decomposition proposed in the 1980s [2]. There is a need for a paradigm shift that accelerates the development of chemical processes.

Machine learning (ML) and, more generally, artificial intelligence (AI) have great potential for chemical process design but usually require big data [3, 4]. Recent breakthroughs in ML led to success in games, computer vision, healthcare, finance, etc., even surpassing human performance in numerous tasks [5]. This great surge of AI applications often stems from the accessibility of big data, i.e., big in volume, variety, and velocity (cf., discussion on definitions of big data [6]) [7].

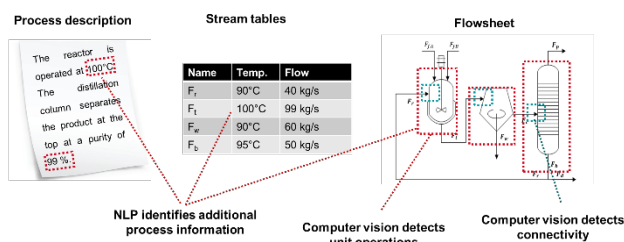
While engineers use a variety of data and knowledge to design chemical processes, most ML approaches that are used in the context of chemical process design currently do not rely on big data [8]. Instead,

most ML approaches used for process synthesis rely on regression models that are trained on manually collected datasets for specific applications [4]. For example, there exist numerous works that train surrogate models on process simulations and subsequently optimize the process design using superstructure formulations (e.g., [9–14]). Although these works frequently deal with a large number of data points (i.e., big in volume), there is a lack of variability in the data. In particular, most previous approaches do not consider a variety of process topology data from different processes but rather keep the considered topology fixed. Since ML models cannot extrapolate, these isolated approaches are limited to their specific process applications and validity domains [15].

ML methods have the potential to learn from process typologies and assist the process design in the future. Interestingly, a few pioneering works presented methods that have the potential to learn from multiple processes. Gani and coworkers extended a group contribution method to flowsheet graphs to estimate the process performance and guide decision-making [16, 17]. Also, Sahinidis and co-workers identified common patterns in flowsheet graphs [18, 19]. Recently, we also proposed new generative AI algorithms for the

autocompletion of flowsheets [20], autocorrection of flowsheets [21], and automatic prediction of control structures [22]. However, these methods have not yet unfolded their full potential because they have only been applied to datasets with few instances of flowsheet graphs and lack physical knowledge. In this article, I will focus on the methods to make more information accessible to the ML algorithms through data mining.

Scientific literature and patents provide much information about chemical engineering processes (cf. big scholarly data [23]). Flowsheets are the most important building blocks to define and communicate the structure of chemical processes [24]. As shown in **Figure 1**, they are schematic drawings describing overall process design, i.e., interconnection and type of unit operations. Depending on the development phase of a process, there exist different flowsheets with varying levels of detail (i.e., ranging from block flow diagrams (BFDs) to process flow diagrams (PFDs) and piping and instrumentation diagrams (P&IDs)) [24]. There is at least one flowsheet for every chemical process ever developed or built. These flowsheets are commonly available in PDF format in scientific publications, simulation files, patents, and company reports. Most information about industrial processes is confidential and unavailable for public research. However, in this article, I focus on flowsheet information extraction from publicly available patents and scientific publications. In the future, the proposed methods can also be used on industry data. In addition to flowsheet images, process descriptions and stream tables usually provide additional information about flow compositions, operating conditions, and sizing of the unit operations (see **Figure 1**).



**Figure 1.** Illustration of the potential for information extraction from flowsheets, process descriptions, and stream tables. Note that NLP is the abbreviation for natural language processing.

The document-centric workflow in chemical process development is inadequate. Since the existing data is unstructured, necessitating the need for chemical engineers to manually review the literature to learn about existing process designs for their specific application. Manually reviewing, verifying, and utilizing this vast amount of unstructured data is not only cumbersome but also can be inaccurate. Given the sheer number of

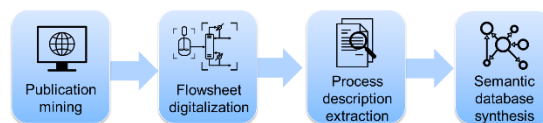
existing flowsheets, no human can comprehend all information that has been incorporated into flowsheets. Understanding how to store, structure, and link this vast amount of chemical flowsheet data and knowledge is key to further progress.

Semantic Web (SW) technologies offer functionality to connect previously isolated pieces of data and knowledge, associate meaning to them, and represent knowledge extracted from them. In particular, SW addresses data variety, by proposing graphs as a unifying data model, to which a data source can be mapped [25]. Such graphs not only contain data, but also metadata and domain knowledge (ontologies containing axioms or rules), all in the same uniform structure, and are then called knowledge graphs (KGs) (i.e., ontology + data = knowledge graph) [26, 27].

I envision that document-centric process information will be transformed into a findable, publicly accessible, interoperable, and reusable (FAIR) [28] knowledge base by representing information through a KG (cf. efforts to structure scholarly information [29]). The first step towards my vision involves the automatic extraction of information from flowsheets and process descriptions in scientific literature and patents. As illustrated in **Figure 1**, the automated extraction of information necessitates a combination of natural language processing (NLP) and computer vision techniques. In this contribution, I propose and concept of “flowsheet mining”. Moreover, I review relevant interdisciplinary literature and outline perspectives for future research.

## FLWSHEET MINING

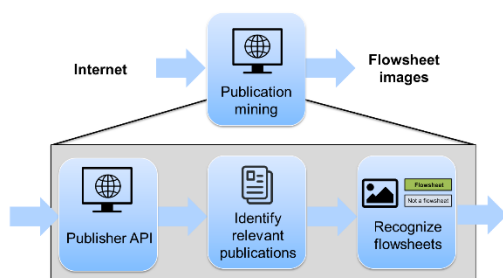
As illustrated in **Figure 2**, I propose a four-step approach for flowsheet mining. In Step 1, publications are automatically downloaded, relevant publications are identified, and flowsheet figures are extracted. In Step 2, flowsheet figures are digitized and saved in a graph format. In Step 3, information is extracted from process descriptions and stream tables. In Step 4, the extracted information is semantically enriched and saved in a knowledge graph.



**Figure 2.** Proposed four-step approach for flowsheet mining.

### Publication mining

The goal of the publication mining step is to identify relevant publications that describe a chemical process and extract flowsheet images (see **Figure 3**).



**Figure 3.** Illustration of Step 1 of the overall flowsheet mining approach: The publication mining.

The automated download of scientific publications and patents is possible through application programming interfaces (APIs). For example, CrossRef provides metadata of scientific publications through a Representational State Transfer (REST) API given a digital object identifier (DOI) [30]. This metadata includes titles, authors, license information, and links to the documents (e.g., publications in PDF or XML format), which are hosted by the corresponding publishers.

Given the large number of scientific publications, the automated identification of relevant publications is crucial. For example, crossref currently stores metadata for over 120 million records (January 2021). To identify chemical engineering publications, I propose to generate a list of all chemical engineering journals.

Only a small fraction of chemical engineering publications describe processes. Thus, the goal is to further identify chemical engineering publications that most likely contain a process flowsheet. I propose to train a topic model on the abstracts, title, and keywords of chemical engineering publications. Topic models are unsupervised ML models that can predict the topics of publications. The common latent Dirichlet allocation (LDA) for instance is a probabilistic topic model that relies on a bag-of-words approach, which means that the model considers the (tokenized) words of the passed documents disregarding their initial order [31]. To predict if an unseen publication contains a flowsheet, a classification model can be trained on the predicted topic distribution (cf. [32]).

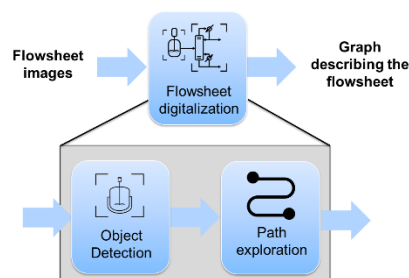
After identifying and downloading relevant publications, the flowsheet images need to be identified. This step necessitates the extraction of all images from a PDF document (e.g., using Python package PyMuPDF or PDF-Figures 2.0 [33]). Afterward, I propose to train a classification algorithm that recognizes flowsheet images. The flowsheet image recognition is a fairly simple classification problem as flowsheets are usually black-white technical drawings that follow conventions (e.g., ISO 10628). Given the success of deep convolutional neural networks (CNNs) and transfer learning in computer vision [34–36], I propose to use pre-trained state-of-the-art CNN architectures (e.g., VGG16 [37]). The results of the whole

publication mining step are the extracted images of flowsheets from scientific publications.

In our recent work, we demonstrate that flowsheet images can be recognized from literature [38]. We trained a CNN on a training set including about 1,000 PFDs and about 13,000 other images. The model showed a good overall performance with a precision of 80.7% and a recall of 94.4%. In a preliminary study, we identified about 2,500 PFDs in the journal *Computers & Chemical Engineering* which corresponds to approximately 4.5% of all images in the journal. Moreover, we identified about 2,300 PFDs in the journal *Chemical Engineering Science* and about 560 PFDs in the book *Ullmann's Encyclopedia of Industrial Chemistry*.

## Flowsheet digitization

The goal of the flowsheet digitization step is to extract the flowsheet topologies from the flowsheet images and save them in a graph format (see **Figure 4**). The digitization of chemical process flowsheets involves an object detection step [39] and a pathway exploration step. In the object detection step, a model identifies the position and type of unit operation on the flowsheet. In the pathway exploration step, the connectivity of the unit operations is explored.

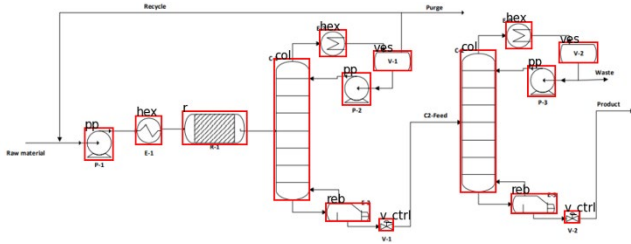


**Figure 4.** Illustration of Step 2 of the overall flowsheet mining approach: The flowsheet digitization.

Digitization of various types of engineering diagrams has been a focus of research in the computer vision domain since the 1980s, but most developed methods relied on classic computer vision methods (e.g., [40–46]). For example, Okazaki et al. [42] combined template matching and feature extraction in a hybrid model for symbol identification in circuit drawings. However, template matching can only be utilized when all drawings are identical in shape. With the emergence of deep CNNs, the field of computer vision has seen great advancements [47].

Recent literature distinguishes one stage from two-stage object detection algorithms [39]. One stage CNNs solve the tasks of (1) localizing and (2) classifying objects within a single network. Common frameworks of this network type are YOLO [48], its successors [49–51] as well as RetinaNet [52]. Single-stage networks have been

employed in the context of P&ID digitization [53–55]. However, single-stage networks do not perform as well as two-stage networks on common benchmark datasets and they often lack accuracy for the detection of small objects [50, 56].



**Figure 5.** Illustration of our preliminary object detection algorithm for the digitization of PFDs.

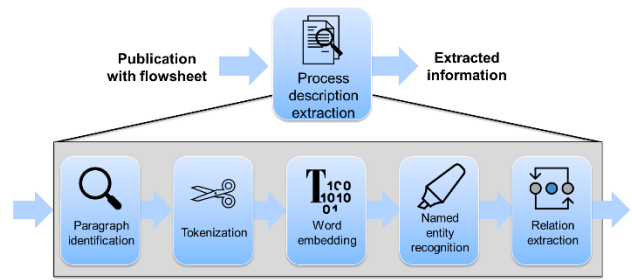
Two-stage detectors extract a feature map from the original image using a backbone network in the first stage, which is then used to find regions of interest and classify them. The most common two-stage detector is Faster R-CNN [57]. Recently, the potential of two-stage object detection for process symbols and piping detection has been demonstrated [58–60]. While there exist established algorithms for object detection, the automated exploration of the connectivity is still limited and relies on rule-based approaches [61]. The final result of the flowsheet digitization should be saved in a standardized data exchange format to enable interoperability (cf. DEXPI initiative for data exchange in the process industry [62, 63]).

In our recent work, we demonstrate that convolutional neural networks can digitize P&IDs and PFDs to a high accuracy. **Figure 5** shows the prediction of our object detection algorithm for the digitization of PFDs [60]. The figure shows the identified bounding boxes and associated object class abbreviations (e.g., hex for heat exchanger).

### Process description extraction

The goal of the process description extraction step is to extract relevant process information from process descriptions in text format.

In the NLP community, the process of automatically extracting structured information from unstructured and/or semi-structured machine-readable documents is called Information extraction (IE) [64, 65]. As illustrated in Figure 6, this includes several sub-steps that are described in the following.



**Figure 6.** Illustration of Step 3 of the overall flowsheet mining approach: The process description extraction.

The process description is usually only a small part of scientific publications. Thus, the relevant paragraphs including the process description need to be identified. This task is closely related to the identification of relevant publications (see Section Publication mining). Consequently, I propose to use topic models to identify relevant paragraphs.

Tokenization splits text into smaller tokens for further processing [66]. The simplest (rule-based) tokenizers split sentences into tokens based on white spaces, punctuations, or grammatical and syntactical rules. Tokenization is well-established in NLP and there exists a broad variety of implementations and comparisons (e.g., [67]).

In the word-embedding step of the document processing pipeline, each token is represented by a real-valued vector that can be processed by subsequent ML models. The goal is to design a mapping that represents the token's meaning as a vector. For example, the word vectors of the words “king” and “queen” should be close to each other because they have a similar meaning. Thus, this step is also referred to as representation learning.

Named entity recognition performs classification on tokens. This is necessary to identify the type of information that is provided. For example, **Figure 7** shows an illustrative example where a named entity algorithm identifies the token “heat exchanger” as a type [UNIT]. In the relation extraction step, possible relationships between entities are identified (i.e., green arrows in **Figure 7**).



**Figure 7.** Illustration of tokenization, named entity recognition and relation extraction.

Relation extraction can also involve relation classification, which is typically formulated as a classification problem to classify the relationship between the entities identified in the text [68]. A classifier takes the contextualized representation of two or more entities (e.g., words) as inputs and predicts possible relations between the entities as output. Feature- or kernel-based methods such as the conditional random field are earlier



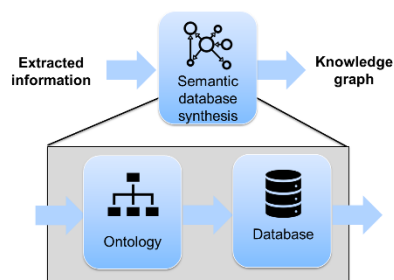
approaches used for named entity recognition and relation extraction in a pipeline setting [68]. However, the performance of traditional approaches heavily depends on manual feature engineering, which requires domain-specific knowledge and a deep understanding of linguistics [68]. Also, several deep learning architectures have been proposed for entity recognition and relation extraction, which are mostly recurrent types of neural networks based on Bidirectional Long Short Term Memory (BiLSTM) or Gated Recurrent Unit cells [69, 70]. These models are often available in NLP tools (e.g., spaCy, Stanford CoreNLP, AllenNLP, and IBM Watson Natural Language Understanding).

Recently, transformer language models [71] have become the de-facto standard for representation learning in NLP allowing for domain-specific transfer learning. Compared to transformer models that benefit from abundant knowledge from pre-training and strong feature extraction capability, approaches based on BiLSTM have shown a lower generalization performance [68] and are less efficient in capturing long-distance context (due to vanishing gradients in the training process) [65]. Specifically, Bidirectional Encoder Representations from Transformers (BERT) [72] is a common language model that utilizes bidirectional attention mechanism and large-scale unsupervised corpora to obtain effective context-sensitive representations of each word in a sentence [68]. Moreover, various improved variants of BERT have been proposed for various downstream NLP tasks. For example, BERT-based approaches have been applied to scientific texts (e.g., SciBERT [73]) and have even been adapted for molecular property prediction (e.g., ChemBERTa [74]). Recently, large language models (LLMs) such as ChatGPT have also solved various IE tasks through prompting.

The majority of traditional IE approaches are focused on informal text (e.g., social media texts), or biomedical text (e.g., PubMedInfo Crawler [75]), while only a little literature on chemical engineering-related text exists (e.g., chemical patents [76]). In particular, very little attention is paid to IE for chemical process design. Xu et al. [77] showed that bio-entity extraction based on BioBERT [78] can significantly outperform other methods. Thus, a great potential for transformer-based language models on chemical engineering entity extraction tasks can be expected. However, relation extraction is highly domain-specific, due to the variety of underlying ontologies and relation types. To overcome this issue, approaches that combine LLMs, information retrieval, and SW technologies are promising (cf. [79]).

### Semantic database synthesis

The goal of the semantic database synthesis step is to build a knowledge graph by integration of the extracted information (see **Figure 8**).



**Figure 8.** Illustration of Step 4 of the overall flowsheet mining approach: The semantic database synthesis.

A KG is an effective means for capturing and structuring a large amount of multi-relational data from various disciplines [26]. KGs are typically represented in the standardized Resource Description Framework (RDF) data model and queried using the SPARQL Protocol and RDF Query Language. It can be formally defined as  $G = \{E, R, T\}$ , where  $G$  is a labeled and directed multi-graph, and  $E, R, T$  are the sets of entities, relations, and triples, respectively. Each triple is formalized as  $(u, e, v) \in T$ , where  $u \in E$  is the head node,  $v \in E$  is the tail node, and  $e \in R$  is the edge connecting  $u$  and  $v$  [27]. KGs are usually stored in a graph database system (e.g., Neo4j or virtuoso).

Ontologies are semantic data models that define the types of things that exist in a domain and the properties that can be used to describe them, including the relationships between them [27]. The main components of an ontology are individuals (instances or objects) of classes (distinct types of concepts that exist in the data), relationships (ways to relate classes and individuals), and attributes (aspects, properties, features, characteristics that objects and classes can have). RDF Schema (RDFS) and OWL (Web Ontology Language) are two different language models that enable the construction of quantified statements in the form of RDF graphs. Currently, a few ontologies have been developed for process systems [80, 81], process safety [82, 83], process operation [84–86], and kinetic reactions [87, 88]. However, these ontologies are mostly generalized data models that have not yet been used in conjunction with big data to form a KG. Currently, there exists only one publicly available chemical engineering KG, which creates a digital twin of the eco-industrial park on Jurong Island, Singapore [89]. To build the process flowsheet KG, I recommend extending the ONTOCAPE ontology [80, 81]. Furthermore, I recommend linking or integrating flowsheet information into existing open scholarly KGs, e.g., the open research knowledge graph (ORKG) [90].

## CONCLUSIONS

In the current setting, information about chemical



processes is often not easily findable, accessible, interoperable, and reusable (FAIR). In particular, process design information is depicted in flowsheets and described in text format (e.g., in scientific literature, patents, or company reports). This lack of structured data is a major hurdle for the development of processes. Also, the lack of structured data is a bottleneck for process design through AI algorithms.

This article describes the steps of a larger ongoing research agenda that aims to build an open Chemical Engineering Knowledge Graph (ChemEngKG). I envision that all (publicly available) chemical processes will be accessible in the ChemEngKG. To achieve this goal, I proposed the concept of “flowsheet mining” that will enable the autonomous extraction of chemical process information from scientific literature. The proposed concept mines scientific literature and patents, identifies and digitizes flowsheet images, extracts information from process descriptions, and saves all information in a semantic database. Thus, flowsheet mining requires the integration and further development of algorithms from several domains. I review relevant interdisciplinary literature in data mining, computer vision, natural language processing, and semantic web technologies. Moreover, I highlighted the potentials for the development of algorithms that are relevant for flowsheet mining.

Finally, I envision that the ChemEngKG will be an enabler technology for innovative generative AI algorithms facilitating chemical process design. There exist a multitude of powerful ML algorithms that have already demonstrated breakthrough results in other domains such as molecular design. I believe that these methods have also a great potential for AI-assisted process design. First methods already show the potential of AI for the auto-completion, autocorrection, and auto-generation of flowsheets. In the future, further development is needed to integrate engineering knowledge, big data, and AI for process design.

## ACKNOWLEDGEMENTS

The author acknowledges NWO funding for the Veni talent programme.

## REFERENCES

1. A. Mitsos, N. Asprion, C. A. Floudas, M. Bortz, M. Baldea, D. Bonvin, A. Caspari, and P. Schäfer, *Comput Chem Eng* 113:, 209–221 (2018)
2. D. W. Green and M. Z. Southard, *Perry's chemical engineers' handbook*, McGraw-Hill Education (2019)
3. V. Venkatasubramanian, *AIChE Journal* 65:, 466–478 (2019)
4. J. H. Lee, J. Shin, and M. J. Realff, *Comput Chem Eng* 114:, 111–121 (2018)
5. Y. LeCun, Y. Bengio, and G. Hinton, *Nature* 521:, 436–444 (2015)
6. A. De Mauro, M. Greco, and M. Grimaldi, *Library Review* (2016)
7. L. Zhou, S. Pan, J. Wang, and A. V Vasilakos, *Neurocomputing* 237:, 350–361 (2017)
8. M. Wiedau, G. Tolksdorf, J. Oeing, and N. Kockmann, *Chemie Ingenieur Technik* 93:, 2105–2115 (2021)
9. C. A. Henao and C. T. Maravelias, *AIChE Journal* 57:, 1216–1232 (2011)
10. I. Fahmi and S. Cremaschi, *Comput Chem Eng* 46:, 105–123 (2012)
11. M. Jones, H. Forero-Hernandez, A. Zubov, B. Sarup, and G. Sin, Superstructure optimization of oleochemical processes with surrogate models, in *Computer Aided Chemical Engineering*, Elsevier (2018), pp. 277–282
12. H. A. Pedrozo, S. B. R. Reartes, Q. Chen, M. S. D'Áz, and I. E. Grossmann, *Comput Chem Eng* 141:, 107015 (2020)
13. W. R. Huster, A. M. Schweidtmann, and A. Mitsos, *Optimization and Engineering* 21:, 517–536 (2020)
14. D. Rall, A. M. Schweidtmann, B. M. Aumeier, J. Kamp, J. Karwe, K. Ostendorf, A. Mitsos, and M. Wessling, *J Memb Sci* 600:, 117860 (2020)
15. A. M. Schweidtmann, J. M. Weber, C. Wende, L. Netze, and A. Mitsos, *Optimization and Engineering* 1–22 (2021)
16. L. d'Anterrosches and R. Gani, *Fluid Phase Equilib* 228:, 141–146 (2005)
17. A. K. Tula, M. R. Eden, and R. Gani, *Comput Chem Eng* 81:, 245–259 (2015)
18. T. Zhang, N. V Sahinidis, and J. J. Siirola, *AIChE Journal* 65:, 592–603 (2019)
19. C. Zheng, X. Chen, T. Zhang, N. V Sahinidis, and J. J. Siirola, *Comput Chem Eng* 107676 (2022)
20. G. Vogel, L. Schulze Balhorn, and A. M. Schweidtmann, *Comput Chem Eng* 171:, 108162 (2023)
21. L. S. Balhorn, M. Caballero, and A. M. Schweidtmann, Toward autocorrection of chemical process flowsheets using large language models, (2023)
22. E. Hirtreiter, L. Schulze Balhorn, and A. M. Schweidtmann, *AIChE Journal* n/a:, e18259
23. F. Xia, W. Wang, T. M. Bekele, and H. Liu, *IEEE Trans Big Data* 3:, 18–35 (2017)
24. G. Nasby, *Chem Eng Prog* 108:, 36–44 (2012)
25. P. Hitzler, M. Kröttsch, and S. Rudolph, *Foundations of semantic web technologies*, CRC press (2009)
26. X. Wilcke, P. Bloem, and V. De Boer, *Data Science* 1:, 39–57 (2017)
27. A. Hogan, E. Blomqvist, M. Cochez, C. d'Amato, G.

- de Melo, C. Gutierrez, J. E. L. Gayo, S. Kirrane, S. Neumaier, A. Polleres, and others, *arXiv preprint arXiv:2003.02320* (2020)
28. M. D. Wilkinson, M. Dumontier, Ij. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, and others, *Sci Data* 3:, 1–9 (2016)
  29. S. Auer, V. Kovtun, M. Prinz, A. Kasprzik, M. Stocker, and M. E. Vidal, Towards a knowledge graph for science, (2018), pp. 1–6
  30. R. Lammey, *Insights* 28: (2015)
  31. D. M. Blei, A. Y. Ng, and M. I. Jordan, *Journal of machine Learning research* 3:, 993–1022 (2003)
  32. M. Pavlinek and V. Podgorelec, *Expert Syst Appl* 80:, 83–93 (2017)
  33. C. Clark and S. Divvala, PDFFigures 2.0: Mining figures from research papers, (2016), pp. 143–152
  34. A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, *Comput Intell Neurosci* 2018: (2018)
  35. A. Khan, A. Sohail, U. Zahoor, and A. S. Qureshi, *Artif Intell Rev* 53:, 5455–5516 (2020)
  36. W. Rawat and Z. Wang, *Neural Comput* 29:, 2352–2449 (2017)
  37. K. Simonyan and A. Zisserman, *arXiv preprint arXiv:1409.1556* (2014)
  38. L. Schulze Balhorn, Q. Gao, D. Goldstein, and A. M. Schweidtmann, Flowsheet Recognition using Deep Convolutional Neural Networks, in *In Press: Computer Aided Chemical Engineering*, Elsevier (2022)
  39. Z.-Q. Zhao, P. Zheng, S. Xu, and X. Wu, *IEEE Trans Neural Netw Learn Syst* 30:, 3212–3232 (2019)
  40. H. Bunke, Automatic Interpretation of Lines and Text in Circuit Diagrams, (1982), pp. 297–310
  41. C. Howie, J. Kunz, T. Binford, T. Chen, and K. H. Law, *Advances in Engineering Software* 29:, 563–570 (1998)
  42. A. Okazaki, T. Kondo, S. Tsunekawa, and E. Kawamoto, *IEEE Trans Pattern Anal Mach Intell* 10:, 331–341 (1988)
  43. F. C. A. Groen, A. C. Sanderson, and J. F. Schlag, *Pattern Recognit Lett* 3:, 343–350 (1985)
  44. C.-S. Fahn, J.-F. Wang, and J.-Y. Lee, *Comput Vis Graph Image Process* 44:, 119–138 (1988)
  45. M. K. Gellaboina and V. G. Venkoparao, Graphic Symbol Recognition Using Auto Associative Neural Network Model, (2009), pp. 297–301doi:10.1109/ICAPR.2009.45
  46. C. F. Moreno-García, E. Elyan, and C. Jayne, *Neural Comput Appl* 31:, 1695–1712 (2019)
  47. A. Krizhevsky, I. Sutskever, and G. E. Hinton, *Adv Neural Inf Process Syst* 25:, 1097–1105 (2012)
  48. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, You only look once: Unified, real-time object detection, (2016), pp. 779–788
  49. J. Redmon and A. Farhadi, YOLO9000: Better, Faster, Stronger at <https://arxiv.org/pdf/1612.08242>
  50. A. Farhadi and J. Redmon, Computer Vision and Pattern Recognition, cite as (2018)
  51. A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, YOLOv4: Optimal Speed and Accuracy of Object Detection at <https://arxiv.org/pdf/2004.10934>
  52. T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, Focal loss for dense object detection, (2017), pp. 2980–2988
  53. B. Sezen, *Evaluation of Machine Learning Algorithms for Object Detection in Technical Drawings like P&IDs and Circuit Diagrams*, Technical University of Munich Press (2019) at <https://www.researchgate.net/publication/340307540\_Evaluation\_of\_Machine\_Learning\_Algorithms\_for\_Object\_Detection\_in\_Technical\_Drawings\_like\_PIDs\_and\_Circuit\_Diagrams>
  54. E. Elyan, L. Jamieson, and A. Ali-Gombe, *Neural Netw* 129:, 91–102 (2020)
  55. Yu, Cha, Lee, Kim, and Mun, *Energies (Basel)* 12:, 4425 (2019)
  56. T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, Feature Pyramid Networks for Object Detection at <https://arxiv.org/pdf/1612.03144>
  57. S. Ren, K. He, R. Girshick, and J. Sun, *IEEE Trans Pattern Anal Mach Intell* 39:, 1137–1149 (2016)
  58. W. Gao, Y. Zhao, and C. Smidts, *Progress in Nuclear Energy* 128:, 103491 (2020)
  59. D.-Y. Yun, S.-K. Seo, U. Zahid, and C.-J. Lee, *Applied Sciences* 10:, 4005 (2020)
  60. M. F. Theisen, K. N. Flores, L. Schulze Balhorn, and A. M. Schweidtmann, *Digital Chemical Engineering* 6:, 100072 (2023)
  61. S. Mani, M. A. Haddad, D. Constantini, W. Douhard, Q. Li, and L. Poirier, Automatic digitization of engineering diagrams using deep learning and graph search, (2020), pp. 176–177
  62. M. Wiedau, L. von Wedel, H. Temmen, R. Welke, and N. Papakonstantinou, *Chemie Ingenieur Technik* 91:, 240–255 (2019)
  63. S. Fillinger, H. Bonart, W. Welscher, E. Esche, and J.-U. Repke, *Chemie Ingenieur Technik* 89:, 1454–1463 (2017)
  64. E. D. Liddy, *Encyclopedia of Library and Information Science, 2nd Ed* (2001) at <http://surface.syr.edu/cgi/viewcontent.cgi?article=1019&context=cnlp>
  65. D. Jurafsky and J. H. Martin, *Speech and language processing, second edition*, Harlow, Pearson Education (2014)
  66. J. J. Webster and C. Kit, Tokenization as the initial

- phase in NLP, (1992)
67. Y. He and M. Kayaalp, Bethesda, MD: The Lister Hill National Center for Biomedical Communications 48: (2006)
  68. K. Xue, Y. Zhou, Z. Ma, T. Ruan, H. Zhang, and P. He, Fine-tuning BERT for joint entity and relation extraction in chinese medical text, (2019), pp. 892–897
  69. S. Zheng, F. Wang, H. Bao, Y. Hao, P. Zhou, and B. Xu, *arXiv preprint arXiv:1706.05075* (2017)
  70. M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer, *arXiv preprint arXiv:1802.05365* (2018)
  71. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, Attention is all you need, (2017), pp. 5998–6008
  72. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, *arXiv preprint arXiv:1810.04805* (2018)
  73. I. Beltagy, K. Lo, and A. Cohan, SciBERT: Pretrained Language Model for Scientific Text, (2019)
  74. S. Chithrananda, G. Grand, and B. Ramsundar, *arXiv preprint arXiv:2010.09885* (2020)
  75. A. Kanwal, S. Fazal, A. I. Bhatti, M. Ullah, and M. A. Khalid, *Meta Gene* 20:, 100550 (2019)
  76. C. Sun, Z. Yang, L. Wang, Y. Zhang, H. Lin, and J. Wang, *arXiv preprint arXiv:2009.01560* (2020)
  77. J. Xu, S. Kim, M. Song, M. Jeong, D. Kim, J. Kang, J. F. Rousseau, X. Li, W. Xu, V. I. Torvik, and others, *arXiv preprint arXiv:2005.04308* (2020)
  78. J. Lee, W. Yoon, S. Kim, D. Kim, S. Kim, C. H. So, and J. Kang, *Bioinformatics* 36:, 1234–1240 (2020)
  79. R. Anantharangachar, S. Ramani, and S. Rajagopalan, *arXiv preprint arXiv:1302.1335* (2013)
  80. J. Morbach, A. Yang, and W. Marquardt, *Eng Appl Artif Intell* 20:, 147–161 (2007)
  81. W. Marquardt, J. Morbach, A. Wiesner, and A. Yang, *OntoCAPE: A Re-Usable Ontology for Chemical Process Engineering*, Springer Publishing Company (2010)
  82. S. Natarajan, K. Ghosh, and R. Srinivasan, *Comput Chem Eng* 46:, 124–140 (2012)
  83. M. Rodriguez and J. Lagua, *Chem Eng Trans* 77:, 67–72 (2019)
  84. R. Batres and Y. Naka, Process plant ontologies based on a multi-dimensional framework, (2000), pp. 433–437
  85. R. Batres, A. Aoyama, and Y. Naka, *Comput Chem Eng* 26:, 487–498 (2002)
  86. E. Muñoz, A. Espuña, and L. Puigjaner, *Comput Chem Eng* 34:, 668–682 (2010)
  87. F. Farazi, J. Akroyd, S. Mosbach, P. Buerger, D. Nurkowski, M. Salamanca, and M. Kraft, *J Chem Inf Model* 60:, 108–120 (2019)
  88. F. Farazi, M. Salamanca, S. Mosbach, A. Eibeck, L. K. Aditya, A. Chadzynski, K. Pan, X. Zhou, S. Zhang, and others, *ACS Omega* 5:, 18342–18348 (2020)
  89. A. Eibeck, M. Q. Lim, and M. Kraft, *Comput Chem Eng* 131:, 106586 (2019)
  90. M. Y. Jaradeh, A. Oelen, K. E. Farfar, M. Prinz, J. D'Souza, G. Kismihók, M. Stocker, and S. Auer, Open research knowledge graph: next generation infrastructure for semantic scholarly knowledge, (2019), pp. 243–246

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# From Process to Systems Design: A Perspective on the Future of Design Education

Victor M. Zavala<sup>a\*</sup>

<sup>a</sup> Department of Chemical and Biological Engineering, University of Wisconsin-Madison

\* Corresponding Author: [victor.zavala@wisc.edu](mailto:victor.zavala@wisc.edu)

---

## ABSTRACT

Chemical engineers are natural “systems-thinkers”; this is a skill that allows us to analyze highly complex processes that involve heterogeneous components, phenomena, and scales. Systems-thinking skills are fostered in the chemical engineering curriculum via integrative and project-based courses, such as process/product design and laboratories. However, existing curricula tends to focus scope to product/process boundaries, offering limited opportunities to capture connections to behavior occurring at small scales (e.g., atomistic and molecular) and at large scales (e.g., supply chains, policy, markets, and infrastructures). This limit in scope can hinder our ability to appreciate how products/processes that we develop impact society, markets, and the environment (e.g., the opioid addiction crisis, environmental impacts of forever chemicals and chemical fertilizers, and electricity markets). This limit in scope can also hinder our ability to appreciate how emerging tools from the molecular sciences can help us design better products/processes. Expanding the boundaries of our thinking is essential in overcoming these limitations. In this perspective, I discuss how emerging concepts and technologies from machine learning, data science, environmental sciences, molecular simulations, and mathematics provide powerful tools to help foster systems-thinking over a broad set of scales and to help establish connections with non-traditional disciplines (e.g., social sciences). In addition, I discuss the need to create new conceptual frameworks, case studies, and software that can help foster systems-thinking.

**Keywords:** Education, Chemical Engineering, Design, Systems Engineering

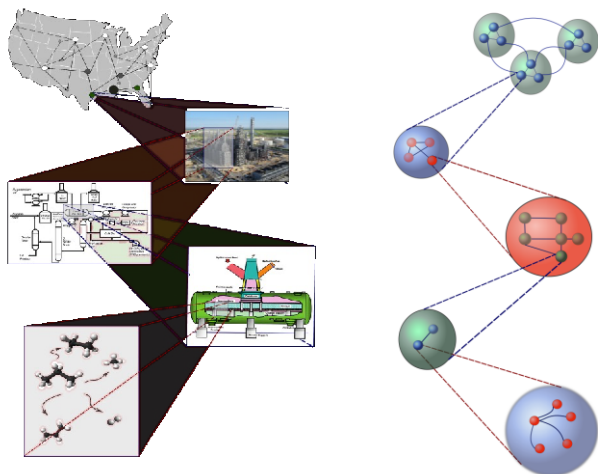
## INTRODUCTION

A “system” is defined as a collection of interconnected components; these components can be *heterogeneous* (have different functionalities) and can interact in a *hierarchical* manner (over multiple spatial and temporal scales). A chemical process is a *complex system*; chemical processes involve unit operations that carry out distinct functionalities (e.g., mixing, reaction, separation) and that are tightly interconnected. A unit operation (e.g., a reactor) is also a complex system that integrates components (e.g., reactants, inerts, catalyst, cooling/heating system) to form diverse products/byproducts. A catalyst is also a complex system that integrates components (e.g., promoter, support, and active materials) at a molecular level to facilitate reactions. A chemical process can be seen as a component of a larger system (e.g., a refinery), which is in turn a component of an even larger

system (e.g., a supply chain). This supply chain interacts with other complex systems (e.g., the environment, the power grid, and energy markets). This highlights how systems exhibit hierarchical structure that links behavior over multiple scales, see **Figure 1**. For instance, the disruption of a chemical process (e.g., due to weather) can collapse an entire supply chain; similarly, the design of a catalyst can impact the performance of a chemical process (e.g., energy use due to more difficult separation of products/byproducts).

The definition of a system is powerful in that it is *abstract* and thus general: *everything* is technically a system (e.g., the cell, the human body, human society). Having an abstract way of thinking is critical, as this facilitates transferring of knowledge across disciplines. For instance, chemical engineers leverage systems-thinking skills when applying principles of thermodynamics for analyzing systems that are outside our “traditional” domain

of expertise. This is one of the reasons why chemical engineers have been impactful in a broad range of domains (e.g., biomedical, electrochemical, materials, environmental, pharmaceutical).



**Figure 1.** Hierarchical graph representation of a complex system involving multiple scales.

I would argue that “systems-thinking” has been at the core of chemical engineering since our profession was created (it is in our DNA). Let’s recall that chemical engineering was created as a branch of mechanical engineering (an already established profession), with the goal of designing manufacturing processes for chemicals. This required, *from the get-go*, a type of engineer that could fuse concepts from diverse disciplines. Modern chemical engineers combine concepts of transport, thermodynamics, conservation, chemistry, chemical kinetics, mathematics, and so on for analyzing and designing products and processes for manufacturing such products. Modern chemical engineers actively collaborate with scientists and engineers from a broad range of disciplines; this collaborative nature is also the result of systems-thinking, which helps establish connections/bridges with other domains.

The need to navigate the complexity of chemical processes has required chemical engineers to develop *conceptual abstractions*. I would argue that the first of such abstractions was the notion of a “*unit operation*,” which is an abstraction that compartmentalizes chemical processes to understand functionality and connectivity/interactions of different components. The unit operation is a powerful *unifying abstraction* because it can be applied to a broad range of systems beyond chemical processes (e.g., the kidneys and the lungs can be seen as unit operations of the human body). The concept of the unit operation is so *ingrained* in our chemical engineering psyche that it is difficult for us to think in other ways (we naturally compartmentalize systems); this way of thinking has been challenged by our need to intensify

processes via tight integration of phenomena (e.g., reactive separations).

Chemical engineers have developed powerful conceptual abstractions by leveraging mathematical modeling techniques (e.g., transport models, thermodynamic models, optimization models, dynamics/control models). The use of mathematical modeling marked the origin of our research field of process systems engineering and marked the origin of other related fields such as computational fluid dynamics, computational chemistry, multi-scale modeling, and molecular simulations.

*Society is becoming increasingly complex*, linking humans, technology, policy, economy, and the environment in a complex manner. This coupling is giving rise to a broad range of problems that need to be addressed (e.g., air/water pollution, addiction, mental health, resource scarcity, climate change). It is important to recognize and acknowledge the role that chemical engineers have played in creating such problems (e.g., design of highly addictive drugs and foods, design of fuels and chemicals that rely on fossil sources, design of plastics that are difficult to recycle, socio-political conflicts due to oil exploitation). I would argue that many of the problems that we have created are, in part, the result of a *lack of “holistic” thinking*. For instance, in some cases, we have failed to anticipate connections of the products/processes that we create with other systems (we have taken a myopic view of problems). This failure has not always been intentional; in some cases, connections remained hidden from us due to limits in technology and scientific understanding; for instance, heroin was found to be addictive *years* after commercial manufacturing and PFAS contaminants are present in the environment at concentrations of *parts-per-trillion* (there are limited technologies that can detect this).

Chemical engineers have played and will play a critical role in providing solutions that benefit society; but, in doing so, *we need to think holistically*. This will require us to expand the domain of our thinking and to anticipate the impact of our products/processes with the environment and with society at large scales (e.g., economics and policy) and at molecular scales (e.g., toxicology). Expanding our domain of thinking can also help us identify innovative solutions to pressing problems. For instance, fostering molecular-level understanding can help us develop more effective drugs and can help us design better catalysts, but it is important to understand the impact of such drugs and catalysts on overarching systems (e.g., healthcare costs, materials availability, and carbon emissions). Moreover, fostering infrastructure-level understanding can help us develop more effective energy and manufacturing systems (e.g., electrochemical processes that interact with the power grid and technologies that valorize plastic waste). To enable more holistic thinking, it is necessary to adapt chemical engineering education



by expanding the current focus on product/process design to discover non-obvious interconnections with other systems and across diverse scales. This is not a straightforward task, as it requires *thinking deeply* about how to contextualize and link topics covered in the curriculum and outside of the curriculum.

In this short perspective, I argue that emerging technologies from machine learning, environmental sciences, molecular simulations, and mathematics can help us develop *new and powerful conceptual abstractions* that can help chemical engineers reinforce systems-thinking skills and solve problems of increasing complexity. Specifically, I discuss how *graph theory* provides a powerful unifying abstraction to represent complex systems and their connectivity across scales. Moreover, I discuss how *machine learning* tools can help develop models that link behavior across scales and how *lifecycle assessment tools* can help define and navigate domain boundaries that span multiple scales. My discussion will be inspired by problems that I have encountered in my own research and based on recent experiences in teaching topics of supply chains, machine learning, markets, and life-cycle assessment in my courses. I will also emphasize why I think that we need a new generation of modeling and simulation tools that help students visualize and navigate *hierarchical arrangement* and *multi-scale interactions*.

## UNIFYING ABSTRACTIONS

The notion of a unit operation provides a unifying abstraction that helps navigate complexity of diverse systems (e.g., by separating functionalities). The notion, importance, and use of the unit operation can be understood from the perspective of *graph theory*; specifically, we can think of a chemical process as a graph/network in which unit operations represent nodes and their connectivity is expressed as edges. The graph-theoretic view of a chemical process provides a powerful unifying abstraction that allows us to capture multi-scale interactions, see **Figure 1** [1]. For instance, going in the “up” direction, one can think of a chemical process as a node that forms part of a bigger graph/network that represents an entire chemical complex (a chemical complex involves processes that generate different types of products and that are exchanged with other processes). A chemical complex can also be seen as a node that forms part of a bigger graph that represents a local supply chain, the local supply chain is a node of a regional supply chain, and the regional supply chain is a node of a global supply chain (all these supply chains are linked and form a chemical supply chain). A chemical supply chain can also be seen as a node that forms part of a bigger graph that represents an *infrastructure network* (e.g., power grid and natural gas network); this reveals how a supply chain is connected to energy markets and can be affected by

disruptions of infrastructure. A supply chain can also be seen as a node that forms part of a bigger graph that represents society or the environment; this reveals how a supply chain is connected to consumer demands, resource availability, and weather disruptions. This also reveals that a supply chain can impact the environment and human health; for instance, a pandemic can suddenly increase demands for chemical products from the supply chain (sounds familiar?). Now, going in the “down” direction, one can think about molecules as graphs in which nodes represent atoms and edges represent bonds. One can represent environments that molecules form as graphs in which nodes represent an entire molecule and edges represent intermolecular forces (e.g., hydrogen bonds). One can also represent a molecular environment as a node that affects mesoscale behavior observed in a unit operation (e.g., a chemical reactor).

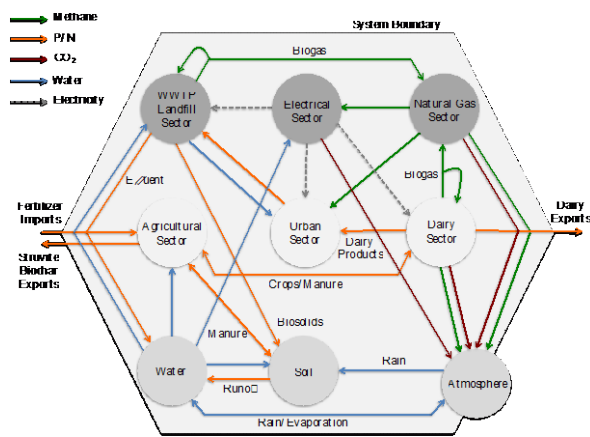
The previous discussion highlights how a graph is a conceptual abstraction that helps us think about how systems and phenomena are linked across scales. Computational and algorithmic tools from graph theory can be used to visualize/analyze such links and gain interesting insights. For instance, one can aim to identify nodes that are more tightly connected to others (and are thus more critical). Graph representations for molecules can also be used by machine learning (ML) models to predict emergent properties such as thermodynamic and toxicity properties. These machine learning models thus establish a link between molecular-scale behavior (non-observable) to meso-scale behavior (observable); in other words, ML models can be used to create bridges/edges across scales. ML can also be used to create graph representations of different resolution; for instance, one can create an ML model for a unit operation or an ML model an entire chemical process. ML thus allows us to aggregate nodes of a graph and thus to construct system representations of different levels of resolution (which can be critical in navigating and handling complexity). For example, when aiming to understand the behavior a supply chain, it might be sufficient to have a simple “input-output” model of a chemical process (there is no need to capture all the internal complexity of the process). Similarly, when aiming to understand the interdependencies of the power grid with the manufacturing sector, one can capture the sector using a simple input-output model.

Graph abstractions can also be used to capture connectivity of chemical processes with other systems (e.g., environment and policy). Capturing this connectivity is important, as this can help understand the impact that these externalities can have on the economic viability of the process or can help us understand how the process impacts markets and the environment. For instance, we can capture how a carbon tax impacts the selling price of a chemical, how carbon emissions impact the environment, and how the environment inherently restricts the

functionality/viability of a process (e.g., availability of resources such as water). Graph abstractions can also help engineers to *imagine or anticipate* what types of impacts their products and processes can have on humans and the environment (e.g., anticipate cascading effects). Uncovering non-obvious connections between systems can also be used to identify new markets and needs for products (e.g., one can make chemicals from cow manure via pyrolysis).

## ILLUSTRATIVE EXAMPLES

I now proceed to draw examples on how systems-thinking is necessary to tackle emerging societal challenges. In doing so, I will draw examples from my own research and teaching experiences.

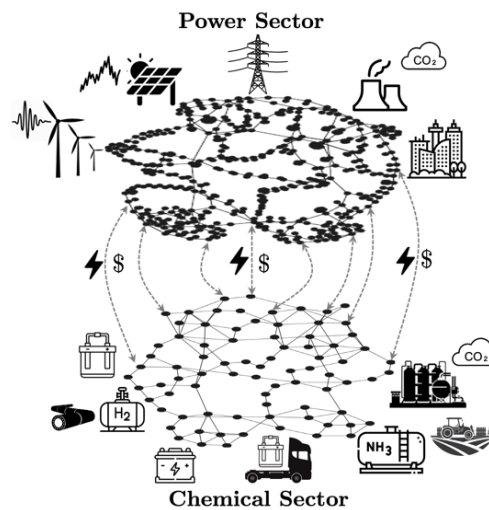


**Figure 2.** Interconnections of the dairy sector with other industrial sectors.

### Example I. Fertilizer Pollution

The U.S. dairy industry is a multibillion-dollar industry that provides essential food products (e.g., milk, cheese, yogurt) [2]. This industry operates a massive manufacturing and supply chain infrastructure that manages millions of cows that produce milk; cows are fed corn stover and grass, thus demanding substantial amounts of agricultural land and crops. Supplying feed for cows demands large amounts of nutrients, which are provided in the form of chemical fertilizer and cow manure. Cow manure is a natural fertilizer, and its use enables a certain degree of circularity for nutrients; however, loss of nutrients (e.g., fertilizer run-off) requires the use of synthetic chemical fertilizers such as ammonia and diammonium phosphate. Overapplication of fertilizer creates nutrient accumulation in soil, and uncontrolled run-off pollutes waterbodies (e.g., triggers algae blooms). Nutrients in waterbodies can be transported over long distances by following hydrological networks. For instance, nutrient pollution in Pennsylvania can propagate

all the way to the Mississippi river; this indicates that communities can be affected by pollution that is generated outside their jurisdiction (this is known as *environmental injustice*). In addition to water pollution, application of manure and synthetic fertilizer leads to significant emissions/leaks of methane and ammonia to the environment; for instance, organic matter in manure is degraded by bacteria to generate methane and ammonia injected in the soil (a gas) cannot be fully absorbed by the soil. The dairy industry is a sector that is tightly connected to other sectors (chemical industry and agricultural industry) and with the environment in complex ways, see **Figure 2**. The chemical industry can help address environmental impacts of the dairy sector by designing new fertilizer products that minimize run-off/emissions or by designing new crops that require less fertilizer. Nutrient run-off can also be mitigated by identifying technologies that can separate nutrients from manure streams in a scalable and economic manner.

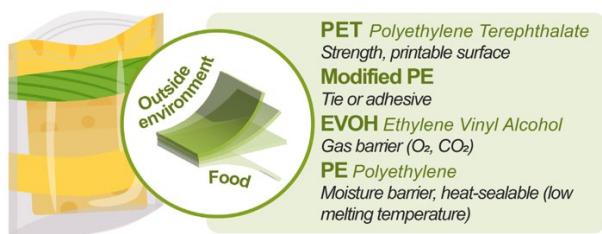


**Figure 3.** Interdependence between the power grid and the chemical sector.

### Example II. Exploiting Electricity Markets

Decarbonization of the chemical industry requires *electrification*, as this pathway can help leverage the use of renewable power [3]. At the same time, decarbonization of the power grid via adoption of renewable power requires that power demands/loads *are more flexible*, as such flexibility is key for mitigating intermittency of renewable power. Electrochemical technologies such as electrolysis will play a crucial role in decarbonizing both the chemical and power sectors, as these technologies help bridge/connect these sectors, see **Figure 3**. For instance, one can produce clean (carbon-free) hydrogen via water splitting systems to be used as feedstock for other chemical processes such as methanol synthesis, which currently uses large amounts of (carbon-intensive)

hydrogen generated via steam reforming and natural gas. Electrolyzers can exploit *dynamics of electricity markets*, which currently exhibit high fluctuations and, in some regions, high frequency of negative prices (consumers get paid to use power). It has recently been shown that hydrogen produced via electrolysis (and via exploitation of electricity market dynamics) can be cheaper than hydrogen produced via steam reforming. Moreover, because electrolysis is modular, one can envision creating hybrid systems that produce hydrogen on-demand and in locations where building natural gas delivery infrastructure is difficult. Identifying these types of opportunities requires that chemical engineers think about how their *processes connect with infrastructure* (power grid and natural gas networks) and associated markets. At the same time, chemical engineers should work with electrical engineers to identify how to *leverage electrolysis flexibility* in power grid operations. Electrochemical technologies also offer an opportunity to shift power demands over multiple temporal scales (e.g., from hours to days to seasons); visualizing how to schedule these shifts is challenging.



**Figure 4.** Design of a multi-layer plastic film for food packaging.

### Example III. Food Packaging Design

*Food packaging* is one of the main uses of plastic materials around the world. These packaging systems consist of *multi-layer plastic films*, which comprise different types of plastics that conduct different functionalities (e.g., oxygen barrier, moisture barrier, heat isolation, and mechanical strength), see **Figure 4** [4]. The engineering of these packaging materials helps minimize food spoilage and helps maximize access to food supply (better packaging leads to more flexible transportation and inventory management). Unfortunately, multi-layer plastic films cannot be recycled using mechanical recycling technologies. *Solvent-based separation technologies* offer a promising pathway to enable recovery of plastic materials from multi-layer films; in such processes, solvents are carefully selected to separate specific plastic films in a selective manner. The selection of the solvent plays a critical role in the efficiency of the separation process, in resource use (e.g., energy), and ultimately in recycling cost. Moreover, the selection of the solvent separation sequence is dictated by the nature of the multi-layer film

to be separated (and waste streams tend to have high variability). As such, when designing this recycling processes, it is important to think about how to design the food packaging materials, the solvent, and the process *simultaneously*. The selection of the solvent can be accelerated by leveraging molecular simulations tools such as *COSMO-RS*. Chemical engineers can leverage these types of tools to inform product/process design. Moreover, chemical engineers should use *lifecycle analysis tools* to quantify the environmental impact of recycling processes and to compare this with virgin plastic production; for instance, designing an ineffective recycling process can do more harm than good to the environment (e.g., can lead to a high carbon footprint or to higher ecotoxicity). Lifecycle analysis is a technique that fosters multi-scale, systems-thinking, as this requires the careful selection of domain boundaries, careful definition of pathways and steps followed to obtain products, and careful identification of interdependencies of these pathways and external systems (e.g., sources of energy and associated carbon footprints). Notions of graph theory can be valuable here, as they can help visualize interdependencies between products, technologies, and impacts. For instance, graphs can be used to identify all steps needed to arrive to plastic materials from oil and to identify all environmental impacts generated along the way.

## FUTURE OUTLOOK AND NEEDS

As society becomes more complex and it becomes increasingly difficult to assess how chemical products and processes interact with other systems, it will be necessary to identify better conceptual frameworks to capture connectivity of complex systems across multiple spatial and temporal scales. Graphs can be used to represent shifts of power in space and time, a principle that is being used to manage the power loads of power-intensive data centers. In addition to new conceptual frameworks, it is also necessary to develop modeling and simulation tools that can help engineers visualize multi-scale interdependencies arising in complex systems. This is critical in navigating complexity and in fostering understanding of impacts of products and processes at a large scale. It is also necessary to develop modeling tools that can link process simulators to molecular scale phenomena (e.g., microkinetic models and molecular simulations), to understand inherent physical limitations that processes might have. In enabling this couple, it is needed to develop tools that can seamlessly replace complex simulation models with machine learning models, which can be much easier to manage and manipulate.

## ACKNOWLEDGEMENTS

This material is based upon work supported by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Bioenergy Technologies Office under Award Number DEEE0009285. We also acknowledge support from the U.S. National Science Foundation under award number CMMI-2328160.

## REFERENCES

1. Jalving, J., Cao, Y., & Zavala, V. M. (2019). Graph-based modeling and simulation of complex systems. *Computers & Chemical Engineering*, *125*, 134-154.
2. Sampat, A. M., Ruiz-Mercado, G. J., & Zavala, V. M. (2018). Economic and environmental analysis for advancing sustainable management of livestock waste: A Wisconsin Case Study. *ACS sustainable chemistry & engineering*, *6*(5), 6018-6031.
3. Ma, J., Rebarchik, M., Bhandari, S., Mavrikakis, M., Huber, G. W., & Zavala, V. M. (2023). Exploiting electricity market dynamics using flexible electrolysis units for retrofitting methanol synthesis. *Energy & Environmental Science*, *16*(5), 2346-2357.
4. Sánchez-Rivera, K. L., Zhou, P., Kim, M. S., González Chávez, L. D., Grey, S., Nelson, K., ... & Huber, G. W. (2021). Reducing Antisolvent Use in the STRAP Process by Enabling a Temperature-Controlled Polymer Dissolution and Precipitation for the Recycling of Multilayer Plastic Films. *ChemSusChem*, *14*(19), 4317-4329.

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Part 1

## Peer-Reviewed Articles

Section 2: Advances in PSE Design





# Development of Mass/Energy Constrained Sparse Bayesian Surrogate Models from Noisy Data

Samuel Adeyemo and Debangsu Bhattacharyya\*

West Virginia University, Department of Chemical and Biomedical Engineering, Morgantown, West Virginia, USA

\* Corresponding Author: [Debangsu.Bhattacharyya@mail.wvu.edu](mailto:Debangsu.Bhattacharyya@mail.wvu.edu).

---

## ABSTRACT

This paper presents an algorithm for developing sparse surrogate models that satisfy mass/energy conservation even when the training data are noisy and violate the conservation laws. In the first step, we employ the Bayesian Identification of Dynamic Sparse Algebraic Model (BIDSAM) algorithm proposed in our previous work to obtain a set of hierarchically ranked sparse models which approximate system behaviors with linear combinations of a set of well-defined basis functions. Although the model building algorithm was shown to be robust to noisy data, conservation laws may not be satisfied by the surrogate models. In this work we propose an algorithm that augments a data reconciliation step with the BIDSAM model for satisfaction of conservation laws. This method relies only on known boundary conditions and hence is generic for any chemical system. Two case studies are considered—one focused on mass conservation and another on energy conservation. Results show that models with minimum bias are built by using the developed algorithm while exactly satisfying the conservation laws for all data.

**Keywords:** System Identification, Machine Learning, Algorithms, Design Under Uncertainty, Optimization

## 1.0 INTRODUCTION

Surrogate models are of great use for many process systems when developing the first-principles model is complex and/or time-consuming or the repeated simulation of the first-principles model for optimization/control is computationally expensive and/or difficult to converge reliably [1-3]. In recent years, greater access to numerous sensors, ease of collection and storage of large amount of data for many process systems and increased computational power with the emergence of supercomputers incentivized the development of data-driven models.

In the area of data-driven models, there are significant works on artificial neural network (ANN) models, with many powerful tools being readily available in the public domain. However, due to the black box nature, ANN models suffer from the lack of model interpretability which is a desired property of surrogate modeling. In addition, ANN models often require large amount of data for training while the developed models are often characterised with limited extrapolation capabilities. One approach to address these limitations is to develop data-driven

interpretable models by using well-defined basis functions such as those employed in Sparse Identification of Nonlinear Dynamics (SINDy) [4], Automatic Learning of Algebraic Models (ALAMO) [5] and Algebraic learning via elastic net (ALVEN)[6]. However, these existing approaches can perform poorly when trained with the noisy data especially for the case of correlated noise which is common for the industrial data. For the SINDy algorithm in particular, several methods including ensemble modeling [7] and implicit SINDy [8] have been proposed to enhance the robustness of the algorithm to the noisy data. Still, the level of noise that is acceptable for satisfactory model building can be highly limited. In addition, these approaches often result in more complicated models leading to some loss of desired model interpretability and sparsity. In our previous work [9], we proposed the Bayesian identification of Dynamic Sparse Algebraic Models (BIDSAM) algorithm that addresses the challenges of sparsity, model interpretability and robustness to noisy data. The developed algorithm employs Bayesian inferencing implemented in the expectation maximization (EM) framework for simultaneous model parameter estimation and uncertainty quantification that

explicitly accounts for correlation in the measurement noise, if exists. In that algorithm, system nonlinearities are approximated by linear combinations of linear and nonlinear basis functions that are transformations of input variables and their interactive effects among themselves and the output variables. The optimal sparsest subset of the resulting large family of basis functions is selected using the branch and bound algorithm which returns a set of hierarchically ranked sparse models by minimizing an information criterion that not only incentivizes model fitness but also accounts for parameter estimability.

It is desired that models of chemical systems satisfy mass/energy constraints. Else the predicted results by a surrogate model can be meaningless. To this end, there exists several works in which physics-informed machine learning algorithms are developed for satisfying some physics constraints [10-12]. Most of these are not for satisfying mass or energy constraints, but for satisfying other physical constraints such as thermodynamic constraints. In addition, these approaches are highly specific requiring detailed knowledge of the system, which may not be available for a system where data-driven models are being developed.

In this work, we propose an approach that extends the capabilities of the BIDSAM algorithm for identification of robust sparse models that satisfy mass/energy conservation laws exactly both for the forward and inverse problems without loss of sparsity and model interpretability. The proposed approach guarantees the satisfaction of the conservation laws for all predictions while relying only on the knowledge of boundary conditions.

## 2.0 THEORY

### 2.1 Sparse Model Selection and Parameter Estimation

Consider a general nonlinear system with states  $x$ , inputs  $u$ , measured outputs  $y$ , and model parameters  $\theta$  represented by the following set of differential and algebraic equations:

$$\dot{x} = f(x, u, \theta) \quad (1)$$

$$y = g(x) \quad (2)$$

The BIDSAM algorithm proposed in our previous work [9] approximates the system nonlinearities with a linear combination of linear and nonlinear transformations of input variables and their interactive effects among themselves and the measurements. The resulting model takes the following form:

$$y_k = \mathbf{A}y_{k-1} + \mathbf{C}U_{k-1}^* + \omega_k \quad (3)$$

where  $U_{k-1}^*$  represents the set of basis functions,  $\omega_k$  stands for additive noise and  $\mathbf{A}, \mathbf{C}$  are the model

parameters. Using Bayesian inferencing implemented in the Expectation-Maximization (EM) framework, uncertainty quantification is undertaken while estimating model parameters. A branch and bound algorithm is developed to search for the sparsest subset of basis functions that best approximates the system behavior given the available data. The BIDSAM algorithm returns desired numbers of hierarchically ranked models based on a modified Akaike information criteria that is used as the model fitness criterion.

### 2.1 Model Update for Satisfaction of Conservation Laws

The models resulting from the BIDSAM algorithm show superior performance compared to existing algorithms especially when training data is corrupted with correlated noise. But the BIDSAM models are not guaranteed to satisfy conservation laws. This becomes very important especially when the training data are significantly noisy and violate the conservation laws. In this work, the BIDSAM model is augmented with a dynamic data reconciliation (DDR) step and an algorithm for updating model parameters so that final model results satisfy the mass/energy conservation laws. The overall model structure is given in Figure 1. The approach includes a linear transformation of the output from the BIDSAM model,  $x$ , by introducing a transformation that brings in more degrees of freedom for the optimization.

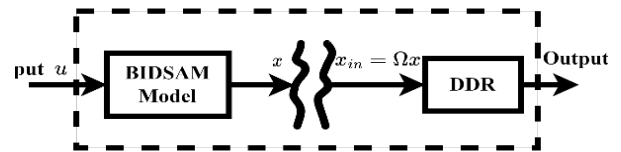


Figure 1: Mass/energy constrained model

#### 2.2.1 Inverse problem

The inverse problem involves updating the BIDSAM model parameters in an iterative manner while solving the following optimization problem.

$$\min_{\Omega, y_{ddr}} \sum_i (\Omega x^k - y_{ddr})^T W_1^{-1} (\Omega x^k - y_{ddr}) + \sum_i (y_{data} - y_{ddr})^T W_2^{-1} (y_{data} - y_{ddr}) \quad (4)$$

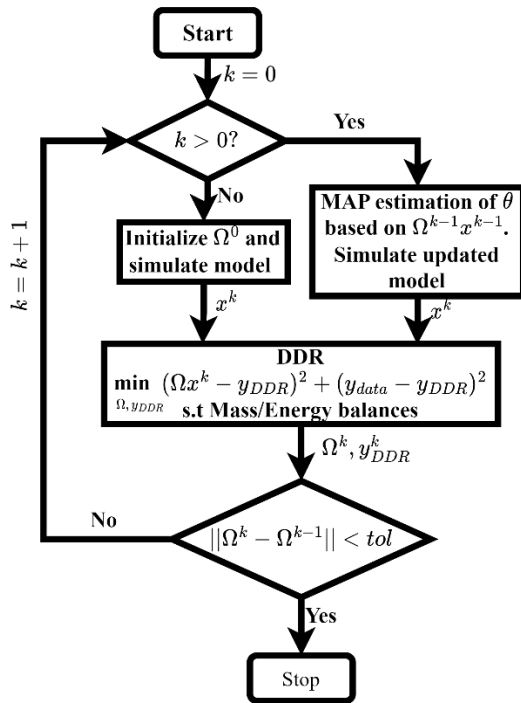
s. t. Mass/Energy balances

In Equation 4,  $W_i$ 's represent the weighting matrices,  $y_{data}$  denote the measured data and  $y_{ddr}$  denote the final model output that are desired to satisfy mass/energy balances. For the case of mass conservation, atom balance can take the following general form:

$$\sum n_{j,p_i}(u_p(k)) = \sum n_{j,q_o}(y_{q,adr}(k)) \quad (5)$$

where  $n_{j,p_i}$  is the number of atoms of element  $j$  with respect to input variable  $p$ , and  $n_{j,q_o}$  are the number of atoms of element  $j$  with respect to output variable  $q$ . For energy conservation, a similar approach is followed by considering enthalpy in and out for the system as discussed later in section 3.2 for the second case study. Although equation 5 has been written for a general case of multiple component situation, the same conditions apply to scenarios in which only total mass flow is known at the inlet or outlet of the system. The situation in which only total flows in and out of the system are known is considered trivial and mass is assumed to be generally conserved conditioned on the known mass flow rate into the system as shown in the second case study.

The overall algorithm for solving the inverse problem is shown in Figure 2.



**Figure 2.** Sequential algorithm for inverse problem.

In the first iteration, initial values are assumed for the hyperparameters  $\Omega$ , then we solve the optimization problem in Equation 4 using the Interior point (IPOPT) algorithm [13]. In subsequent iterations, model parameters are updated based on the optimal values obtained for  $\Omega$  in the previous iteration. This is done by undertaking a maximum a posteriori (MAP) estimates of these parameters using Bayesian inferencing and the EM algorithm as detailed in our previous work [9]. The iteration continues till

convergence.

### 2.2.2 Forward problem

The overall structure in Figure 1 remains the same for the forward problem. However, in this case, the parameters  $\Omega$  and  $\theta$  assume the optimal values obtained from the inverse problem and remain unchanged while DDR is undertaken just one time for each set of desired prediction. The following objective function replaces that in Equation 4 for the forward problem:

$$\min_{\Omega, y_{adr}} \sum_i (\Omega x^k - y_{adr})^T W_1^{-1} (\Omega x^k - y_{adr}) \quad (6)$$

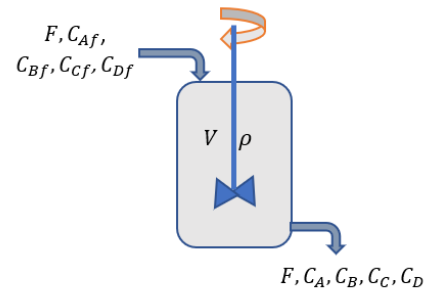
The constraints for this optimization remain the same as for the inverse problem.

## 3.0 CASE STUDIES

The efficiency of the developed algorithm is tested on two case studies: an isothermal CSTR and a lumped parameter superheater system.

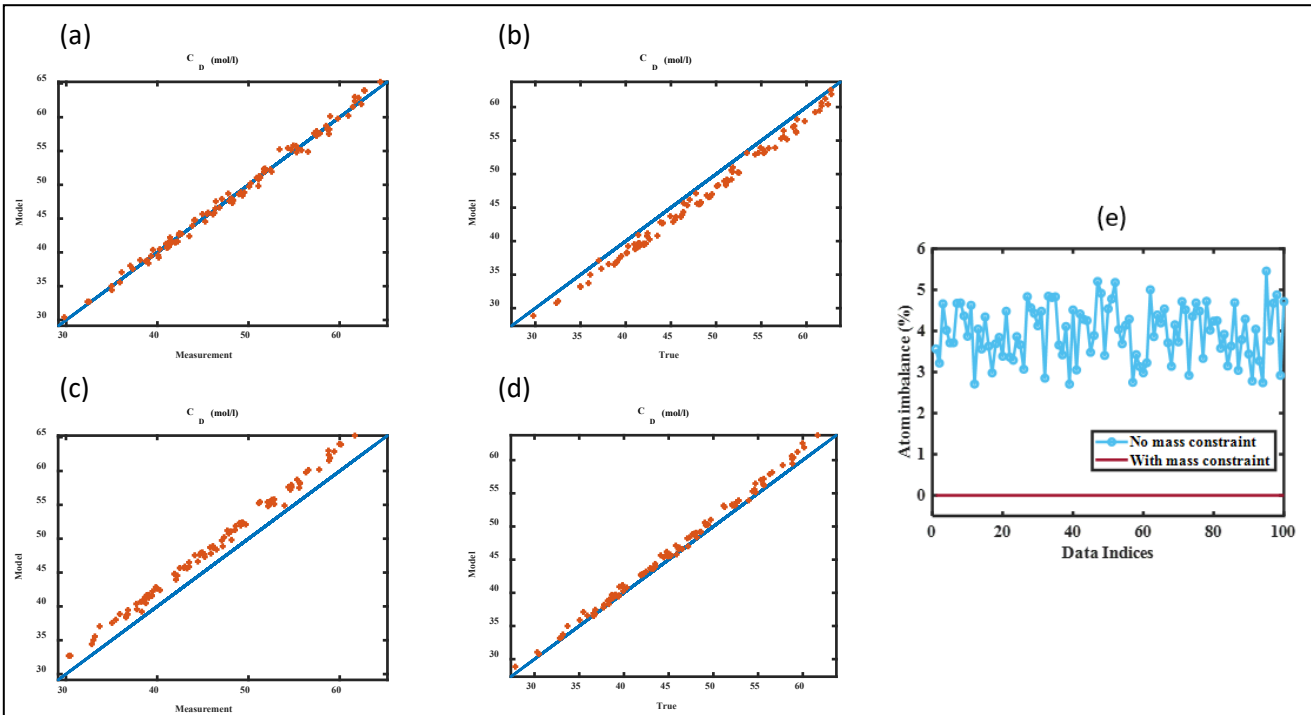
### 3.1 Isothermal CSTR

For this case study, the performance of the developed algorithm is examined for the satisfaction of mass conservation. A schematic of the CSTR is shown in Figure 3. Chemical reactions for this system are as follows:



**Figure 3.** Schematic diagram of isothermal CSTR.

For this modeling purpose, the input variables considered are the reactor feed rate ( $F$ ) and species concentrations at the inlet ( $C_{Af}$ ,  $C_{Bf}$ ,  $C_{Cf}$  and  $C_{Df}$ ). We seek to predict the concentration of each species ( $C_A$ ,  $C_B$ ,  $C_C$  and  $C_D$ ) at the reactor outlet based on changes in the input variables. Training data are simulated by solving the differential equations used to model the system from first-principles. Simulated correlated noise and constant bias



**Figure 4.** Validation plots for isothermal CSTR model: (a)-(b) BIDSAM, (c)-(d) BIDSAM with mass constraint.

**Table 1.** Performance measures of the models of the isothermal CSTR.

Mass constrained?	No. of parameters	% RMSE			
		$C_A$	$C_B$	$C_C$	$C_D$
<b>Trained using True Data</b>					
No	24	1.32	0.64	1.47	1.59
Yes		1.44	0.75	1.41	1.02
<b>Trained using Noisy Data with Bias</b>					
No	24	4.36	4.04	4.45	4.25
Yes		2.05	1.81	2.00	2.12

are added to this data which is then used for model training. Validation plots for a model obtained when training is done by using noisy data with bias are shown in Figure 4.

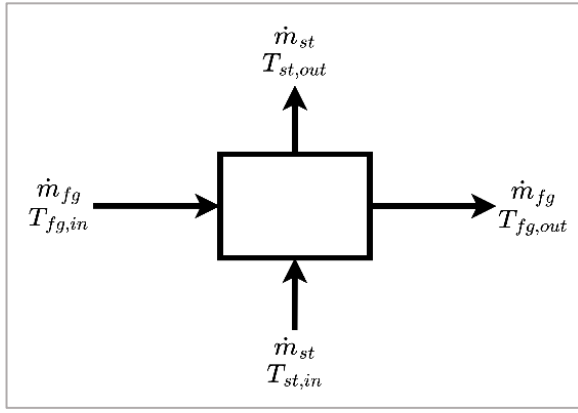
Figures 4(a-b) show that while the results from the BIDSAM model without constraints match the training data quite well but there is bias with respect to the truth. Figures 4(c-d) show that the results from the model built by using the algorithm with mass constraint exhibit bias with respect to the biased noisy data, but are closer to the truth. Figure 4(e) shows that while the model built without considering mass conservation constraints violates the mass balance, the algorithm with mass constraint included results in exact satisfaction of the conservation law. For training with noisy data with bias, the maximum %RMSE for the unconstrained algorithm is 4.45% while for the constrained algorithm we have

2.05%.

### 3.2 Adiabatic superheater system

This case study evaluates the performance of the BIDSAM model with energy constraints included. A superheater from a power plant, shown in Figure 5, where steam is superheated by using the hot flue gas is considered. The system has a cross-flow configuration. The input variables considered are the steam and flue gas mass flowrates ( $\dot{m}_{st}$ ,  $\dot{m}_{fg}$ ), temperatures ( $T_{st,in}$ ,  $T_{fg,in}$ ) and pressures ( $P_{st,in}$ ,  $P_{fg,in}$ ). The output variables are the steam and flue gas outlet temperatures ( $T_{st,out}$ ,  $T_{fg,out}$ ).

The energy balance constraints to be included as equality constraints in the optimization problem given by Equation 4 are as follows:



**Figure 5.** Schematic of lumped superheater system.

$$\begin{aligned} & \dot{m}_{fg}(k) C_{p_{fg}}(k) (T_{fg,in}(k) - T_{fg,out}(k)) \\ & = \dot{m}_{st}(k) (h_{st,out}(P_{st}, T_{st,out}, k) - h_{st,in}(P_{st}, T_{st,in}, k)) \quad (8) \end{aligned}$$

Here, it is assumed that the flue gas is treated as an ideal gas with constant specific heat capacity while specific enthalpies for steam are computed by using the IAPWS R7-97 correlations. Similar to the first case study, model training is done by using the noisy data with constant bias. Plots comparing results from the models built with and without energy constraints are shown in Figure 6.

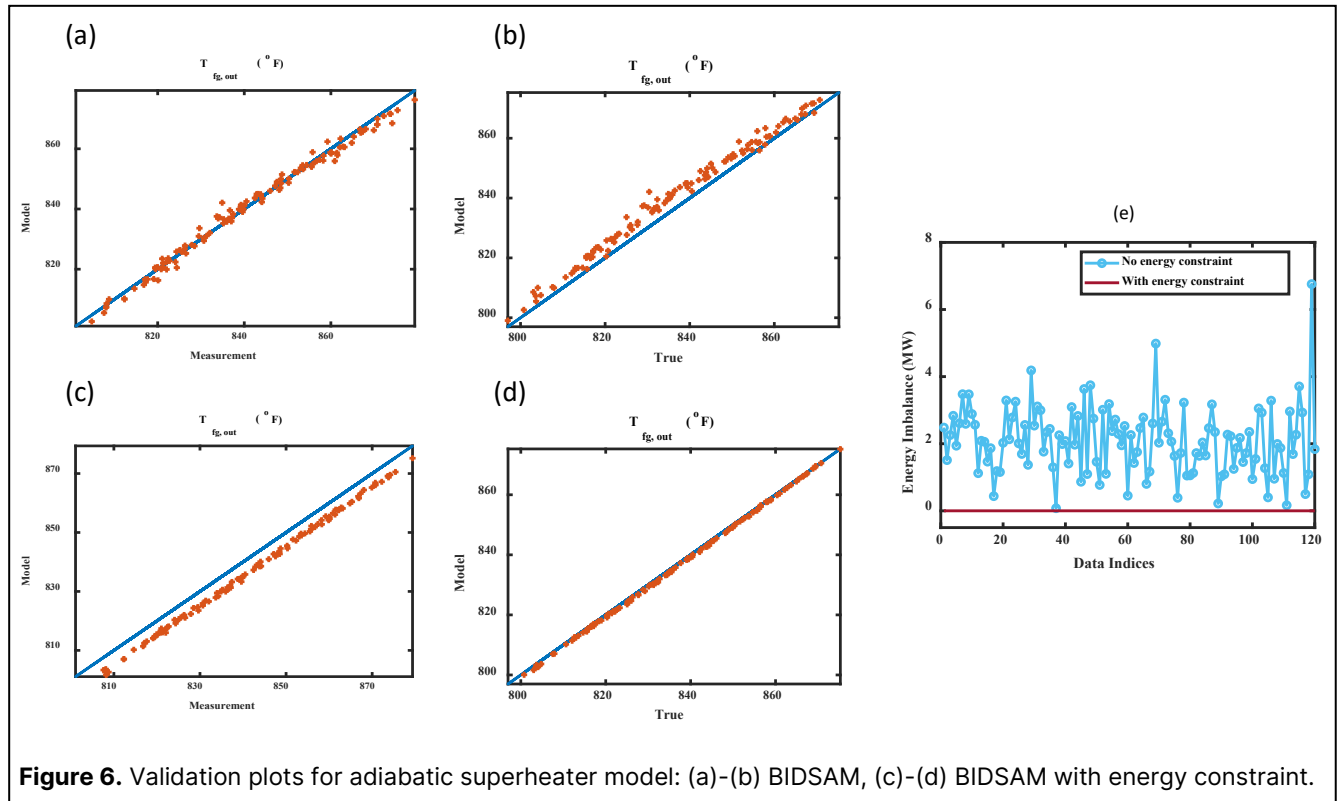
Figures 6(a-b) show the results from the model built by using the algorithm without energy constraints while

Figures 6(c-d) compare the results from the model built by using the algorithm with energy constraints. Figure 6(c-e) show that although the algorithm with energy constraints included leads to biased estimate with respect to the data, it has practically negligible bias with respect to the truth and also satisfies the energy conservation constraints. This is reverse for the model built by using the algorithm without energy constraints.

The key performance measures are shown in Table 2. It is observed that while the RMSE for both output variables have increased when trained using the data with bias for both the models built with or without considering energy constraints, the relative increase is higher for the model built without considering energy constraints. Particularly for the flue gas temperature ( $T_{fg,out}$ ) there is considerable difference in the results obtained by using the algorithm with energy constraints.

**Table 2.** Performance measures of the models for the adiabatic superheater system.

Mass constrained?	No. of parameters	% RMSE	
		$T_{st,out}$	$T_{fg,out}$
<b>Trained Using True Data</b>			
No	10	1.35	1.98
Yes		0.99	0.23
<b>Trained Using Noisy Data with Bias</b>			
No	8	4.10	4.31
Yes		2.61	0.75



**Figure 6.** Validation plots for adiabatic superheater model: (a)-(b) BIDSAM, (c)-(d) BIDSAM with energy constraint.



## CONCLUSIONS

An algorithm has been developed that guarantees the satisfaction of conservation laws by sparse models obtained by using the enhanced BIDSAM algorithm. The results for the case studies show that the sparse models built by using the algorithm guarantee the satisfaction of mass/energy balance even when trained by using data that are noisy with constant bias with respect to the truth. For the isothermal CSTR case study, %RMSE of the models built with and without mass constraints remains similar when the training is done by using the true data but there was considerable difference in results from the models when training is done by using the noisy data with bias. Similar observations are made for the adiabatic superheater case study as well, there was reduction in the model error with the constraints where energy constraints are satisfied by the enhanced BIDSAM algorithm. One limitation of the proposed algorithm is in the relatively larger training time for the selection of appropriate basis function using the BIDSAM approach. Detailed analysis of this can be found in our previous work [9]. In the future, the algorithm will be extended to application for dynamic model building and for simultaneous conservation of mass and energy while also improving on the computational efficiency for model selection.

## ACKNOWLEDGEMENTS

The work was funded by the U.S DOE through the project titled "Boiler Health Monitoring Using a Hybrid First Principles-Artificial Intelligence Model" (Grant # DE-FE0031768). The DOE financial support is gratefully acknowledged by the authors of this paper.

## REFERENCES

1. McBride K, Sundmacher K. Overview of Surrogate Modeling in Chemical Process Engineering. *Chemie-Ingenieur-Technik* 91:228–239 (2019).
2. Gonçalves GFN, Assen B, Yuyi L, Lachlan RM, Indrani P, Omar KM. Data-driven surrogate modeling and benchmarking for process equipment. *Data-Centric Eng.* 1, (2020).
3. Panahi M, Khezri V, Yasari E, Skogestad S. Application of surrogate models as an alternative to process simulation for implementation of the self-optimizing control procedure on large-scale process plants—a natural gas-to-liquids (gtl) case study. *Ind. Eng. Chem. Res.* 60:4919–4929 (2021).
4. Brunton SL, Proctor JL, Kutz JN, Bialek W. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Natl. Acad. Sci. U. S. A.* 113:3932–3937 (2016).

5. Cozad A, Sahinidis NV, Miller DC. A combined first-principles and data-driven approach to model building. *Comput. Chem. Eng.* 73:116–127 (2015).
6. Sun W, Braatz, RD. ALVEN: Algebraic learning via elastic net for static and dynamic nonlinear model identification. *Comput. Chem. Eng.* 143:107103 (2020).
7. Fasel U, Kutz JN, Brunton BW, Brunton SL. Ensemble-SINDy: Robust sparse model discovery in the low-data, high-noise limit, with active learning and control. *Proc. R. Soc. A Math. Phys. Eng. Sci.* 478, (2022).
8. Kaheman K, Kutz JN, Brunton SL. SINDy-PI: A robust algorithm for parallel implicit sparse identification of nonlinear dynamics: SINDy-PI. *Proc. R. Soc. A Math. Phys. Eng. Sci.* 476, (2020).
9. Adeyemo S, Bhattacharyya D. Optimal nonlinear dynamic sparse model selection and Bayesian parameter estimation for nonlinear systems. *Comput. Chem. Eng.* 180: 108502 (2024).
10. Karniadakis GE, Kevrefidis IG, Lu L, Perdikaris P, Wang S, Yang L. Physics-informed machine learning. *Nat. Rev. Phys.* 3:422–440 (2021).
11. Sansana J, Joswaik MN, Castillo I, Wang Z, Rendall R, Chiang LH, Reis MS. Recent trends on hybrid modeling for Industry 4.0. *Comput. Chem. Eng.* 151:107365 (2021).
12. Bradley W, Kim J, Kilwein Z, Blakely L, Eydenberg M, Jalvin J, Laird C, Boukouvala, F. Perspectives on the integration between first-principles and data-driven modeling. *Comput. Chem. Eng.* 166: 107898 (2022).
13. Wächter A, & Biegler LT. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming* vol. 106 (2006).

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Guaranteed Error-bounded Surrogate Framework for Solving Process Simulation Problems

Chinmay M. Aras<sup>a</sup>, Ashfaq Iftakher<sup>a</sup>, and M. M. Faruque Hasan<sup>a,b\*</sup>

<sup>a</sup> Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, Texas 77843-3122, USA

<sup>b</sup> Texas A&M Energy Institute, Texas A&M University, College Station, Texas 77843-3122, USA

\* Corresponding Author: [hasan@tamu.edu](mailto:hasan@tamu.edu).

## ABSTRACT

Process simulation problems often involve systems of nonlinear and nonconvex equations and may run into convergence issues due to the existence of recycle loops within such models. To that end, surrogate models have gained significant attention as an alternative to high-fidelity models as they significantly reduce the computational burden. However, these models do not always provide a guarantee on the prediction accuracy over the domain of interest. To address this issue, we strike a balance between computational complexity by developing a data-driven branch and prune-based framework that progressively leads to a guaranteed solution to the original system of equations. Specifically, we utilize interval arithmetic techniques to exploit Hessian information about the model of interest. Along with checking whether a solution can exist in the domain under consideration, we generate error-bounded convex hull surrogates using the sampled data and Hessian information. When integrated in a branch and prune framework, the branching leads to the domain under consideration becoming smaller, thereby reducing the quantified prediction error of the surrogate, ultimately yielding a solution to the system of equations. In this manner, we overcome the convergence issues that are faced by many simulation packages. We demonstrate the applicability of our framework through several case studies. We first utilize a set of test problems from literature. For each of these test systems, we can find a valid solution. We then demonstrate the efficacy of our framework on real-world process simulation problems.

**Keywords:** Surrogate Model, Modelling and Simulations, Algorithms, Data-Driven

## INTRODUCTION

Systems of equations arise in many fields of practical interest such as engineering, economics, physics, and chemistry. A system of equations in  $N$  -dimensions is represented as:

$$\mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_N(\mathbf{x}))^T$$

Simulation of chemical processes involve numerically solving such systems of equations. Solving such a system means finding a vector  $\mathbf{x}$  such that  $\mathbf{F}(\mathbf{x}) = 0$  where  $\mathbf{x} = (x_1, x_2, x_3, \dots, x_n)$ , which is not trivial, but there are many existing methods that have been developed to solve such systems traditionally. These methods can be categorized under three general classes: i) Homotopy continuation methods; ii) Interval-Newton methods; iii) Newton-type methods.

Homotopy continuation methods involve starting with a system of equations  $\mathbf{G}(\mathbf{x})$ , whose solution is known, embedded in a homotopy function  $\mathbf{H}(\mathbf{x}, t)$  along with the original system of equations  $\mathbf{F}(\mathbf{x})$  where  $t$  is the homotopy parameter. The idea here is to solve a series of problems as  $t$  increases from 0 to 1 to obtain the solutions of  $\mathbf{F}(\mathbf{x})$  by following the solution paths starting from  $\mathbf{G}(\mathbf{x})$  [1,2]. Interval-Newton methods involve finding intervals of the solution space where solutions can exist with mathematical guarantee [2,3].

Newton's methods [4] or Newton-type methods are iterative techniques that generate increasingly improved approximations of the root with each iteration. These methods have a faster rate of convergence but there are a few issues associated with Newton-type methods [5], e.g., i) Newton-type methods require computation of the Jacobian while some require the inverse of the Jacobian

as well which may be computationally expensive to obtain for large systems of equations at every iteration; ii) This may also lead to numerical instabilities with convergence if the Jacobian is very close to zero; iii) The convergence of such methods also depends upon the initialization. It is difficult to know apriori if a particular initialization converges to a root within a desired neighborhood. If initialized at a point which is far away from a root, convergence to that desired root may not be obtained.

Several other algorithms have been developed which do not exactly fall under any one of these categories but are nevertheless intelligent algorithms to find solutions to systems of equations. Grosan and Abraham [6] developed an evolutionary algorithm where they convert the system of equations into a multi-objective optimization problem and use pareto dominance relationships to find solutions. Ramos and Vigo-Aguiar [5] developed an approach which involves replacing a non-linear equation with 2 associated equations with the idea being that the system of associated equations is easier to solve than the original equation using Newton's methods. Ramos and Monteiro [7] further developed this approach to solve systems to non-linear equations.

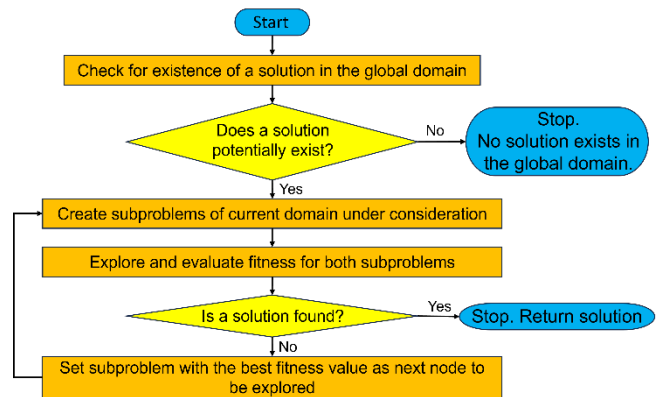
This work aims to propose and study the GEMS framework which we have developed to solve systems of equations. GEMS stands for Guaranteed Error-bounded Modeling of Surrogates which involves replacing each of the original  $C^2$ -continuous equations by surrogates. The maximum error between these equations and their respective surrogates is bounded by an error metric,  $\epsilon_{max}$ , which can be derived through the theory of guaranteed estimators and using interval arithmetic methods [8]. One of the key features of the GEMS framework is that it does not depend on the Jacobian of the system of equations but does require a bound on the diagonal elements of the Hessian. Another key feature is that it allows the user to define domain bounds to selectively search the space for solutions. Lastly, it defines a convergence criterion which ensures that a solution is found.

The contents of this work are organized as follows: the methodology section introduces the overall flow of the framework along with the theoretical requisites. In the next section, we present case studies to demonstrate the applicability of the framework. The computational results are reported in the next section, finally followed by the conclusions.

## METHODOLOGY

We propose the GEMS framework to solve systems of equations with one of its key ideas being that it can check whether a solution exists in a search space using sampled data, the error metric  $\epsilon_{max}$ , and interval arithmetic. From the point of view of the user, GEMS takes in input data in the form of the systems of equations  $F(x)$

and the global domain bounds and returns a solution to  $F(x)$ . When provided with a system of equations and the global domain bounds, the framework first checks for the existence of a solution in the global search space. If a solution can exist, the search space (domain) is then divided into two subproblems which are explored. Exploration of a subproblem involves checking for the existence of a solution which, if exists, necessitates the evaluation of the fitness for the subproblem. The domain of the subproblem with the best fitness is then set as the domain to be explored next. This is done iteratively until a solution is found. A high-level flow of the GEMS framework is shown in Figure 1.



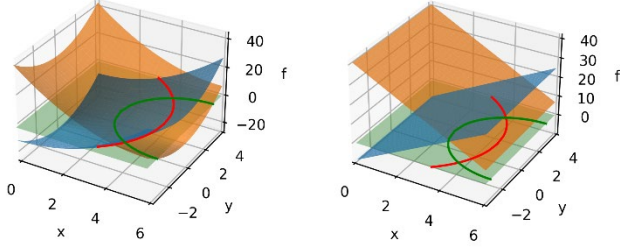
**Figure 1:** Overall flow of the GEMS framework.

The subsequent subsections explain the theories and subroutines that are utilized in the GEMS framework.

## Data Sampling and Generation of Convex Hull Surrogates

GEMS replaces the original system of equations with convex hull surrogates. Given a search space i.e., a bounded domain, the functions are evaluated (sampled) at the domain bounds so that the set of points used to generate the convex hull can cover even the extremities of the search space. An  $N$ -dimensional space has  $2^N$  vertices. The number of vertices increases drastically as  $N$  increases thereby making it impractical to have a function evaluation at every vertex for higher dimensional problems. Caratheodory's theorem states that given an  $N$ -dimensional space, if any point  $x$  lies within a convex hull defined by a set of points, then it is possible to express  $x$  as a convex combination of at most  $N + 1$  points in that set of points [9]. In other words,  $N + 1$  points are sufficient to generate a convex hull surrogate in any domain. When obtaining  $F(x)$  is expensive, this theorem can be used to reduce the number of required function evaluations. We consider both cases where  $2^N$  and  $N + 1$  points are used to generate the convex hull surrogates. To illustrate, the 2-D functions  $f_1(x, y) = x^2 + y^2 - 25$  and  $f_2(x, y) = (x - 6)^2 + y^2 - 9$  with their convex hull

surrogates  $c_1(x, y)$  and  $c_2(x, y)$  which are constructed using the vertex evaluations are shown in Figure 2.



**Figure 2:**  $f_1(x, y)$  (blue) and  $f_2(x, y)$  (orange) shown in a given domain (left). Their respective convex hulls  $c_1(x, y)$  (blue) and  $c_2(x, y)$  (orange) shown over the same domain (right). The red and green curves represent the zeros of  $f_1(x, y)$  and  $f_2(x, y)$  respectively.

## Interval Arithmetic

We use our in-house subroutine to compute Hessian bounds for a  $C^2$ -continuous function in a closed form involving the standard arithmetic, trigonometric, logarithmic, and exponential operators. We first generate the computational graph for a given function, and then implement interval arithmetic rules in all nodes of the graph to efficiently compute bounds on the output. An automatic differentiation type scheme is then used to compute bounds on the gradients. The computational graph that corresponds to the gradient is dynamically created to store all dependencies. Using these graphs, we then automatically compute bounds on the Hessian. A key advantage of this approach is the guaranteed estimation of Hessian bounds that can not be achieved using point-wise evaluation techniques. In some cases, we also use interval arithmetic on the computational graph to compute a lower bound,  $lbmin$  and upper bound,  $ubmax$  for each function in the case where enough samples aren't available.

## Guaranteed Estimators

Now that Hessian bounds can be obtained, the concept of guaranteed estimators is utilized which is used to compute the error metric  $\epsilon_{max}$ . Namely, we use the guaranteed edge-concave underestimator [10],  $L(x)$  defined as:

$$L(x) = f(x) - \sum_{i=1}^N \theta_i^L (x_i - x_i^G)^2 \quad (1)$$

where  $\theta_i^L = \max\left\{0, \left[\frac{d^2 f}{dx^2}\right]^U\right\}$

The edge-concave definition for a function  $f$  is equivalent to:

$$\frac{d^2 f}{dx^2} \leq 0 \quad \forall i = 1, \dots, N \quad (2)$$

For an edge concave function,  $\theta_i^L = 0 \forall i = 1, \dots, N$

Similarly, we use the edge-convex overestimator,  $U(x)$  defined as [11]:

$$U(x) = f(x) + \sum_{i=1}^N \theta_i^U (x_i - x_i^G)^2 \quad (3)$$

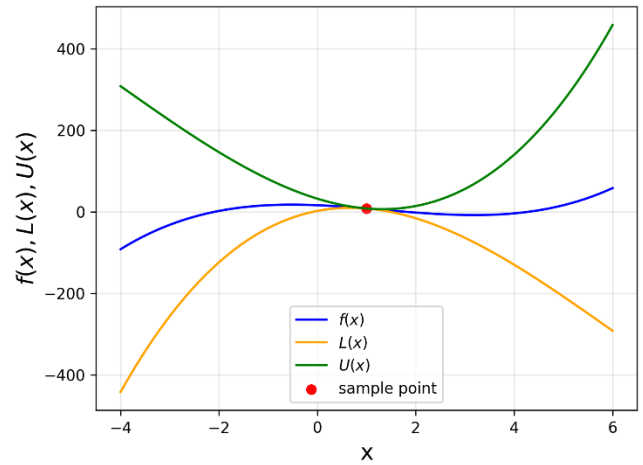
where  $\theta_i^U = \min\left\{0, -\left[\frac{d^2 f}{dx^2}\right]^L\right\}$

The edge-convex definition for a function  $f$  is equivalent to:

$$\frac{d^2 f}{dx^2} \geq 0 \quad \forall i = 1, \dots, N \quad (4)$$

For an edge convex function,  $\theta_i^U = 0 \forall i = 1, \dots, N$

$x_i^G$  is the point from which the estimator is generated. The estimators are illustrated in Figure 3.



**Figure 3:** The function  $f(x)$  shown along with its guaranteed underestimator  $L(x)$  (shown in orange) and its guaranteed overestimator  $U(x)$  (shown in green) generated from the sample point (shown in red).

## Error-boundedness

$\epsilon_{max}$  is the maximum possible error between the equations and their respective surrogates and therefore it is desired to have the smallest  $\epsilon_{max}$  as possible. If the closed form of the equations is convoluted in the sense that it contains nested terms, interval arithmetic may result in extremely large (loose) values of  $\epsilon_{max}$ .

Based on the nature of the function i.e., based on its  $\theta_i^L$  and  $\theta_i^U$  values, each of the functions in the original system can be classified into one of three categories:

- i) Category 1 - The function is neither edge-convex nor edge-concave in all dimensions,
- ii) Category 2 - The function is both edge-concave and edge-convex in all dimensions,
- iii) Category 3 - The function is edge-convex or

edge-concave in some but not all dimensions.

The value for  $\epsilon_{max}$  for each of these categories when we consider the set of all  $2^N$  vertices and then a subset of  $N + 1$  of those  $2^N$  are shown in Tables 1 and 2 respectively.

**Table 1:**  $\epsilon_{max}$  values when all  $2^N$  vertices are evaluated.

Category 1	$\epsilon_{max} = \sum_{i=1}^N (\theta_i^l + \theta_i^u) (x_U^i + x_M^i)$
Category 2	$\epsilon_{max} = \max\{f_{eval}\} - \min\{f_{eval}\}$
Category 3	$\epsilon_{max} = f_{IAmax} - f_{IAmin}$

**Table 2:**  $\epsilon_{max}$  values when  $N + 1$  vertices are evaluated.

Category 1	$\epsilon_{max} = \sum_{i=1}^N (\theta_i^l + \theta_i^u) (x_U^i + x_M^i)$
Category 2	$\epsilon_{max} = f_{IAmax} - f_{IAmin}$
Category 3	$\epsilon_{max} = f_{IAmax} - f_{IAmin}$

Here,  $\max\{f_{eval}\}$  is the maximum of the vertex function evaluations,  $\min\{f_{eval}\}$  is the minimum of the vertex function evaluations,  $f_{IAmax}$  is the upper bound on the function obtained through interval arithmetic and  $f_{IAmin}$  is the lower bound on the function obtained through interval arithmetic.

### Solving the system of surrogates

Now that we have our system of error-bounded convex hull surrogates, we wish to find a point  $x^*$  in the search space such that the distance  $\tau$  between furthest the convex hull surrogate and the zero "plane" is minimized. The following linear programming (LP) model CH enables us to do that:

$$\begin{aligned} & \min \tau \\ \text{s. t.} \quad & x_n = \sum_{v \in cv} \lambda_{j,v} \hat{x}_{v,n} \quad \forall n \in N, j \in J \end{aligned} \quad (5)$$

$$CH_j = \sum_{v \in cv} \lambda_{j,v} f_j(\hat{x}_{v,n}) \quad \forall j \in J \quad (6)$$

$$\sum_{v \in cv} \lambda_{j,v} = 1 \quad \forall j \in J \quad (7)$$

$$CH_j \leq \tau \quad \forall j \in J \quad (8)$$

$$-CH_j \leq \tau \quad \forall j \in J \quad (9)$$

$$0 \leq \lambda_{j,v} \leq 1 \quad \forall j \in J, v \in V \quad (10)$$

In this formulation, Equations 5, 6 and 7 are referred to as the convex combination constraints. Equations 7 and 10 ensure that the point lies inside the domain under consideration. Here,  $J$  is the set of functions and  $V$  is the set of vertices.

Note that,  $\hat{x}_{v,n}$  is the known co-ordinates of vertex  $v$  in dimension  $n$  and  $f_j(\hat{x}_{v,n})$  is the value of the function

evaluation for a function  $j$  at the point  $\hat{x}_{v,n}$ .  $\lambda_{j,v}$  is the weight assigned to vertex  $v$  for the function  $j$ .

Solving this formulation, we obtain a point ( $x^*$ ) at which we evaluate fitness defined as follows:

$$Fitness = \sum_{j=1}^N f_j^2(x^*)$$

This fitness allows us to define a metric for quantifying the "closeness to zero" or quality of the obtained solution. It is desirable to have a fitness function value close to zero.

### Branch and Prune

With the theoretical requisites established, GEMS involves a branch and prune subroutine (which is inspired by the classical branch and bound) where, given an initial search space, the idea is to repeat the following steps iteratively:

- 1) Prune regions of the search space based of the existence of a solution in that space.
- 2) Branching to generate two subproblems if a solution could exist.
- 3) Choose subproblem with the best (smallest) fitness value to be explored next.

Each subproblem in the branch and prune tree is also called a node. The current node is branched at the midpoint of the domain of the variable having the longest edge. For systems of equations where the domain bound ranges are unequal, we scale the domain bounds relative to the global domain bounds and branch at the midpoint of the domain of the variable having the scaled longest edge. To check whether a solution exists in each search space (each node), the GEMS framework does the following: for each function, we perform a loose pruning check which involves computing a lower bound,  $lbmin$  and an upper bound,  $ubmax$  based solely using interval arithmetic. For any function, if its lower and upper bounds have the same sign, it is not possible for that function to have a solution in that space. Each function has a distinct lower and upper bound. If any of the functions have  $lbmin$  and  $ubmax$  of the same sign, it is impossible for the entire system of equations to have a solution in that space, thereby allowing us to discard that node from further consideration. Note that, we are using the  $lbmin$  and  $ubmax$  based solely on IA which are loose. However, this check helps eliminate unnecessary computation. We move to the next step of node exploration if a solution can exist in the node i.e., if it cannot be pruned based on the loose pruning condition. We then perform function evaluations and Hessian bound estimations to obtain tighter  $lbmin$  and  $ubmax$  values for another tighter pruning check. Figure 4 shows the node exploration subroutine. The  $lbmin$  and  $ubmax$  for each of the categories when we consider the set of all  $2^N$  vertices are



shown in table 3. Similarly, the values of  $lbmin$  and  $ubmax$  when we consider a subset of  $N + 1$  vertices are shown in table 4. Note that for functions which are both edge-concave and edge-convex in all dimensions (Category 2), when we have function evaluations at all  $2^N$  vertices, Tardella's theorem [12,13] enables us to obtain tightest  $lbmin$  and  $ubmax$ .

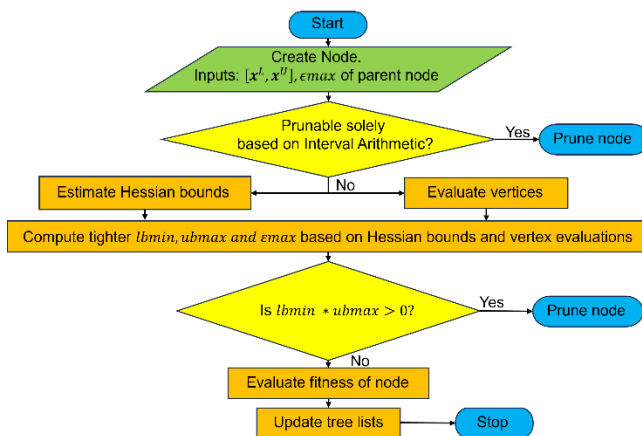
**Table 3:**  $lbmin$  and  $ubmax$  values when all  $2^N$  vertices are considered.

Category 1	$lbmin = \min\{f_{eval}\} - \epsilon_{max}$ $ubmax = \max\{f_{eval}\} + \epsilon_{max}$
Category 2	$lbmin = \min\{f_{eval}\}$ $ubmax = \max\{f_{eval}\}$
Category 3	$lbmin = f_{IAmin}$ $ubmax = f_{IAmax}$

**Table 4:**  $lbmin$  and  $ubmax$  values when all  $N + 1$  vertices are considered.

Category 1	$lbmin = \min\{f_{eval}\} - \epsilon_{max}$ $ubmax = \max\{f_{eval}\} + \epsilon_{max}$
Category 2	$lbmin = f_{IAmin}$ $ubmax = f_{IAmax}$
Category 3	$lbmin = f_{IAmin}$ $ubmax = f_{IAmax}$

If  $lbmin$  and  $ubmax$  have different signs, a solution may possibly exist, and we move to the next step of node fitness evaluation. The node fitness evaluation involves generation of the convex hull surrogates and using the LP formulation to find the point at which we evaluate fitness. If the fitness of a node falls below a certain threshold ( $<1e-4$ ) we have found a point for which the value of each function in  $F(x)$  is extremely close to zero which basically means we have found a solution to the original system of equations.



**Figure 4:** The node exploration subroutine.

## Convergence

As  $\epsilon_{max}$  reduces, the error between the function and its convex hull surrogate reduces. The error metric  $\epsilon_{max}$  is dependent on the Hessian bounds and domain sizes which means the  $\epsilon_{max}$  of a child node can never be greater than the  $\epsilon_{max}$  of its parent. In other words, the error-bound for any subproblem is less than or equal to its parent problem. The sum of all  $\epsilon_{max}$ s for the child node will always be greater than the sum of  $\epsilon_{max}$ s of the parent node. It can be considered an equivalent to the gap in the classical branch and bound technique used in optimization. Also, having a small enough value of  $\epsilon_{max}$  can lead to efficient pruning of sub-regions.

This concludes the overall flow of the GEMS framework.

## CASE STUDIES

We consider some test problems having a single solution as well as a few benchmark problems from literature along with examples of systems of equations arising from test problems in chemical engineering literature.

### Test problems

- 1) Hypersphere systems – The following system of equations can be scaled to N-dimensions while still maintaining the characteristic of having a single solution. We consider the 10, 50 and 100-dimensional systems for our studies. Here, R is the radius of the hyperspheres.

$$F_{HSN}(x) = \begin{cases} \sum_{i=1}^N x_i^2 - R = 0 \\ (x_1^2 - 2R)^2 + \sum_{i=2}^N x_i^2 - R = 0 \\ \sum_{i=1}^N x_i^2 + (x_j - R)^2 - R = 0, \forall j = (3, \dots, N) \end{cases}$$

### Benchmark systems in literature

- 1) Interval Arithmetic [14] – The following system has been considered as one of the benchmark problems from interval arithmetic:

$$F_{IA}(x) = \begin{cases} x_1 - 0.25428722 - 0.18324757x_4x_3x_9 = 0 \\ x_2 - 0.37842197 - 0.16275449x_1x_{10}x_6 = 0 \\ x_3 - 0.27162577 - 0.16955071x_1x_2x_{10} = 0 \\ x_4 - 0.19807914 - 0.15585316x_7x_1x_6 = 0 \\ x_5 - 0.44166728 - 0.19950920x_7x_6x_3 = 0 \\ x_6 - 0.14654113 - 0.18922793x_8x_5x_{10} = 0 \\ x_7 - 0.42937161 - 0.21180486x_2x_5x_8 = 0 \\ x_8 - 0.07056438 - 0.17081208x_1x_7x_6 = 0 \\ x_9 - 0.34504906 - 0.19612740x_{10}x_6x_8 = 0 \\ x_{10} - 0.42651102 - 0.21466544x_4x_8x_1 = 0 \end{cases}$$

- 2) Neurophysiology Application [15] – The

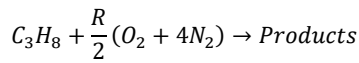
following system arises from an application in neurophysiology:

$$F_N(x) = \begin{bmatrix} x_1^2 + x_3^2 = 1 \\ x_2^2 + x_4^2 = 1 \\ x_5x_3^2 + x_6x_4^3 = c_1 \\ x_5x_1^3 + x_6x_2^3 = c_2 \\ x_5x_1x_3^2 + x_6x_2x_4^2 = c_3 \\ x_5x_3x_1^2 + x_6x_4x_2^2 = c_4 \end{bmatrix}$$

where the constants  $c_i$  can be arbitrarily chosen. We consider all  $c_i = 0$ .

## Systems arising from process engineering models

- 1) Combustion Application [16] – The following system arises from the combustion of propane in air. The stoichiometric equation is:



where R is the relative amounts of air to fuel. Combustion Application – The following system arises from the combustion of propane in air. The variables  $x_i$  for  $i = 1$  to 10 are the number of moles of product  $i$  formed per every mole of propane combusted.

$$F_{CA}(x) = \begin{bmatrix} x_1 + x_4 = 3 \\ 2x_2 + 2x_5x_6x_7 = 8 \\ 2x_3 + x_9 = 40 \\ 2x_1 + x_2 + x_4 + x_7 + x_8 + x_9 + 2x_{10} = 10 \\ x_1x_5 = K_5x_2x_4 \\ \sqrt{40}x_6x_1^{0.5} = K_6\sqrt{x_2x_4N_{Tot}} \\ \sqrt{40}x_7x_4^{0.5} = K_7\sqrt{x_1x_2N_{Tot}} \\ 40x_8x_4 = K_8x_1N_{Tot} \\ \sqrt{40}x_9x_4 = K_9x_1\sqrt{x_3N_{Tot}} \\ 40n_{10}n_4^2 = K_{10}n_1^2N_{Tot} \end{bmatrix}$$

where,

$$N_{Tot} = \sum_{i=1}^{10} x_i$$

The first four equations in the system arise from the mole balances while the rest arise from the equilibrium relations.  $K_i$  for  $i = 5$  to 10 are the equilibrium constants.

- 2) Chemical Equilibrium [17] – The mathematical structure of the combustion application system is analysed and reformulated by using the notion of ‘element variables’ (surrogates for atomic combinations). The first four equations arise from the reformulation of the mole balances in the combustion application system whereas the equilibrium relations reduce to a single equation because of the element variable substitutions

which leads to the following system:

$$F_{CEQ}(x) = \begin{bmatrix} x_1x_2 + x_1 - 3x_5 = 0 \\ 2x_1x_2 + x_1 + x_2x_3^2 + R_8x_2 - Rx_5 \\ + 2R_{10}x_2^2 + R_7x_2x_3 + R_9x_2x_4 = 0 \\ 2x_2x_3^2 + 2R_5x_3^2 - 8x_5 \\ + R_6x_3x_4 + R_7x_2x_3 = 0 \\ 2x_4^2 - 4Rx_5 + R_9x_2x_4 = 0 \\ x_1(x_2 + 1) + R_{10}x_2^2 + x_2x_3^2 + R_8x_2 \\ + R_5x_3^2 + x_4^2 - 1 + R_6x_3 + R_7x_2x_3 \\ + R_9x_2x_4 = 0 \end{bmatrix}$$

Note that the variables  $x_i$  for  $i = 1$  to 5 have been redefined along with the constant R for  $i = 5$  to 10 to account for the element variable substitutions.

- 3) Recycle and Purge in the Synthesis of Ammonia – The ammonia synthesis process involving a recycle and a purge stream is considered. Given the feed stream specifications and the reactor conversion, it is desired to calculate the overall conversion of nitrogen and the ratio of moles of gas purged to the moles of gas leaving the condenser.

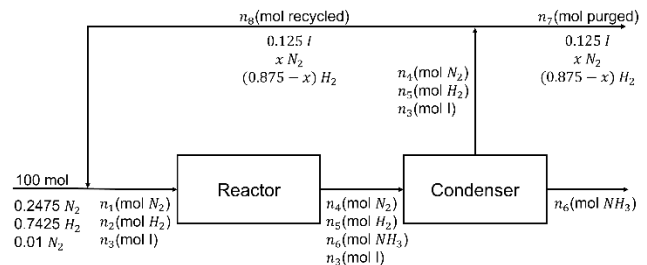


Figure 4: Flowsheet for the ammonia synthesis process.

The following system of equations arises from the overall mole balances, stoichiometric relations, and material balances at the mixing and split points.

$$F_{AS}(x) = \begin{bmatrix} 1 - 0.125n_6 = 0 \\ 148.5 - 3n_5 - 1.75n_6 + 2n_6x = 0 \\ 49.5 - 2n_6x - n_5 = 0 \\ n_6 - 0.75n_1 = 0 \\ 0.5n_1 - n_5 = 0 \\ n_6 - xn_7 - n_4x = 0 \\ 74.25 + 0.875n_7 - xn_7 - n_2 = 0 \\ 1 + 0.125n_7 - n_3 = 0 \end{bmatrix}$$

The domain ranges considered for each of these problems are shown in Table 5.

Table 5: Domain bounds for the test problems

Problem Instance	Domain Ranges
10D-hyperspheres	$x_i \in [-5,5] \forall i = 1, \dots, 10$

50D-hyperspheres	$x_i \in [-5,5] \forall i = 1, \dots, 50$
100D-hyperspheres	$x_i \in [-5,5] \forall i = 1, \dots, 100$
Interval Arithmetic	$x_i \in [-4,4] \forall i = 1, \dots, 10$
Neurophysiology Application	$x_1 \in [-10,5]$ $x_2 \in [-10,6]$ $x_3 \in [-10,7]$ $x_4 \in [-10,8]$ $x_5 \in [-10,9]$ $x_6 \in [-10,9]$
Combustion Application	$x_i \in [-5,5] \forall i = 1, \dots, 10$
Chemical Equilibrium	$x_i \in [-5,5] \forall i = 1, \dots, 5$
Ammonia Synthesis	$n_i \in [0,500] \forall i = 1, \dots, 7$ $x \in [0,1]$

## COMPUTATIONAL EXPERIMENTS AND RESULTS

All runs were performed on an Intel Xeon Gold 6248R 3GHz processor running on Linux (CentOS 7). The algorithm was developed in Python v3.6.8 and used CPLEXv20.1.0.1 in GAMSv35.1 to solve the LP model CH for the system of surrogates.

**Table 6:** Fitness of solution obtained, number of nodes explored and number of function evaluations for the test problems.

Instance	Fitness of Solution	Nodes Explored	Function Evaluations
10-D hyperspheres	7e-5	228	1487
50-D hyperspheres	1e-4	1278	35827
100-D hyperspheres	4e-5	813	51713
Interval Arithmetic	1e-5	64.48	1823
Neurophysiology Application	6e-5	307	1301
Combustion Application	1e-4	21650	97856
Chemical Equilibrium	8e-5	837	3149
Ammonia Synthesis	1e-8	383	54483

We have reported the solutions for each of the test problems using a sampled data set of  $N + 1$  vertices except for the system of equations arising from the ammonia synthesis model where we have used  $2^N$  vertices. The solution of the LP (to minimize the distance to the furthest convex hull surrogate) highly depends on choice of the  $N + 1$  sampled vertices. In other words, the choice of the set of the sampled data affects the path of the branch

and prune search.

## CONCLUSIONS

In this work, we report a data-driven branch and prune approach that has the potential to unlock new avenues for finding solutions to systems of equations without the requirement of an initial guess. Furthermore, the user can specify the domain bounds if solutions within a desired search space are required. Multiple potential sampling strategies need to be investigated which may aid in faster convergence. Further work is needed to reduce the number of required function evaluations.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge support from the NSF CAREER (CBET-1943479), and the DOE AICHE RAPID Institute (DE-EE0007888-09-03) grants. Part of the research was conducted with the computing resources provided by Texas A&M High Performance Research Computing.

## REFERENCES

- Wayburn, T. L., & Seader, J. D. Homotopy continuation methods for computer-aided process design. *Computers & Chemical Engineering*, 11(1), 7-25. (1987)
- Maranas, C. D., & Floudas, C. A. Finding all solutions of nonlinearly constrained systems of equations. *Journal of Global Optimization*, 7, 143-182. (1995)
- Hansen, E. R., & Greenberg, R. I. An interval Newton method. *Applied Mathematics and Computation*, 12(2-3), 89-98. (1983)
- Remani, C. Numerical methods for solving systems of nonlinear equations. Lakehead University Thunder Bay, Ontario, Canada, 77. (2013)
- Ramos, H., & Vigo-Aguiar, J. The application of Newton's method in vector form for solving nonlinear scalar equations where the classical Newton method fails. *Journal of computational and applied mathematics*, 275, 228-237. (2015)
- Grosan, C., & Abraham, A. A new approach for solving nonlinear equations systems. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 38(3), 698-714. (2008)
- Ramos, H., & Monteiro, M. T. T. A new approach based on the Newton's method to solve systems of nonlinear equations. *Journal of Computational and Applied Mathematics*, 318, 3-13. (2017)
- Alefeld, G., & Mayer, G. Interval analysis: theory and applications. *Journal of computational and*

- applied mathematics, 121(1-2), 421-464. (2000)
9. Floudas, C. A. Deterministic global optimization: theory, methods and applications (Vol. 37). Springer Science & Business Media. (2013)
  10. Hasan, M. M. F. An edge-concave underestimator for the global optimization of twice-differentiable nonconvex problems. *Journal of Global Optimization*, 71(4), 735-752. (2018)
  11. Iftakher, A., Aras, C. M., Monjur, M. S., & Hasan, M. F. Data-driven approximation of thermodynamic phase equilibria. *AIChE Journal*, 68(6), e17624. (2022)
  12. Tardella, F. On a class of functions attaining their maximum at the vertices of a polyhedron. *Discrete applied mathematics*, 22(2), 191-195. (1988)
  13. Tardella, F. On the existence of polyhedral convex envelopes (pp. 563-573). Springer US. (2004)
  14. Moore, R. E. Methods and applications of interval analysis. Society for Industrial and Applied Mathematics. (1979)
  15. Verschelde, J., Verlinden, P., & Cools, R. Homotopies exploiting Newton polytopes for solving sparse polynomial systems. *SIAM Journal on Numerical Analysis*, 31(3), 915-930. (1994)
  16. Shacham, M., Brauner, N., & Cutlip, M. B. A web-based library for testing performance of numerical software for solving nonlinear algebraic equations. In *Computer Aided Chemical Engineering* (Vol. 9, pp. 291-296). Elsevier. (2001)
  17. Meintjes, K., & Morgan, A. P. A methodology for solving chemical equilibrium systems. *Applied Mathematics and Computation*, 22(4), 333-361. (1987)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Process Flowsheet Optimization with Surrogate and Implicit Formulations of a Gibbs Reactor

Sergio I. Bugosen<sup>a,b</sup>, Carl D. Laird<sup>b</sup>, and Robert B. Parker<sup>a,c,\*</sup>

<sup>a</sup> Information Systems and Modeling, Los Alamos National Laboratory, Los Alamos, NM, USA

<sup>b</sup> Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA, USA

<sup>c</sup> Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, NM, USA

\* Corresponding Author: [rbparker@lanl.gov](mailto:rbparker@lanl.gov).

---

## ABSTRACT

Alternative formulations for the optimization of chemical process flowsheets are presented that leverage surrogate models and implicit functions to replace and remove, respectively, the algebraic equations that describe a difficult-to-converge Gibbs reactor unit operation. Convergence reliability, solve time, and solution quality of an optimization problem are compared among full-space, ALAMO surrogate, neural network surrogate, and implicit function formulations. Both surrogate and implicit formulations lead to better convergence reliability, with low sensitivity to process parameters. The surrogate formulations are faster at the cost of minor solution error, while the implicit formulation provides exact solutions with similar solve time. In a parameter sweep on the autothermal reformer flowsheet optimization problem, the full-space formulation solves 33 out of 64 instances, while the implicit function formulation solves 52 out of 64 instances, the ALAMO polynomial formulation solves 64 out of 64 instances, and the neural network formulation solves 48 out of 64 instances. This work demonstrates the trade-off between accuracy and solve time that exists in current methods for improving convergence reliability of chemical process flowsheet optimization problems.

---

**Keywords:** Surrogate modeling, Machine learning, Chemical process design, Chemical process optimization, Nonlinear optimization.

## INTRODUCTION

The selection of optimal operating conditions is a fundamental task in chemical process design. This requires the minimization or maximization of an objective function subject to nonlinear and nonconvex constraints. While local solvers such as CONOPT [1] and IPOPT [2] can handle these problems, their convergence can be very sensitive to model formulation, initial guess, and scaling factors. In addition, flowsheet design equations usually contain complex nonlinear expressions, including logarithms, high-degree polynomials and bilinear terms. At certain variable values, these equations can be undefined, and their Jacobian can become singular, hindering convergence.

Some methods to improve convergence of a full space flowsheet optimization problem and address sensitivity to initial parameters include using (1) sophisticated initialization routines and (2) model reformulation

strategies, such as surrogate models and implicit functions.

In the case of (1), it is well known that an effective initialization strategy determines how easily the optimization algorithm converges to a solution [3]. However, finding a good initialization is a cumbersome task because some unit operations do not easily converge for a set of specifications, and there is no systematic way to determine good initial values [4]. In addition, there is a significant computational cost associated with trying multiple initialization methods for each problem instance one attempts to solve.

Regarding (2), surrogate models aim to be simple models that approximate the input and output behaviour of complex systems (in our case study, a Gibbs reactor) over a specific input domain. A surrogate model typically relaxes accuracy in exchange for lower dimensionality and more reliable convergence [5]. Even if the development of these models requires computationally



expensive simulations over a range of input values, they are computationally cheaper to evaluate once they are embedded into a larger interconnected system of equations, such as a chemical process flowsheet. Two surrogates that have been recently used in chemical engineering are neural networks and ALAMO polynomials [6]. Neural networks have been widely used in process control, modeling, and optimization. For instance, Henao and Maravelias use neural networks to model the production of maleic anhydride in a superstructure optimization problem [7]. The ALAMO framework [8] is a recently developed tool used to build simple and accurate polynomial surrogates from a minimal set of training data. It makes use of an integer programming technique to choose basis functions of the input variables and compute the output variables as a linear combination thereof [6,9]. Surrogate optimization using the ALAMO framework has been used in superstructure optimizations for carbon capture systems [10], in global optimization of poly-generation systems [11], and distillation sequences [9].

Reformulating a model with the implicit function theorem, as proposed by [12], aims to exploit the fact that the difficulty of converging a large-scale nonlinear program (NLP) may be due to the effort required to converge the system of nonlinear equality constraints corresponding to specific unit operations in the flowsheet. Solving these units separately from the original formulation can significantly improve convergence reliability. In this work, we take advantage of non-singularity of the Gibbs reactor equations. Given that the Jacobian of these equations with respect to the Gibbs reactor's output variables is non-singular, solving the Gibbs reactor equations yields a unique solution for these output variables. This implies that there exists a function mapping state and input variables to the outputs of the Gibbs reactor. The utility of this property is that complicated algebraic equations from a unit operation can be removed from the NLP using implicit functions. This leaves a smaller set of equations to be seen by the solver and fewer variables to initialize and scale [12].

The surrogate and implicit reformulation strategies are similar in that they both remove complicating equations that describe a difficult-to-converge unit model. In this work, we design an autothermal reformer flowsheet (ATR) in the IDAES process modeling framework [13] and we formulate a full space optimization problem, as shown in Section 3. We then compare convergence reliability, solution quality, and solve time among full-space, implicit function, and surrogate formulations of this problem, and study the trade-offs among these formulations.

## BACKGROUND

### Full Space Formulation

Vector  $x$  is a vector of state variables corresponding to each unit model in the flowsheet, while  $u$  is a vector of manipulated inputs. Function  $G$  describes operational constraints, function  $H$  describes the equality constraints corresponding to each unit model in the flowsheet not including the unit model we intend to replace, and function  $R$  describes the equality constraints only for this unit model. Vector  $y$  is a vector containing the outlet variables corresponding to this unit and any additional intermediate variables used only by this unit. Eq. (1d) will be replaced by either a surrogate model or an implicit function.

$$\max f(x, u) \quad (1a)$$

$$\text{s. t. } G(x, u) \leq 0 \quad (1b)$$

$$H(x, y, u) = 0 \quad (1c)$$

$$R(x, y, u) = 0 \quad (1d)$$

### Surrogate Formulation

This formulation is identical to the one shown in Eq. (1) except for Eq. (2d), which instead of representing first-principles design equations for the unit model of interest, contains a surrogate model  $\bar{R}(x, y, u)$  that approximates this unit's behavior. While this surrogate can be obtained using a variety of methods, as described in [6], in this paper we use neural network and polynomial surrogates.

$$\max f(x, u) \quad (2a)$$

$$\text{s. t. } G(x, u) \leq 0 \quad (2b)$$

$$H(x, y, u) = 0 \quad (2c)$$

$$\bar{R}(x, y, u) = 0 \quad (2d)$$

### Implicit Formulation

The implicit function theorem states that Eq. (1d) can be reformulated as  $y = R_y(x, u)$  if  $\nabla_y R$  is non-singular for all values of state and input variables. Under this condition, Eq. (3) is an exact reformulation of Eq. (1). We solve for  $y$  externally as a square system of equations in a separate interface and the resulting values and derivatives are communicated back to the optimization solver, specifically to Eq. (3c), which links the outlet variable  $y$  calculated by the implicit function to the inner optimization problem. Since this formulation keeps Eq. (1d) feasible and the unit model equations are not seen by the NLP solver, we expect it to lead to better convergence reliability.

$$\max f(x, u) \quad (3a)$$

$$\text{s. t. } G(x, u) \leq 0 \quad (3b)$$

$$H(x, R_y(x, u), u) = 0 \quad (3c)$$

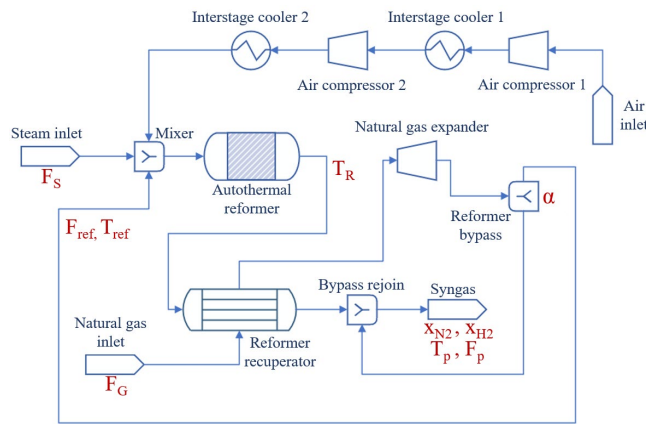
To solve this formulation with a second-order optimization method (e.g., IPOPT), the routines implemented

in [12] are used to calculate the constraint Jacobian, objective gradient, and Hessian of the Lagrangian. As an example, Eq. (4) shows the Jacobian of  $y$  as a function of  $x$  and  $u$ , which is used to calculate the former derivative matrices.

$$\nabla_{x,u}y = -\nabla_y R^{-1} \nabla_{x,u} R \quad (4)$$

## PROBLEM STATEMENT

An autothermal reforming flowsheet is used as an example in the Optimization & Machine Learning Toolkit (OMLT) [14]. The main objective of the autothermal reformer (ATR) is to produce syngas, mainly composed of  $H_2$ ,  $CO$ ,  $CH_4$  and  $CO_2$ . The process is shown in Figure 1. First, a mixture of natural gas, steam, and air is fed into the ATR (modeled as a Gibbs reactor). The hot syngas is then circulated through a shell and tube heat exchanger, also called the reformer recuperator, to heat the natural gas feed. This natural gas feed is then expanded to generate electrical power and is finally fed into the reactor, closing the loop.



**Figure 1.** Process Flow Diagram of the ATR process.

In this study, the objective of all optimization problems is to maximize the hydrogen concentration in the syngas stream, as shown in Eq. (5),

$$\max x_{H_2} \quad (5)$$

The operational constraints are specified in Eq. (6).

$$T_R \leq 1200 [K] \quad (6a)$$

$$T_p \leq 650 [K] \quad (6b)$$

$$F_p \geq 3500 \left[\frac{mol}{s}\right] \quad (6c)$$

$$x_{N_2} \leq 0.3 \quad (6d)$$

$$1120 \leq F_G \leq 1250 \left[\frac{mol}{s}\right] \quad (6e)$$

$$0 \leq \alpha \leq 1 \quad (6f)$$

$$200 \leq F_s \leq 350 \left[\frac{mol}{s}\right] \quad (6g)$$

Where  $T_R$  is the autothermal reformer outlet temperature and  $T_p$ ,  $F_p$  and  $x_{N_2}$  correspond to product's temperature, molar flow rate and nitrogen concentration respectively. The manipulated variables are  $F_G$ ,  $F_s$  and  $\alpha$ , corresponding to inlet molar flow rate of natural gas, inlet molar flow rate of steam, and bypass fraction respectively. The full space optimization problem is composed of 898 variables and 895 equality constraints describing material, energy, and momentum balances for each unit model, where 55 of these correspond to algebraic equations describing the Gibbs reactor, such as Gibbs energy minimization, and enthalpy, element, pressure, total flow, and component flow balances. As an example, the Gibbs minimization equation is shown in Eq. (7a).

$$g_{partial,j} + \sum_e (L_e \times \beta_{j,e}) = 0, \quad \forall j \in J \quad (7a)$$

$$J = \{H_2, CO, H_2O, CO_2, CH_4, C_2H_6, C_3H_8, C_4H_{10}, O_2\} \quad (7b)$$

$$J_{inert} = \{N_2, Ar\} \quad (7c)$$

Here,  $g_{partial,j}$  is the partial molar Gibbs energy of component  $j$ ,  $L_e$  is the Lagrange multiplier of element  $e$  and  $\beta_{j,e}$  is the number of moles of element  $e$  in one mole of component  $j$  [13].

In this work, we implement full-space, surrogate, and implicit function formulations for solving this optimization problem. The Gibbs reactor, Eq. (1d), will be replaced by a surrogate and an implicit function.

## IMPLEMENTATION

We used Pyomo 6.7.1 [15], an open-source optimization modeling language, and IDAES 2.4.0 [13], a process modeling framework, to design the ATR flowsheets in Python 3.9.6. The optimization problems were solved with IPOPT 3.14.11 [2] using linear solver MA27, called via the Cylpopt interface. The surrogates were obtained using the ALAMO machine learning framework [8] and TensorFlow 2.15. The neural network was embedded into the optimization problem using OMLT 1.1 [14]. Derivative computations are performed via the AMPL solver library (ASL) [16] via the PyNumero interface [17]. Results were produced on a machine with an Apple M1 Max processor and 32 GB of RAM running macOS Ventura 13.6.6.

### Dataset Generation

A dataset describing the Gibbs reactor was generated to train the ALAMO polynomials and the neural network. This data was generated from 625 samples of the four-dimensional input space: five samples in each input arranged in a regular grid. The ranges of sampled inputs are shown in Table 1. The outputs (13 in total) correspond to reactor's outlet temperature ( $T_{out}$ ), outlet molar flow rate ( $F_{out}$ ), and outlet compositions for the eleven

components. Time to acquire this data was 400 seconds.

**Table 1:** Inputs and ranges of the ALAMO and neural network surrogates.

Input Variable	Unit	Range
Steam molar flow rate ( $F_s$ )	[mol/s]	200 – 350
Natural gas temperature ( $T_{ref}$ )	[K]	600 – 900
Natural gas flow rate ( $F_{ref}$ )	[mol/s]	600 – 900
CH <sub>4</sub> Conversion in ATR ( $X$ )	[%]	80 – 95

This data set was partitioned into 80% training and 20% validation data to train the ALAMO and neural network surrogates and then gauge their accuracy. We consider that a surrogate is accurate if the coefficient of determination ( $R^2$ ) is greater than 0.8 for each of the 13 parity plots obtained from the validation dataset.

### ALAMO surrogate

The Gibbs reactor was replaced by a surrogate block containing simple algebraic equations determined by the ALAMO machine learning framework. To balance the bias-variance trade-off and calculate a model with low nonlinear complexity, the four basis models we consider are a quadratic, a cubic, a linear variable and a constant. We have the option to include bilinear terms and higher degree polynomials to achieve a higher surrogate accuracy. However, for the purpose of this research, convergence could be hindered with the inclusion of those terms, particularly if we observe that a linear combination of simple basis functions can effectively approximate the reformer's behavior. The surrogate model is composed of 13 equations and 3 variables, in contrast to the 55 equations and variables that model this first-principles Gibbs reactor. A subset of these 13 equations is displayed in Eq. (8). Training time to acquire this surrogate model was 1.6 s.

$$T_{out} = 8.2 \times 10^{-4} F_s + 0.41 F_{ref}^3 + 897.4 X \quad (8a)$$

$$F_{out} = 3.9 F_{ref} + 1.1 F_s - 7.9 \times 10^{-7} F_{ref}^2 + 685.0 X^2 \quad (8b)$$

$$x_{H_2} = 5.3 \times 10^{-4} F_{ref} - 1.5 \times 10^{-10} F_{ref}^3 + 0.14 X^3 \quad (8c)$$

$$x_{CO} = -6.2 \times 10^{-10} F_s^3 + 8.1 \times 10^{-4} F_{ref} - 4.1 \times 10^{-7} F_{ref}^2 + 0.2 X - 0.35 \quad (8d)$$

$$x_{CH_4} = 1.5 \times 10^{-5} F_{ref} - 6.0 \times 10^{-6} F_s - 0.33 X^2 + 0.16 X^3 + 0.16 \quad (8e)$$

### Neural network surrogate

The autothermal reformer was also replaced by a surrogate block containing a neural network. Hyperparameter tuning was performed to obtain the neural network with the lowest validation loss, quantified with the mean squared error. Hyperparameter ranges are shown in Table 2. The sigmoid and tanh activations were chosen because they are smooth functions, matching our setting

of nonlinear continuous optimization. The Adam optimizer was used for training.

**Table 2:** Hyperparameter ranges used to train the neural network.

Hyperparameter	Range/value
Activation function	Sigmoid & tanh
Number of layers	2 – 5
Number of neurons	20 – 35
Epochs	500

The training time to run every hyperparameter combination in Table 2 was 550 s. Ultimately, the neural network that best approximates the Gibbs reactor uses the tanh activation function and has 4 hidden layers with 32 neurons each. The time to train this neural network in isolation was 9.5 s. The optimization formulation uses the full space option provided by OMLT, as it was found to converge more reliably than the reduced space formulation. In the full-space formulation, variables and constraints corresponding to interior nodes in the neural network are explicitly included in the optimization problem. While only a single neural network surrogate is considered in this work, a comparison of optimization problems with many different architectures (and equation-based formulations) of embedded neural networks would be an interesting future study.

### Implicit function

The theoretical formulation given in Eq. (3) can be implemented as shown in Eq. (9).

$$\max f(x, u) \quad (9a)$$

$$\text{s.t. } G(x, u) \leq 0 \quad (9b)$$

$$H_{internal}(x, u) = 0 \quad (9c)$$

$$H_{linking}(x, y, u) = 0 \quad (9d)$$

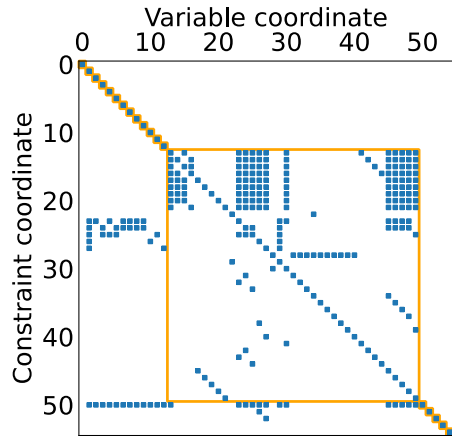
$$R(x, y, u) = 0 \quad (9e)$$

Here, Eq. (9e) solves for  $y$  externally as an implicit function  $y = R_y(x, u)$  using the PyNumero interface and the resulting values are communicated back to Eq. (9d), which is exposed to the NLP solver. In this case, Eq. (9d) is referred to as a set of linking equality constraints that link the externally obtained outlet variables  $y$  to the inlet variables of the reformer recuperator, which is downstream of the Gibbs reactor (See Figure 1). The dimension of  $y$  equals the dimension of  $R$  and  $\nabla_y R$  is non-singular.

The Gibbs reactor is replaced by an implicit function  $y = R_y(x, u)$  that solves the reactor equations as a parameterized system of equations. The PyNumero interface solves this system at every iteration of the nonlinear optimization solver with a decomposition that partitions variables and equations into block-lower triangular form

before solving the resulting blocks independently. The block triangular partition is computed by the approach of Duff and Reid [18] using the Incidence Analysis Pyomo extension [19]. The PyNumero interface also computes the derivatives required by IPOPT.

The system of equations for  $R_y(x,u)$  decomposes into 55 diagonal blocks, where 18 have dimension  $1 \times 1$  and one has dimension  $37 \times 37$ . The block triangular form of this system's incidence matrix is shown in Figure 2.



**Figure 2.** Incidence matrix of the square system corresponding to the Gibbs reactor.

Each independent system of equations is solved with SciPy's `fsolve`, a wrapper around MINPACK's implementation of Powell's hybrid trust region method [20].

## RESULTS

To compare convergence reliability, solve time and solution quality between the full space, implicit function, and surrogate-based formulations, we perform a parameter sweep varying the inlet natural gas absolute pressure and its conversion in the Gibbs reactor. We attempt to solve the optimization problem using IPOPT [2] for every combination of these two parameters for a total of 64 problem instances. These have identical initialization methods and solver options, where the maximum number of iterations are 300 and each must converge to a tolerance of  $10^{-7}$ . The code used to implement these formulations and reproduce the results can be found at <https://github.com/Robbybp/surrogate-vs-implicit>.

The convergence status for each formulation is shown in Figure 3. An unsuccessful run is due to the optimization solver reaching the iteration limit, converging to an infeasible point, or a failure due to repeated function evaluation errors. Here, a function evaluation error may be an error in a scalar-valued function, such as attempting to evaluate the logarithm of a negative number, or a more complicated error such as a failure to solve the square system that defines the implicit function  $R_y(x,u)$ .

A successful run indicates not only that the optimization problem converged, but also that the calculated input variables yield no constraint violations when used to simulate the full-space model.

In this experiment, the full space formulation was able to successfully converge 33 out of 64 instances, the implicit function formulation solved 52, the surrogate-based formulation using ALAMO converged 64 instances, and the surrogate-based formulation using a neural network converged 48. Regarding the full space formulation, many of these failed instances are due to large residuals in calculating energy balances in the autothermal reformer and the reformer recuperator.

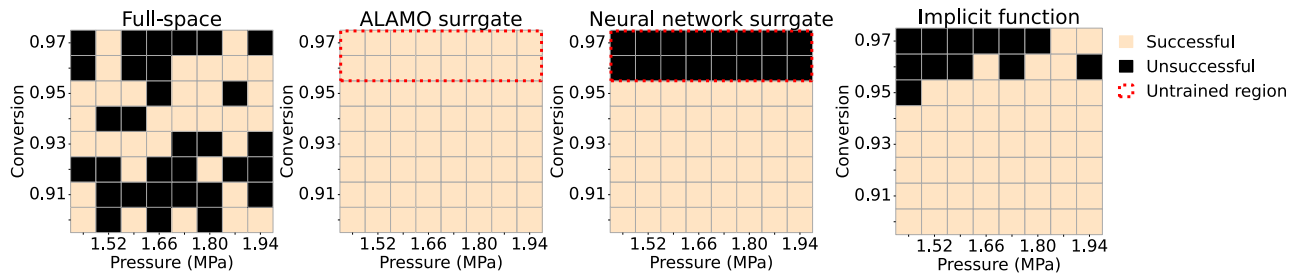
The implicit function formulation obtains the same objective and values for the manipulated inputs as the full space formulation but converged 58% more instances. The unsuccessful instances at higher conversion are due to function evaluation errors.

In the failing instances, the implicit function formulation experiences evaluation errors in the Gibbs minimization equations. Given the high conversion, the optimization algorithm calculates a value near zero for the outlet molar composition of propane in the reactor, which is used to calculate the entropy of the ideal gas. Here, one of the terms attempts to calculate the natural logarithm of this composition, causing numerical stability issues, as shown in Eq. (10). In consequence, the Newton solver fails to calculate  $g_{partial,C_3H_8}$ , see Eq. (7), and the outer optimization does not converge. The entropy of the ideal gas is used to calculate the ideal partial Gibbs energy of propane, which is then corrected with a departure function.

$$S_{C_3H_8}^0 = \int_{298.15}^T \frac{A+BT+CT^2+DT^3}{T} dT + \Delta S_{form}^{298.15} - R \ln \frac{P}{P_{ref}} - R \ln x_{C_3H_8} \quad (10)$$

The ALAMO and neural network formulations introduce only a minor increase in solution error. To evaluate solution quality, we solved a square system for the original flowsheet, where the fixed inlet natural gas, steam molar flow rate, and bypass fraction correspond to the values calculated by the ALAMO and the neural network formulations, separately. Then, we compared the objective value given by these simulations with the objective value calculated by the full space optimization problem. This comparison was done for the 25 instances for which full space, ALAMO, neural network, and implicit formulations all converged successfully. The average relative objective function difference for the ALAMO formulation was 1.9%, with a maximum difference of 2.1%. For the neural network formulation, the average relative objective function difference was 0.96%, with a maximum difference of 2.4%.

It is important to note that even though the ALAMO models were trained in the range of conversion of 80 to



**Figure 3:** Convergence status for each formulation. “Untrained region” indicates conversions above 0.95, the upper bound used for surrogate training data. The results indicate that while all three alternative formulations are more reliable than the full-space formulation, the ALAMO surrogate is the most reliable.

**Table 3:** Problem statistics for each formulation.

Formulation	N. Iterations	Avg. solve time (s)	Max. solve time (s)	Std. dev. (s)
Full Space	95	1.7	5.1	0.9
Implicit	46	1.5	3.2	0.4
ALAMO	30	0.5	0.6	0.03
Neural Network	51	0.8	0.9	0.05

**Table 4:** Qualitative results for each formulation applied to the autothermal reformer optimization problem.

Formulation	Solution accuracy	Solve time	Training time	Reliability
Full Space	High	Moderate	N/A	Low
Implicit	High	Moderate	N/A	Moderate
ALAMO	Moderate	Low	Moderate	High
Neural Network	Moderate	Low	High	Moderate

95%, they were able to find accurate solutions with conversions of up to at least 97%. We note that there is no increase in error when conversions above 95% are tested. However, this situation may not be replicated when modeling another process. The neural network surrogate formulation provides virtually the same solution quality as the embedded ALAMO polynomials. Nonetheless, in contrast to those surrogates, it fails to converge when the conversion is 96% or higher, that is, outside training bounds. One interpretation of these convergence failures is that the neural network surrogate has been over-trained to accurately represent the reactor model in the training region, while the comparatively simple and sparse surrogate generated by ALAMO has resisted this over-training. Additionally, hyperparameter tuning is required to obtain an accurate neural network, which is a computationally costly task.

A breakdown of problem statistics is shown in Table 3, where “N. Iterations” refers to the *average* number of iterations it takes for the optimization solver to converge and “Std. dev (s)” refers to the standard deviation of the average solve time from the successful instances. As before, these statistics were computed for the intersection of successful instances among the four formulations.

Despite relatively expensive function evaluations

that solve a square system at every iteration, the implicit function formulation converges slightly faster than the full-space formulation. This is because the implicit function formulation requires fewer IPOPT iterations to converge the optimization problem.

The standard deviation of solve time for the full space formulation shows that different flowsheet parameters have a significant impact on the optimization solver’s ability to converge. Conversely, the implicit and especially surrogate formulations exhibit more uniform solve times and converge for a higher percentage of instances, supporting the idea that they are robust and less sensitive to process parameter values.

Finally, a summary of qualitative results for the four formulations is shown in Table 4. Our qualitative assessment of training time includes time required for data generation and hyperparameter tuning. We note that surrogate training time could be further reduced by using more efficient sampling techniques, such as Latin Hypercube Sampling [21], which would potentially generate a smaller and more representative dataset of the entire experimental region, which might also lead to higher solution accuracies.

## CONCLUSIONS



We have presented four different formulations for optimization of a chemical process flowsheet using the IDAES modeling framework. The implicit function approach demonstrates an improved convergence reliability in contrast to the full space approach. The ALAMO surrogate formulation is the fastest and most reliable optimization alternative in this study that results in low solution errors for the objective value and the manipulated inputs. Nonetheless, the implicit function formulation may be preferred in cases where the design specifications of a process are not able to tolerate the introduction of errors into the calculations, or where producing enough simulation data to train an accurate surrogate model is computationally prohibitive. For instance, large chemical process flowsheets involving wastewater treatment units or gas scrubber systems, where environmental regulations specify a strict threshold for pollutant compositions in water or gas being released into the atmosphere, may not be appropriate for a surrogate formulation.

In this case study, the ALAMO surrogates were able to effectively approximate the behavior of a Gibbs Reactor. Nevertheless, more complex unit models with polar components, such as multi-component distillation columns, or stripping and absorbing columns, might require neural networks to provide a good approximation of the associated differential-algebraic systems. Finally, the implicit function and surrogate formulations presented in this paper constitute important approaches to optimize a chemical process flowsheet when the classical method fails, and their advantages would be more prominent in the design phase of complex, large scale chemical processes where superstructure optimization (MINLP) might be involved.

## ACKNOWLEDGEMENTS

We gratefully acknowledge the support of the U.S. Department of Energy through the Los Alamos National Laboratory (LANL) LDRD program and the Center for Nonlinear Studies (CNLS) for this work.

Los Alamos National Laboratory is operated by Triad National Security, LLC, for the National Nuclear Security Administration of U.S. Department of Energy (Contract No. 89233218CNA000001). This work is approved for unlimited release under LA-UR-23-31036.

## REFERENCES

1. Drud A. CONOPT: A GRG code for large sparse dynamic nonlinear optimization problems. *Math Program* 31:153-191 (1985).
2. Wächter A, Biegler LT. On the implementation of an interior-point filter line search algorithm for large-scale nonlinear programming. *Math Program* 106:25-57 (2005).
3. Mazzei MS, Mussati MC, Mussati SF. NLP model-based optimal design of LiBr-H<sub>2</sub>O absorption refrigeration systems. *Int J Refrig* 38:58-70 (2014).
4. Caballero JA, Grossmann IE. An algorithm for the use of surrogate models in modular flowsheet optimization. *AIChE J* 54:2633-2650 (2008).
5. Thombre MN, Preisig HA, Addis MB. Developing surrogate models via computer-based experiments. *12th International Symposium on Process Systems Engineering and 25th European Symposium on Computer Aided Process Engineering* 641-646 (2015).
6. Bhosekar A, Ierapetritou M. Advances in surrogate-based modeling, feasibility analysis, and optimization: A review. *Comput Chem Eng* 108:250-267 (2018).
7. Henao CA, Maravelias CT. Surrogate-based superstructure optimization framework. *AIChE J* 57:1216-1232 (2010).
8. Cozad A, Sahinidis NV, Miller DC. Learning surrogate models for simulation-based optimization. *AIChE J* 60:2211-2227 (2014).
9. Ma K, Sahinidis NV, Bindlish R, Bury SJ, Haghpanah R, Rajagopalan S. Data-driven strategies for extractive distillation unit optimization. *Comput Chem Eng* 167:107970 (2022).
10. Miller DC, Syamlal M, Mebane DS, Storlie C, Bhattacharyya D, Sahinidis NV, Agarwal D, Tong C, Zitney S, Sarkar A, Sun X, Sundaresan S, Ryan E, Engel D, Dale C. Carbon capture simulation initiative: A case study in multiscale modeling and new challenges. *Annu Rev Chem Biomol Eng* 5:301-323 (2014).
11. A. S. Subramanian, T. Gundersen, P. I. Barton, and T. A. Adams, "Global optimization of a hybrid waste tire and natural gas feedstock polygeneration system", *Energy*. 250:123722 (2022).
12. R. Parker, B. Nicholson, J. Sirola, C. Laird, and L. Biegler, "An implicit function formulation for optimization of discretized index-1 differential algebraic systems", *Comput. & Chem. Eng.* 168:108042 (2022).
13. A. Lee, J. H. Ghouse, J. C. Eslick, C. D. Laird, J. D. Sirola, M. A. Zamarripa, D. Gunter, J. H. Shinn, A. W. Dowling, D. Bhattacharyya, L. T. Biegler, A. P. Burgard, and D. C. Miller, "The IDAES process modeling framework and model library—Flexibility for process simulation and optimization", *J. Adv. Manuf. Process.* 3 (2021).
14. F. Ceccon, J. Jalving, J. Haddad, A. Thebelt, C. Tsay, C. D. Laird, and R. Misener, "OMLT: Optimization & machine learning toolkit", *J. Mach. Learn. Res.* 23:1-8 (2022).
15. M. L. Bynum, G. A. Hackebeil, W. E. Hart, C. D. Laird, B. L. Nicholson, J. D. Sirola, J.P. Watson, and

- D. L. Woodruff. Pyomo — Optimization modeling in Python. Springer Nature (2021).
16. D. M. Gay, “Hooking your solver to AMPL”, tech. rep., Computing Sciences Research Center (1997).
  17. J. S. Rodriguez, R. B. Parker, C. D. Laird, B. L. Nicholson, J. D. Siirola, and M. L. Bynum, “Scalable parallel nonlinear optimization with PyNumero and Parapint”, *INFORMS J. Comput.* 35:509–517 (2023).
  18. I. S. Duff and J. K. Reid, “An implementation of Tarjan’s algorithm for the block triangularization of a matrix”, *ACM Trans. Math. Soft.* 4:137–147 (1978).
  19. R. B. Parker, B. L. Nicholson, J. D. Siirola, and L. T. Biegler, “Applications of the Dulmage-Mendelsohn decomposition for debugging nonlinear optimization problems”, *Comput. & Chem. Eng.* 178:108383 (2023).
  20. M. Powell, “A hybrid method for nonlinear equations”, in Numerical Methods for Nonlinear Algebraic Equations. Ed: P. Rabinowitz. Gordon and Breach (1970).
  21. C. Kamath, “Intelligent sampling for surrogate modeling, hyperparameter optimization, and data analysis,” *Machine Learning with Applications*, vol. 9. Elsevier BV, p. 100373, Sep. 2022.

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# A Novel Cost-Efficient Tributyl Citrate Production Process

Andres F. Cabeza<sup>a,b</sup>, Alvaro Orjuela<sup>a\*</sup>, and David E. Bernal Neira<sup>b,c,d</sup>

a Universidad Nacional de Colombia, Department of Chemical and Environmental Engineering, Bogotá, Colombia

b Purdue University, Davidson School of Chemical Engineering, West Lafayette, IN, USA

c Universities Space Research Association, Research Institute of Advanced Computer Science, Mountain View, CA, USA

d Quantum Artificial Intelligence Laboratory, NASA Ames Research Center, Moffett Field, CA, USA

\* Corresponding Author: aorjuelal@unal.edu.co

## ABSTRACT

Phthalates are the most widely used plasticizers in the polymers industry; however, their toxicity and environmental impacts have led to their ban in various applications. This has driven the search for more sustainable alternatives, including biobased citrate esters, especially tributyl citrate (TBC) and its acetylated form. TBC is typically produced by refined citric acid (CA) esterification with 1-butanol (BuOH). However, the high energy and materials-intensive downstream purification of fermentation-derived CA involves high production costs, thus limiting the widespread adoption of TBC as a plasticizer. This work presents an innovative approach for TBC production using calcium citrate as feedstock instead of pure CA. The process involves a simultaneous acidification-esterification stage and further hydration of calcium sulfate, thus reducing costs by avoiding multiple CA refining steps. The approach proceeds via a solid-solid-liquid reaction of calcium citrate with sulfuric acid in butanol, releasing CA, which is simultaneously esterified to form TBC. The resultant calcium sulfate aids in water removal to enhance esterification conversion. Based upon experimentally validated models and rigorous simulations, the proposed approach was evaluated, and it exhibited significant reductions in processing times and operating costs, with savings of at least 46% in utilities compared to traditional TBC production. The novel approach was found suitable and promising for industrial deployment.

**Keywords:** Process integration, Process Intensification, Modelling and Simulations, Calcium citrate, Tributyl Citrate

## INTRODUCTION

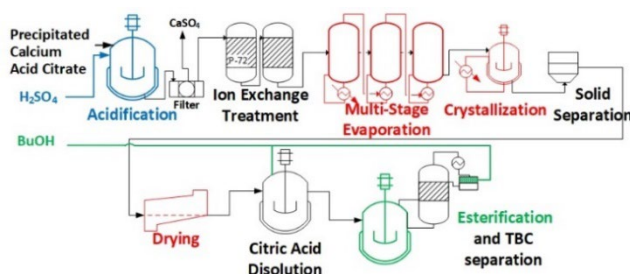
Polymeric materials and plastic products have significantly contributed to the advancement of modern society. Because of their malleability and flexibility, these materials have been used in many applications. These characteristics are mainly possible by incorporating plasticizers, which decrease molecular interactions among polymer chains, increase elasticity, and reduce the glass transition temperature (T<sub>g</sub>) (1). Plasticizers can constitute a significant fraction of some polymers (up to 80% wt.), being the phthalates, the most widely used general-purpose plasticizers. In particular, phthalates are mainly employed in polyvinyl chloride (PVC) products, including flexible films and sheets, hoses, pipes, windows, cable jacketing, flooring, rainwear, shoes, etc. Considering their significant use and that they are not chemical-bonded to the polymers, they can migrate from the material. This

has increased concerns regarding environmental pollution and public health due to the potential toxicity of plasticizers (2). In particular, the negative health impacts caused by phthalates has led to restrictions on their use in sensitive applications, prompting the search for alternative molecules that are non-toxic, renewable, and biodegradable. Among these, bio-based plasticizers like citric acid esters have gained attention.

Citric acid esters, particularly tributyl citrate (TBC) and its acetylated form have performed similarly to phthalates in different applications (3). However, citrates have the advantage of being biocompatible, biodegradable, and renewable. TBC is produced via esterification of citric acid (CA) with a large excess of 1-butanol (BuOH) using acid catalysts [4,5]. The process involves long processing times in a semi-batch operation and high-purity anhydrous CA (6). This last requirement is a major challenge because purification of fermentation-derived CA

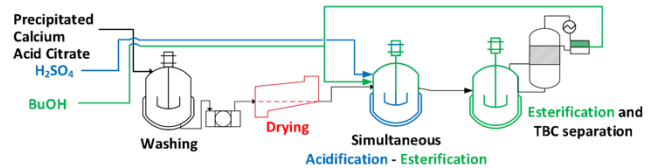
requires highly materials- and energy-intensive processes. As a result, the cost of TBC is highly dependent on that of CA, and it is usually higher than that of phthalates; this limits an extended industrial use of TBC.

Production of bioderived anhydrous CA requires a complex train of physicochemical separation operations. The fermentation broth is filtered to remove biomass and then treated with lime to precipitate CA as a calcium salt. The salt is then separated by filtration, resuspended in water, and finally acidified with sulfuric acid. At this point gypsum is produced and precipitated, releasing CA in the aqueous solution. This aqueous solution is then taken to ion exchange to remove minerals, followed by concentration via multistage evaporation, crystallization, filtration, and drying processes (7). This complex downstream separation train removes nearly 10 kg of water per kg of CA, thus being highly energy intensive. Finally, refined CA is subjected to esterification with BuOH. Anhydrous CA is required as raw material in the esterification process as it produces water, and the reaction conversion is limited by chemical equilibrium. The complete process is depicted in Figure 1.



**Figure 1:** Conventional TBC production process

As an alternative approach, this work proposed and assessed an innovative scheme to overcome most of the steps required to purify CA to produce TBC, by using calcium citrate as raw material. This process is similar to previous reports on the recovery of succinic acid (8) and involves a solid-solid-liquid reaction in which insoluble calcium citrate is suspended in an alcoholic medium (i.e., BuOH), where acidification with sulfuric acid is carried out. As a result, CA is released into the alcoholic medium, where it undergoes simultaneous esterification to TBC catalyzed by the remaining  $H_2SO_4$ . The proposed process, described in Figure 2, denoted as simultaneous acidification-esterification (SAE), avoids the conventional purification of CA. Additionally, a reduction in the dissolution times required in the traditional esterification reaction is expected. Furthermore, calcium sulfate is produced as a by-product in an alcoholic phase, reducing salts' solubility and helping to remove water by forming hydrated salts, thus facilitating the separation of esters.



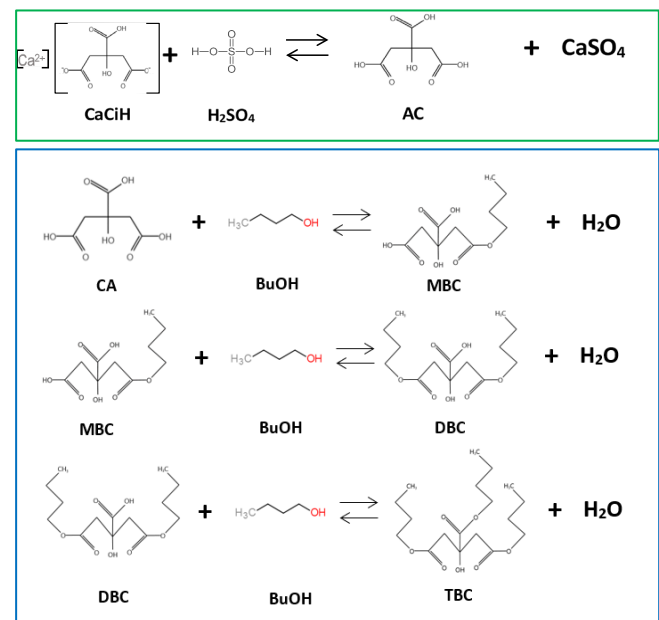
**Figure 2:** Proposed SAE TBC production process

To assess the proposed SAE process against the conventional one using anhydrous CA, rigorous simulation of both processes was carried out in Aspen Plus® using validated kinetic and thermodynamic models [5,10]. The operating conditions and the configuration of the conventional process were based on data from an existing industrial facility for TBC production.

## MODELING AND SIMULATION

### Reaction Kinetics: Conventional Process

In the conventional industrial process (Figure 1), the multistep esterification reaction of CA with BuOH (Figure 3b) is catalyzed using methanesulfonic acid (MSA). A validated kinetic model for the multiple sets of parallel-sequential reactions was previously reported (4) and used for simulations here.



**Figure 3:** Scheme of reactions during the proposed simultaneous acidification-esterification of calcium citrate in 1-butanol using sulfuric acid. a) (top) acidification. b) (bottom) esterification.

### Reaction Kinetics: SAE Process

As previously mentioned, the acidification occurs between solid calcium citrate and liquid  $H_2SO_4$  in an alcoholic media. Then, released CA undergoes esterification with BuOH (Figure 3). The remaining or excess sulfuric

acid acts as a catalyst in this reaction. As a result, a certain degree of conversion to intermediate citrates is achieved (i.e., monobutyl citrate - MBC and dibutyl citrate - DBC). This would facilitate the recovery of CA and reduce the time required for complete esterification [4, 10].

The corresponding kinetic model of the simultaneous acidification-esterification process was previously reported and was used to simulate the process (10). As the esterification to TBC is incomplete, this is finished in a subsequent reactor using MSA as catalyst (4). The kinetics of the acidification-esterification reaction and the corresponding kinetic parameters are presented in the supplementary material.

## TBC Production Conventional Process

The simulation of the conventional process was carried out considering the two main sections involved in the production of TBC. First, there is CA production from precipitated calcium citrate. This section relies upon highly energetic demanding separation stages (e.g., evaporation, crystallization, and drying). Then, the esterification stage, where CA is transformed into the desired plasticizer via a fed-batch operation. Here, reactants loading and operating temperature change follow a pre-programmed protocol as performed in the industry (11).

## Citric Acid Production

The details of the CA process and the production capacity (30,000 tons/year) corresponded to those of an existing industrial plant (11). For simplicity, it is assumed that all the commercialized CA corresponds to the anhydrous form. Approximately 95% of produced CA is sold directly; the remaining fraction is used to produce citrate esters. This small portion corresponds to the remnant CA that doesn't meet specifications regarding particle size or that is recovered from equipment walls like hoppers, filters, centrifuges, or dryers. CA is produced from submerged aerobic fermentation of sugarcane molasses using *Aspergillus niger*. The produced acid is released into the aqueous medium, so precipitation is carried out using  $\text{CaOH}_2$  to generate insoluble calcium citrate. This process separates the citric acid in salt form and allows its extraction from the culture medium via filtration. For comparative purposes, both the modeling of the conventional process and the proposed process considered the resulting stream of precipitated and filtered calcium citrate as the raw material in the TBC production process. As the precipitated calcium citrate corresponds to an intermediate stream in the CA process, its commercial price is unknown. This price was estimated at \$0.43/kg, based on the cost of the raw materials necessary to generate the calcium citrate. Although all high energy-intensive steps in the synthesis of CA were included in the model, polishing the acid and TBC using activated carbon absorption to remove colors and other impurities is not considered.

- **Acidification reaction:** The CA recovery from calcium citrate in an aqueous medium occurs by ion exchange with a stoichiometric amount of  $\text{H}_2\text{SO}_4$ . This solid-liquid reaction is carried out in a stirred reactor at 30 °C with a residence time of 3 hours. As a result, the stoichiometric amount of solid gypsum is produced, and the obtained liquid effluent corresponds to a 30% wt. CA aqueous solution.
- **Filtration:** Conveyor belts or rotary filters are commonly employed to separate the solid gypsum (i.e., calcium sulfate) from the effluent CA solution generated in the acidification process. The filters are connected to vacuum pumps to facilitate filtration and enhance liquid phase recovery. In the simulation a separator filter module is used to describe  $\text{CaSO}_4$  removal.
- **Evaporation:** the aqueous solution is processed in an evaporation train to concentrate up to the crystallization point of the organic acid. Since the process is energetically demanding, multiple interconnected stages are typically employed (1 to 4 steps, depending on the production capacity) to reduce steam consumption. Here, a triple-effect co-current configuration is considered for a capacity of 30,000 tons/year of CA (Figure 4). Multi-effect evaporators were simulated via Heater modules that serve as calandrias for the condensation of the steam, releasing latent heat. This heat is provided to flash vessel modules, where concentration of CA is performed. A vacuum profile in the effects is assumed (0.89, 0.67, and 0.40 bara). The final concentration of the solution depends on the desired hydrated form of CA in the crystallization stage.

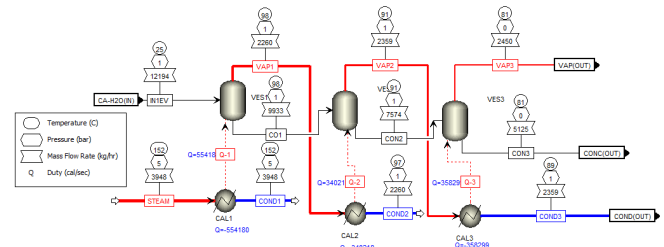


Figure 4: Triple-effect evaporation configuration

- **Crystallization:** CA can precipitate as anhydrous or monohydrate depending on the concentration and temperature. An operating temperature of 45°C is set to obtain anhydrous CA crystals. In this case, a concentration of  $\geq 70\%$  wt. of CA is required to ensure the saturation of the mixture at the operating temperature. Aspen Plus® Crystallizer model and CA solubility data in water (12) were used for the calculation of the saturation conditions in the crystallization modeling.



- **Centrifugation:** The use of two-stage centrifuges is common for the separation of crystallized solids. CFuge model that simulates centrifuge filters (rotating baskets), with no residual solids in the outlet filtrate, is used to separate crystallized CA from the liquor. A residual moisture of 10% by weight on a wet basis was assumed.
- **Drying:** Continuous fluidized bed dryers enable a low moisture content in the final product (< 5% wt.). Steam was used as heating utility to provide the required heat duty for the drying air that is in direct contact with the fluidized crystals. A conveyor belt coupled to a vibratory system allows better air-solid contact and prevents agglomerated solids. In this case, a convective dryer type in the dryer module is used, with a crossflow (through-the-bed) air-solid arrangement.

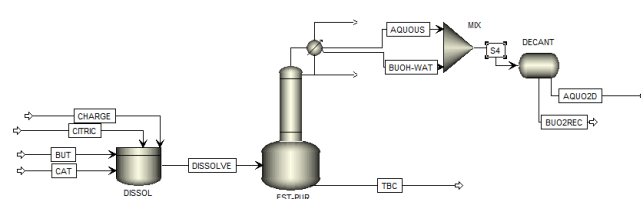
Vacuum was provided using a steam-jet ejector system. The steam required for the vacuum system can be calculated considering the air leakage in the vessels. This can be determined by knowing the operation pressure and the volume of the vessel (13).

### Citric Acid Dissolution-Esterification

Once anhydrous CA is obtained, it is sieve-classified, and the fines fraction is rejected according to market specifications. Additionally, it is common to find solid CA fines adhered to equipment or clumps in some units. First, this material is fed to a 23.7 m<sup>3</sup> dissolution tank alongside with BuOH, to enhance a total homogeneous reaction mix. Afterwards, the mix, together with the catalyst is fed to SS316 jacketed reactor with a capacity of 26 m<sup>3</sup> to perform the esterification reaction. During esterification, removal of produced water is desired to shift the chemical equilibrium. For this reason, the reactor is coupled to a distillation column that operates at the bubble temperature of the BuOH-H<sub>2</sub>O mixture. As the reaction proceeds, a vapor stream of BuOH-H<sub>2</sub>O is generated, and it is separated in the rectifying column up to the azeotropic condition. Since this azeotrope is heterogeneous, two immiscible phases are generated and separated in a decanter after condensing the column's top vapor. The BuOH-rich organic phase is returned to the column as reflux, whereas the aqueous phase is removed from the system to overcome chemical equilibrium limitations. The process of transforming CA into TBC can be divided into three stages:

- **Dissolution:** Due to the low solubility of CA, dissolution must be carried out by dosing the fed solids under heating and constant agitation. In this case, an excess of BuOH and the catalyst (MSA) are added to promote the reaction and to facilitate the dissolution of CA in the alcohol.

- **Esterification:** Once the necessary CA has been dissolved, the temperature of the reactive mixture is increased up to its bubble point to promote reaction. At this point, the vapor outlet line of the reactor is opened to connect with the distillation column, seeking to encourage water removal. The progress of the reaction is monitored through the measurement of the acid value of the mixture over time.
- **Purification:** Once the desired acid value is reached, separating the excess BuOH in the final mixture is necessary. For this, the system's pressure is reduced in the same reactor, and evaporation of all the remaining alcohol and water is carried out. Finally, the catalyst is neutralized using a sodium hydroxide solution.

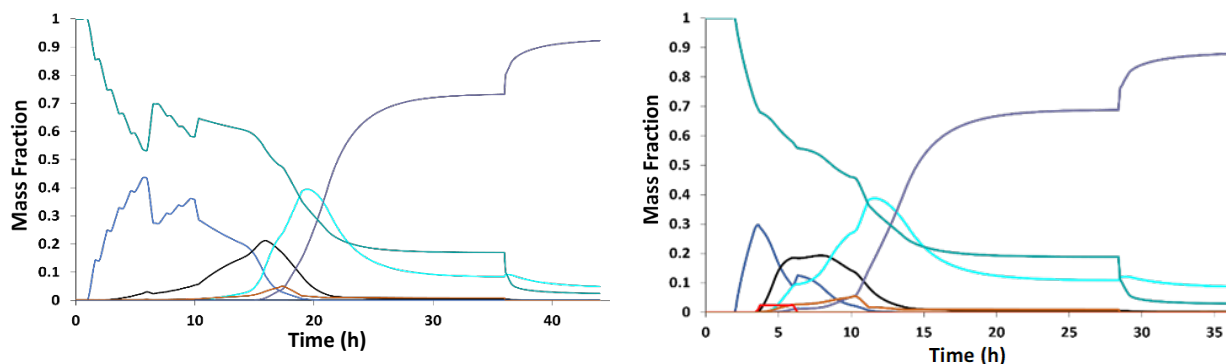


**Figure 5:** Batch and BatchSep module for esterification and acidification-esterification stage modeling.

Based on the above-described stages, the simulation of the citric acid dissolution-esterification stage was performed using both RBatch and BatchSep modules (Figure 5), enabling the dissolution of CA in BuOH, catalyst adding, temperature profile programming, and purification stages as reported in a previous study (11). The CA production is estimated considering a maximum processing capacity of 8t CA per batch. It is also assumed that 180 batches could be processed in a year. According to this, it is estimated that ~ 1440 t/yr of CA fines are generated in the production plant which are used for further esterification.

### Proposed SAE Process

Unlike the conventional process, the proposed process involves directly using calcium citrate from the precipitation stage with CaOH<sub>2</sub> and its subsequent acidification in BuOH, as seen in Figure 2. A pre-treatment stage is considered in which the calcium citrate is washed to remove impurities and ions from the culture medium. Subsequently, it is filtered and dried to avoid the presence of water in the reactive medium. For these pre-treatment stages, the acquisition of new equipment is not considered since these processes can be carried out



**Figure 6:** Mass fraction profiles **a)** conventional process, **b)** SAE process. (—) BuOH, (—) CA, (—) MBC, (—) DBC, (—) TBC, (—) H<sub>2</sub>O, (—) H<sub>2</sub>SO<sub>4</sub>

**Table 1:** Utility costs per batch in the production of TBC for the conventional and novel processes

Utility	Conventional process		Novel Process
	Citric Acid Production	Dissolution - Esterification	Acidification - Esterification
Steam	\$266.58	\$66.88	\$106.914
Electricity	\$107.09	\$244.27	\$209.41
Process Water	\$43.04	\$4.81	\$18.46
Cooling Water	\$2.28	\$0	\$6.15
Solid Waste Disposal	\$310.53	\$0	\$270
<b>TOTAL</b>	<b>\$1045.514</b>		<b>\$610.94</b>

within existing equipment in the CA plant (stirred tank, filters, and fluidized bed dryers).

- **Washing:** Contact with water (1 kg water / 1 kg calcium citrate) is done to remove impurities from the precipitated calcium citrate. A contact time of 1 hour is considered. The batch module is considered to perform the mixing.
- **Filtration:** Conveyor belt filters separate calcium citrate from the washing water.
- **Drying:** The remaining moisture from the calcium citrate is removed through the fluidized bed dryers. This ensures that no water enters the acidification-esterification stage.

Similar to the filtration and drying steps in conventional CA production, solid separator and crossflow convective dryer modules are used to simulate these operations in calcium citrate pretreatment. When the clean calcium citrate is obtained, it is fed into the simultaneous acidification-esterification process. This process is assumed to be carried out in the same dissolution vessel and esterification reactor used in the conventional process. Considering that, as experimentally verified in the solid-liquid reactive system, the solubility of the evolved citric species in

alcohol is rapid (10), the dissolution stage of CA in BuOH is not required. Therefore, the time of addition of the citric species will depend on the feeding capacity of the reactor. Then, a feeding rate of calcium citrate of 4126 kg/h is considered. This feeding rate can be achieved through hoppers coupled to screw conveyors. The mass of calcium citrate fed to the reactor (9583.33 kg) is equivalent to the same citrate moles considered in the previous section.

Once the calcium citrate is introduced into the BuOH, sulfuric acid is added to the reactor to promote the release of CA. A slight excess of H<sub>2</sub>SO<sub>4</sub> ensures complete acidification and facilitates the esterification reaction. The process is halted when the amount of MBC (Monobutyl Citrate) produced is equal to that generated in the dissolution stage of the conventional process. Once this point is reached, a stoichiometric amount of calcium citrate is added to neutralize the slight excess of sulfuric acid. This addition ensures that no residual H<sub>2</sub>SO<sub>4</sub> remains in the subsequent stages of the process.

Subsequently, the acidified and partially esterified mixture is filtrated to remove the calcium sulfate that is produced. It is assumed that the remaining solid fraction, discarded as waste, constitutes 4% by weight of the

reactive mixture. This assumption is made because separation in a filtration process is not 100% efficient. Therefore, two additional batches must be processed to achieve the same amount of TBC production. After filtering the reactive mixture, it is returned to the esterification reactor, and MSA is added. Once the esterification process is completed, the pressure within the system is reduced to remove the remaining BuOH and water. Operating conditions are presented in the supplementary Information.

Similar to the simulation of the CA dissolution and esterification in the conventional process, the novel process is performed using the RBatch and the BatchSep modules. These modules enabled the implementation of the simulation for the novel process and its specific operating policies (e.g., single feeds or scheduled feed configuration) using the same equipment. RBatch module enables modeling mixed, batch, and fed-batch operation conditions employed in the industrial production of TBC. On the other hand, the BatchSep module allows the modeling and scheduling of batch distillation columns when connected to the batch reactor.

Continuous water removal is possible because the more volatile compounds ( $H_2O$ -BuOH) are distilled off from the reactive media during the esterification when operating at the bubble point of the mixture. Nevertheless, BuOH removal is not desired since the reaction driving force would be diminished, striking esterification performance. Thus, assessment of the impact of simultaneous reaction and water removal is paramount to identifying the best operating policies to improve TBC yield.

With all previous considerations, we present a TBC production cost analysis focused on the utility costs of the conventional process concerning those of the SAE process. Capital costs were not considered because it is assumed that the same equipment is used in both processes and they are already available in the studied industrial facility. A composition profile of the two scenarios was merged, allowing the comparison with previous reports on the behavior of the conventional process. Also, a novel fed-batch operating policy developed during the experimental evaluation of simultaneous acidification-esterification (10) was tested and analyzed for the SAE process.

## RESULTS

### TBC Production Conventional Process

The mass concentration profiles over time during the production of TBC using the conventional industrial process are illustrated in Figure 6a. These profiles match those reported in a previous study (11) where a dilution stage ( $t < 18h$ ) enabled MBC production as the main citrate ester and accumulation of water is observed. Due to the low solubility of CA in the medium, alternating feeding

is required to ensure all the acid dissolves. A high temperature ( $\sim 60^\circ C$ ) is also necessary to accelerate the dissolution process.

Once the CA dissolution is completed, the temperature of the reactive medium is increased to the bubble point of the mixture ( $\sim 130^\circ C$ ). From this point, the production of DBC and TBC is increased and a vapor stream is generated, allowing for the removal of water from the top of the distillation column. However, there are differences compared to the previous report (11), as a higher concentration of DBC was observed in the reactor. This difference is caused by the remaining water in the reactive liquid, limiting the conversion towards TBC. Therefore, the final product will have a higher acidity content than required (acidity  $> 0.5\%$ ). The results coincide with what is reported for the industrial process, as the reactor effluent usually requires a subsequent neutralization stage to remove the remaining citric acidity and the catalyst. As observed, the process includes CA loading, dissolution, and esterification and the process lasts around 45 hours.

Regarding the costs derived from the utilities required in the production of TBC, it can be seen that the stage of obtaining anhydrous citric acid constitutes  $\sim 72\%$  of the overall operating costs. This result was expected considering the significant dilution of CA in the fermentation broth and the large amount of water to be evaporated from the CA crystals. Another substantial cost is the handling of solid waste generated in the acidification of the calcium citrate. Notably, a large amount of low-added value gypsum is produced during acidification in the existing industrial process. Although the obtained gypsum could be a salable by-product, purification to meet commercial specifications involves significant costs. So, in this case, this stream was assumed to be a waste and the disposal costs were estimated and internalized.

### TBC Production Proposed Process

As observed in Figure 6b, the direct feeding of calcium citrate reduces dissolution times to carry out the esterification reaction. Additionally, the recovery of the citric species can occur at lower temperatures (e.g.,  $40^\circ C$ ). Using an excess  $H_2SO_4$  ensures that the esterification process concurs as CA is released. As a result, the process with calcium citrate achieves a concentration of 18% wt. MBC in the liquid phase in 6 hours of reaction. In contrast, during the dissolution in the conventional process, 15 hours are required to achieve the same MBC content.

After the acidification-esterification stage, conventional esterification is carried out with MSA as a catalyst, and vacuum purification enables excess BuOH and remaining water removal (Figure 6b). As in the conventional process, achieving the purity required to commercialize the TBC would be difficult, and a neutralizing stage is

needed to remove the remaining acidity. While the process would be completed in 36 hours, the system seems to reach equilibrium in about 18 hours, considerably shorter than the 25 hours required for reaction in the conventional process, counting for a reduction by about 25% in reaction time to reach equilibrium condition.

Finally, as presented in Table 1, it is observed that the cost related to steam consumption for CA purification presented a higher difference. It can be seen that the steam-related costs are the lowest for the novel process compared to the conventional configuration related to the 68% reduction in steam usage per unit of citrate esterified in the SAE process. This is mainly attributed to the direct use of calcium citrate as the CA backbone carrier for TBC synthesis, avoiding all concentration stages commonly used in the purification of CA. The concentration stages in pure CA production (e.g., evaporation, crystallization, and drying) need a great amount of energy to evaporate close to 10 kg of water per kg of CA, to achieve the purity conditions necessary to perform esterification. As seen in Table 1, the cost associated with pure citric acid production accounts for ~70% of the utility costs in the conventional process. The other utility costs in the conventional process summarized in Table 1 do not show a massive reduction compared with the SAE process (mixing, vacuum, solid disposals, etc.) since these operations remain almost unaltered.

Despite both processes generating calcium sulfate, the one obtained from the novel process is in the form of a semi-hydrated form. This form may have potential as a valuable industrial by-product. Moreover, by crystallizing in an organic medium, the particle size distribution and the purity of the obtained crystal may generate commercial interest (14). The cost associated with the utilities used in the SAE process reflects cost savings of 41% compared to those of the conventional process. Such reduction allowed a 1.8% decrease in the TBC operating costs.

The SAE process could also be helpful in a possible continuous configuration for TBC production. As reported previously, continuous reactive distillation (RD) has arisen as a promising alternative for TBC production(5)(10). Since RD allows for the simultaneous esterification reaction and water removal, inherent equilibrium limitations are overcome. In these works, it has been identified that using a pre-reactor stage is beneficial to reduce RD column size or increase TBC production capacity (5). This pre-reaction stage could be performed via SAE processes. As described, this process has a fed-batch configuration. Also, some stages like calcium citrate feed and calcium sulfate removal would make it difficult to couple the SAE process with RD. Some strategies like equipment redundancy or using tanks for pre-reacted mixture storage, could help to overcome these issues. This configuration will be evaluated in future work.

## CONCLUSIONS

This work presented a comparative assessment of a conventional and innovative process for the industrial production of TBC. The novel SAE process employs calcium citrate as raw material instead of CA to produce citric acid esters. Based on experimentally established reaction kinetics and validated thermodynamic models, it was possible to simulate the solid-solid-liquid reaction occurring during the simultaneous esterification-acidification stage in the SAE process. Reduction in the utility costs and in the time to perform the esterification reaction could lead to a significant increase in TBC's production and capacity, alongside an increase in the economic viability of TBC (1.8% cost decrease due to operating costs) using the existing industrial infrastructure. Moreover, although not quantified in this work, the reduction in process units also leads to potential capital cost reduction. As a follow-up to this work, an optimization-based cost analysis will be carried out to minimize the total (fixed and operational) process costs for a new industrial facility using the novel SAE process. This analysis can also be done to retrofit an existing TBC production plant using the conventional process.

## ACKNOWLEDGMENT

D.B.N. was supported by the NASA Academic Mission Services, Contract No. NNA16BD14C. D.B.N. and A.F.C. acknowledge the support of the startup grant of the Davidson School of Chemical Engineering at Purdue University.

## DIGITAL SUPPLEMENTARY MATERIAL

The supplementary material can be found with the LAPSE ID: 2023.36926

## REFERENCES

1. Godwin AD. Plasticizers. Applied plastics engineering handbook. 353-553 (2017) <https://doi.org/10.1016/B978-0-323-39040-8.00025-0>
2. Wang Y, Qian H. Phthalates and their impacts on human health. *Healthcare* 9 (5) 603 (2021) <https://doi.org/10.3390/healthcare9050603>
3. Wypych G. (Ed.), Handbook of Plasticizers, ChemTec Publishing, Toronto (2004)
4. Osorio-Pascuas OM, Santaella MA, Rodriguez G, Orjuela A. Esterification Kinetics of Tributyl Citrate Production Using Homogeneous and Heterogeneous Catalysts. *Ind. Eng. Chem. Res.* 54:12534-12542 (2015) <https://doi.org/10.1021/acs.iecr.5b03608>
5. Santaella MA, Gutiérrez MF, Orjuela A. Tributyl citrate production via reactive distillation: Model

- reconciliation, optimization, scale up and sustainability indicators. *Chem. Eng. J.* 433:133199 (2022)  
<https://doi.org/10.1016/j.cej.2021.133199>
6. Fonseca JD, Latifi AM, Orjuela A, Rodríguez G, Gil ID. Modeling, analysis and multi-objective optimization of an industrial batch process for the production of tributyl citrate. *Comput. Chem. Eng.* 106603 (2020)  
<https://doi.org/10.1016/j.compchemeng.2019.106603>
  7. Kristiansen B, Linden J, Matthey M. Citric Acid Biotechnology. CRC press (1998)
  8. Orjuela A, Orjuela A, Lira CT, Miller DJ. A novel process for recovery of fermentation-derived succinic acid: Process design and economic analysis. *Bioresour. Technol.* 139:235–241 (2013)  
<https://doi.org/10.1016/j.biortech.2013.03.174>
  9. Santaella MA, Suaza A, Berdugo CE, Rivera JL, Orjuela A. Phase Equilibrium Behavior in Mixtures Containing Tributyl Citrate, Citric Acid, Butan-1-ol, and Water. *J. Chem. Eng. Data* 63:3252–3262 (2018) <https://doi.org/10.1021/acs.jced.8b00064>
  10. Cabeza AF. Synthesis of Bio-Plasticizers Derived from Citric Acid from Calcium Citrate (in Spanish)., National University of Colombia, Bogota. (2023)
  11. Osorio-Pascuas OM. Production Of Triethyl Citrate And Tributyl Citrate From The Acid Esterification Of Citric Acid With Ethanol And 1-Butanol (in Spanish)., National University of Colombia, Bogota. (2019)
  12. Nemdili L, *et al.* Crystallization kinetics of citric acid anhydrate. *J. Cryst. Growth* 451:88–94 (2016)  
<https://doi.org/10.1016/j.jcrysgro.2016.07.008>
  13. Gutierrez MF. Sucrose esters production in a solvent-free reaction system by transesterification of sucrose and fatty acid methyl esters., National University of Colombia. (2018)
  14. Tritschler U, Kellermeier M, Debus C, Kempter A, Cölfen H. A simple strategy for the synthesis of well-defined bassanite nanorods. *CrystEngComm* 17:3772–3776 (2015)  
<https://doi.org/10.1039/C5CE00519A>

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to the creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Graph-Based Representations and Applications to Process Simulation

Yoel R. Cortés-Peña<sup>a</sup> and Victor M. Zavala<sup>a\*</sup>

<sup>a</sup> University of Wisconsin Madison, Department of Chemical and Biomolecular Engineering, Madison, Wisconsin, United States

\* Corresponding Author: [zavatejeda@wisc.edu](mailto:zavatejeda@wisc.edu).

---

## ABSTRACT

Rapid and robust convergence of a process flowsheet is critical to enable large-scale simulations that address core scientific questions related to process design, optimization, and sustainability. However, due to the highly coupled and nonlinear nature of chemical processes, efficiently solving a flowsheet remains a challenge. In this work, we show that graph representations of the underlying physical phenomena in unit operations may help identify potential avenues to systematically reformulate the network of equations and enable more robust topology-based convergence of flowsheets. To this end, we developed graph abstractions of the governing equations of vapor-liquid and liquid-liquid equilibrium separation equipment. These graph abstractions consist of a mesh of interconnected variable nodes and equation nodes that are systematically generated through PhenomeNode, a new open-source library in Python developed in this study. We show that partitioning the graph into separate mass, energy, and equilibrium subgraphs can help decouple nonlinearities and guide decomposition algorithms. By employing the graph abstraction on an industrial separation process for separating glacial acetic acid from water, we implemented a new block decomposition scheme in BioSTEAM and demonstrated that this can accelerate convergence over a traditional sequential modular approach.

---

**Keywords:** Process simulation, Graph-Theory, Flowsheet Convergence, Distillation, Liquid Extraction

## INTRODUCTION

The evaluation of thousands of potential scenarios in a chemical process enables researchers to navigate uncertainties in market conditions and technological performance to create optimized designs, form conclusions on sustainability, and chart development pathways for new processing technologies. Automating the evaluation of such a large number of simulations is limited by computational challenges in rapid and robust flowsheet convergence. While many algorithmic paradigms exist (e.g., classical sequential modular simulation [1], parallel modular simulation [2], equation-based simulation [3], dynamic numerical methods [4], and the design of surrogate models [5,6]) only a limited set of approaches exist that are able to leverage the mathematical topology of the underlying phenomenological equations across the flowsheet [7,8].

In classical sequential-modular simulation, each unit operation is treated as a separate model with only

material streams as inputs and outputs. All mass, energy, and thermodynamic equations are formulated and solved independently within each unit operation. While specialized convergence strategies are employed to converge individual units, the convergence of recycle systems can be challenging due to nonlinear coupling between unit operations. Equation-based modeling leverages the sparsity of the full set of equations and employs algebraic differentiation to guide convergence of the entire system. Equation-based modeling may be faster than the sequential modular approach particularly when the initial guess is close to the steady state solution. However, highly coupled and complex networks of equations introduce instabilities and can lead to convergence failure [3].

If a process flowsheet could become aware of the connectivity of the governing phenomenological equations, robust solution strategies could be formulated that aggregate linear relationships and decouple nonlinearities. With the wealth of decomposition algorithms and computing architectures that enable the solution of

complex and large-scale problems, it may be feasible to screen for alternative problem formulations of a production process that result in more rapid and robust convergence. The problem formulation and the selection of a suitable solution method, however, is not a simple task. To address this challenge, we propose the use of graph-theoretic representations of the process equations and variables to better understand the topology of a chemical process at the *phenomenological level*. Leveraging graph abstractions in this manner can provide an avenue to systematically reformulate the equations and enable advanced, topology-aware decomposition schemes of flowsheets.

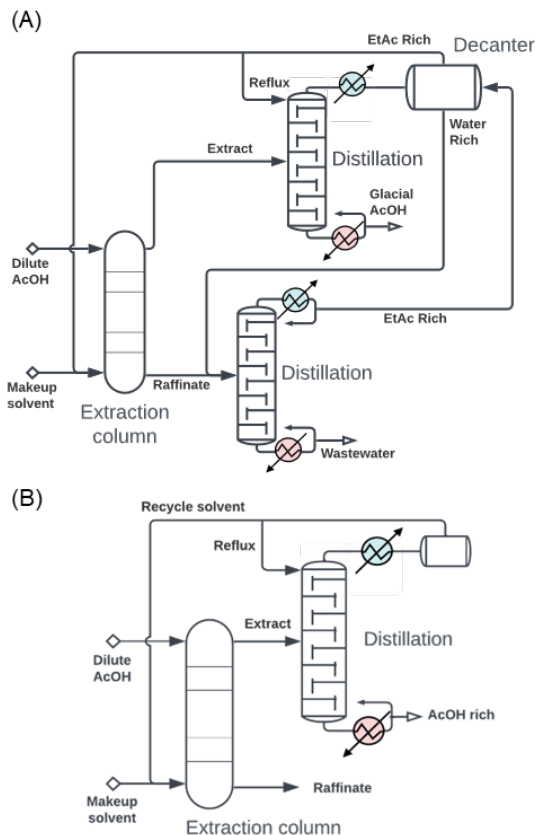
In classical representations of a chemical process, units are represented as nodes and streams as edges. Fundamentally, each unit encapsulates a set of equations with internal variables and each stream carries a set of variables with a unique connection. At the phenomenological level, however, each variable may describe multiple equations and may not have a unique connection. For example, temperature may play multiple roles in thermodynamic phase equilibrium, energy balances, and reaction rates. Graph abstractions of process phenomena may unfold the sets of equations and variables present in unit operations to capture how mass, energy, equilibrium, reaction, and transport equations are related through common variables.

A potential use of graph abstractions is the identification and development of robust decomposition algorithms whereby a subset of variables is decoupled to solve for sets of equations iteratively. Distillation column models are a classical chemical engineering example where decomposition algorithms are commonly employed. For example, the Wang-Henke bubble point method converges all stages by iteratively solving mass, equilibrium, summation and enthalpy (MESH) equations [9]. In fact, MESH partitioning within an equation-oriented approach has already been integrated to solve distillation trains within an equation-oriented approach [7]. At the flowsheet level, however, no unified approach for decomposition exists yet.

In this study, we developed a graph abstraction framework for process phenomena that can be used to systematically represent, analyze, and visualize the structure of a chemical process. To automate the generation of phenomenological graphs of complex chemical processes, we developed PhenomeNode, an open-source library in Python. We leverage PhenomeNode to form graph representations of an industrial separation process for separating glacial acetic acid from water and develop a preliminary convergence algorithm that leverages the mathematical topology. We benchmark this new algorithm against classical sequential modular simulation to understand its benefits and limitations.

## MATERIALS & METHODS

### Benchmark Acetic Acid Purification Process



**Figure 1.** Process flowsheet for the industrial separation of acetic acid from water.

The industrial separation of acetic acid from water (Figure 1A) was chosen as a representative case to benchmark the new convergence algorithm against sequential modular simulation. The system includes highly coupled vapor-liquid equilibrium (VLE) and liquid-liquid equilibrium (LLE) stages and recycle loops that connect the end of the process with the start. In this process, ethyl acetate (EtOH) is used to extract acetic acid from a dilute aqueous mixture [10]. The extract is distilled to recycle the solvent and recover glacial acetic acid. The raffinate is also distilled to recover EtOH from the wastewater. Both distillates are sent to a decanter to separate the aqueous phase that forms after condensation. Additionally, a smaller system composed of a liquid extraction column and a distillation column for solvent recovery (Figure 1B) is used as a benchmark for small, highly coupled systems.

The phenomena graph includes only the equations pertinent to flowsheet convergence, including duties and

material flows across unit operations and stages. Equations for sizing and costing of unit operations are not included as they can be directly calculated based on the steady state results. Process simulations were performed 10 times for each configuration to accurately measure convergence errors and material and energy balance errors as a function of computation time. The convergence error for a given variable at iteration  $k$  is defined as the absolute difference between the current value and the last value (i.e.,  $X_{\text{error}} = |X_k - X_{k-1}|$ ). The material balance error is defined as the difference between outlet flow rates and inlet flow rates (i.e.,  $F_{\text{error}} = \sum_c |\sum_o F_{c,o} - \sum_i F_{c,i}|$ ). The energy balance error is defined as the difference between outlet enthalpy and inlet enthalpy divided by the outlet heat capacity flow (i.e.,  $T_{\text{error}} = |\sum_o H_o - \sum_i H_i| \cdot (\sum_o C_o)^{-1}$ ).

## Phenomenological Graph Architecture

The graph abstractions in this study represent both variables and equations as nodes. Edges between variables and equations provide information on the variables present in each equation. Variables in the nodes are labeled by a shorthand name and their context. The context differentiates variables by information such as the phase and the parent unit operation. For example, the vapor mol fraction within a vapor-liquid equilibrium (VLE) stage may be labeled as  $z_{\text{gas}, s=0}$ , where “ $z$ ” denotes the mol fraction, “gas” denotes the phase, and “ $s$ ” denotes the stage number. Table 1 lists the subscript contexts and Table 2 lists the variable definitions used in this study. The “inlet source” refers to a neighboring stage that feeds the reference stage. Similarly, the “outlet sink” refers to a neighboring stage that is fed by the reference stage. It is possible to obtain different graph representations by reformulating the equations. For example, variable nodes can be eliminated via substitution of equations and new edges may be formed to reflect these changes. Aggregating equations in this manner may be helpful to represent a particular convergence algorithm. The graph representations in this study employ as many variables as possible to provide a starting point for future studies to strategically collapse nodes and/or partition the graph.

**Table 1.** Contexts of variable subscripts.

Subscript	Context	Subscript	Context
c	Chemical	gas	Gas
p	Phase	liq	Liquid
i	Inlet source	ext	Extract
o	Outlet sink	raf	Raffinate

Drawing, positioning, and labeling nodes for a large network of unit operations is a time-consuming task. To automate the systematic generation of graphical abstractions for complex chemical processes, we developed a new open-source Python library called

PhenomeNode [11] that leverages the Graphviz software [12] to generate graphical abstractions from the phenomenological equations within unit operations. In the spirit of expanding the use of graph representations at the phenomena level—not just at level of unit operations or superstructures of units—the PhenomeNode software developed in this study is made open-source and readily available for others to leverage. The roadmap for PhenomeNode includes expanding the built-in equation library, refining core features for aggregation and decoupling of equations, and creating educational tools for student learning of the underlying phenomena behind production processes.

**Table 2.** Variable definitions.

Variable	Definition
F	Molar flow rate
H	Enthalpy flow rate
Q	Duty
V	Molar fraction of vapor phase
$\Phi$	Ratio of top and bottom phase flow rates
L	Molar fraction of extract or liquid phase
T	Temperature
P	Pressure
Z	Molar composition
K	Partition coefficient
C	Heat capacity rate
t	Top phase split fraction
b	Bottom phase split fraction
h	Specific molar enthalpy
f	Fugacity

## Mathematical Relationships within Graphs

The equations describing the physical phenomena within the graph abstractions are agnostic to the thermodynamic property package. Equations for pure component and mixture properties are abstracted by black-box function calls. For example, the VLE criteria is represented by equation 1, where  $f_{c,\text{gas}}$  and  $f_{c,\text{liq}}$  are function calls to calculate the gas and liquid fugacities, respectively, for an arbitrary chemical. Other key equations include the Rachford-Rice, material balance, energy balance, and pressure drop (eqs 2–5, respectively).

$$f_{c,\text{gas}}(z_{c,\text{gas}}, T, P) = f_{c,\text{liq}}(z_{c,\text{liq}}, T, P) \quad (1)$$

$$\sum_c \frac{z_c(K_{c,\text{gas}} - 1)}{1 + V * (K_{c,\text{gas}} - 1)} = 0 \quad (2)$$

$$F_{c,\text{liq}} \left( 1 + \frac{V}{1-V} K_{c,\text{gas}} \right) - F_c = 0 \quad (3)$$

$$F_{\text{liq}} \left( \frac{V}{1-V} h_{\text{gas}} + h_{\text{liq}} \right) - H = 0 \quad (4)$$

$$P_{o=0} = P_{o=1} - \Delta P \quad (5)$$

A liquid-liquid equilibrium (LLE) stage has two liquid phases, “ext” and “raf”, which correspond to the extract

and the raffinate phases, respectively. The material and energy balance equations for an LLE stage are analogous to a VLE stage. The pressure drop across LLE stages is neglected assuming that liquid fugacities are a weak function of pressure. LLE criteria (eq 6) requires that the total Gibb's free energy of all phases is minimized.

$$\min_{z_{c,raf}} G(F_c, z_{c,raf}, z_{c,ext}, T, P) = z_{c,raf} \quad (6)$$

These are the fundamental equations for distillation and liquid-liquid extraction algorithms. However, they can be reformulated to fit a specific solution scheme. For example, in the case of LLE, the enthalpy is a stronger function of temperature than the phase fraction and the energy balance is formulated in terms of temperature (assuming an average heat capacity flow for each phase and a reference temperature and enthalpy):

$$(H_{ext}^{ref} + H_{raf}^{ref}) + (T - T^{ref})(\bar{C}_{ext} + \bar{C}_{raf}) - H = 0 \quad (7)$$

### Phenomena-Oriented Simulation Algorithm

We developed a new flowsheet convergence algorithm that expands and integrates MESH-based multistage equilibrium algorithms together with the sequential modular approach. Additional details on the algorithm architecture and performance are discussed in the RESULTS & DISCUSSION section. The material and energy balance equations are used to tie in all the phenomenological equations together. A stage includes all mixing and splitting of inlet and outlet streams, respectively, without limitation on the number of connected stages. For a given stage,  $F_c$  for all streams can be computed through the material balance equations (8–11):

$$\Sigma_o F_{c,o} - \Sigma_i F_{c,i} = 0 \quad (8)$$

$$\Phi K_{c,gas} \Sigma_o F_{c,liq,o} - \Sigma_o F_{c,gas,o} = 0 \quad (9)$$

$$F_{c,gas,o} - t_o \Sigma_o F_{c,gas,o} = 0 \quad (10)$$

$$F_{c,liq,o} - b_o \Sigma_o F_{c,liq,o} = 0 \quad (11)$$

The mass balance equations apply to both LLE and VLE stages, with either extract and raffinate or gas and liquid phases. Note that subscripts  $i$  and  $o$  refer to upstream and downstream stages, respectively, which are directly connected to the reference stage. For computational efficiency, the energy balance equation solves for the change in linearized variables (i.e.,  $T$  for a LLE stage and  $\Phi$  for a VLE stage) in each iteration (e.g., we solve for  $\Delta T$  where  $T_{i+1} = T_i + \Delta T$ ). Combining equations 4 and 7 (and taking  $\Delta T$  and  $\Delta \Phi$  as the linearized variables) the energy balance for an adiabatic VLE stage becomes:

$$\Delta \Phi h_{gas} \Sigma_o F_{liq,o} - \Sigma_i C_{liq,i} \Delta T_i - \Sigma_i \Delta \Phi_i h_{gas,i} F_{liq,i} = \Sigma_o H_o - \Sigma_i H_i \quad (12)$$

Similarly, the energy balance for an adiabatic LLE stage becomes:

$$\Delta T \Sigma_o C_{liq,o} - \Sigma_i C_{liq,i} \Delta T_i - \Sigma_i \Delta \Phi_i h_{gas,i} F_{liq,i} = \Sigma_o H_o - \Sigma_i H_i \quad (13)$$

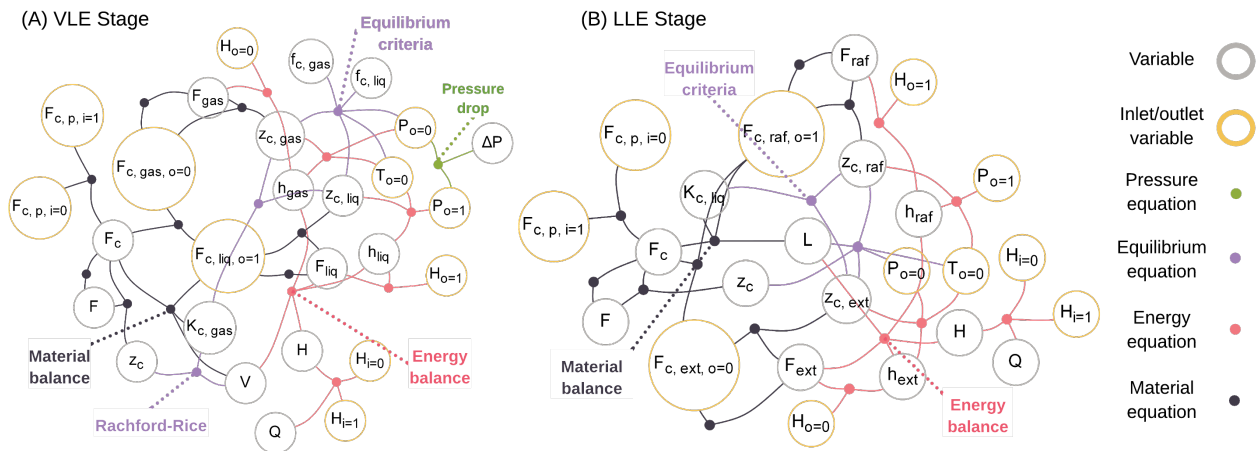
Given the coupling between the material and energy balances through  $F_c$  and  $\Phi$ , solving for  $T$  at each stage may help close the energy balance for large systems with strong compositional dependencies on specific enthalpies:

$$\Delta T \Sigma_o C_o - \Sigma_i C_i \Delta T_i = \Sigma_o H_o - \Sigma_i H_i \quad (14)$$

The VLE criteria (equation 1) is used to solve for  $K$  and  $T$ . The equilibrium temperature of an LLE stage, however, is not unique for a bulk composition. Instead, the LLE criteria (equation 6) is used to solve for  $K$  and  $\Phi$ . In the special case of multistage LLE, where equilibrium is highly sensitive to bulk compositions, the pseudo-equilibrium concept is employed for more rapid and robust convergence of the equilibrium criteria across all stages [13]. Given the equilibrium criteria variables and initial guesses for  $\Phi$  and  $T$ , component flow rates across all stages can be linearly solved by the material balance equations (8–11), and  $\Phi$  of VLE stages and  $T$  of LLE stages can be linearly solved by the energy balance equations (12–13). However, an algorithm employing solely equilibrium, material, and energy equations may fail due to poor conditioning of the initial guess. For this reason, we integrate sequential modular simulation—a robust approach used by leading process simulators—within the simulation algorithm for more robust convergence. All together, we propose the following algorithm as a preliminary architecture:

### Phenomena-oriented simulation algorithm

1. In the absence of an initial guess for  $F_c$ ,  $K$ ,  $T$ , and  $\Phi$  for each stage, run each unit operation sequentially to find initial guesses.
2. For each unit operation:
  - 2.1 Solve the unit operation and update variables  $K$ ,  $T$ , and  $\Phi$ .
  - 2.2 Solve for  $F_c$  across all unit operations as a system of linear equations (eqs 8–11) and update.
  - 2.3 Solve  $\Delta \Phi$  for each VLE stage and  $\Delta T$  for each LLE stage as a system of linear equations (eqs 12–13) and update  $\Phi$  and  $T$ .
3. Solve for the equilibrium criteria variables at each stage. For multistage LLE, solve for all stages simultaneously using the pseudo equilibrium approach.
4. Run steps 2.1, 2.2, then 2.1 again to close the material balance.
5. Solve for  $\Delta T$  as a system of linear equations (eq 14) and update  $T$  to close the energy balance.
6. If partition coefficients, phase ratios, temperatures, and flow rates have not converged under a specified tolerance, repeat steps 2–6.



**Figure 3.** A phenomena graph representation of (A) a VLE stage and (B) a LLE stage. Equations related to material, energy, equilibrium, and pressure are colored black, red, purple, and green, respectively. Each grey ring node and its label represents a variable. Edges denote the variables present equations.

This algorithm was implemented in BioSTEAM —an open-source process simulation software [14]—and benchmarked against sequential modular simulation using a flowsheet for the industrial separation of acetic acid (also modeled in BioSTEAM). The thermodynamic property package estimates phase equilibrium using modified Raoult’s law with activity coefficients estimated through Dortmund UNIFAC interaction parameters [15,16]. Pure component properties (e.g., heat capacity) of fluids are estimated using higher order polynomial fits to the fundamental Helmholtz equation of state (a state-of-the-art property prediction model)[17,18] and recommended correlations from a critical review on thermodynamic properties [19]. Mixture properties are estimated using a molar weighted average of the pure chemical properties.

## RESULTS & DISCUSSION

### Single and Multistage Equilibrium Graph

Material, energy, equilibrium, and summation equations are tightly coupled for both a VLE stage and a LLE stage (Figures 3 A and B, respectively). Note that the designation of what constitutes an equilibrium, material balance, or energy balance equation is based on the unit dimensionality of the equation as well as the decomposition scheme requirements. For example, the Rashford-Rice equation may be regarded as a material balance equation, but it is denoted here as an equilibrium relationship to maintain linear relationships within the material balance (assuming partition coefficients and phase fractions are decoupled).

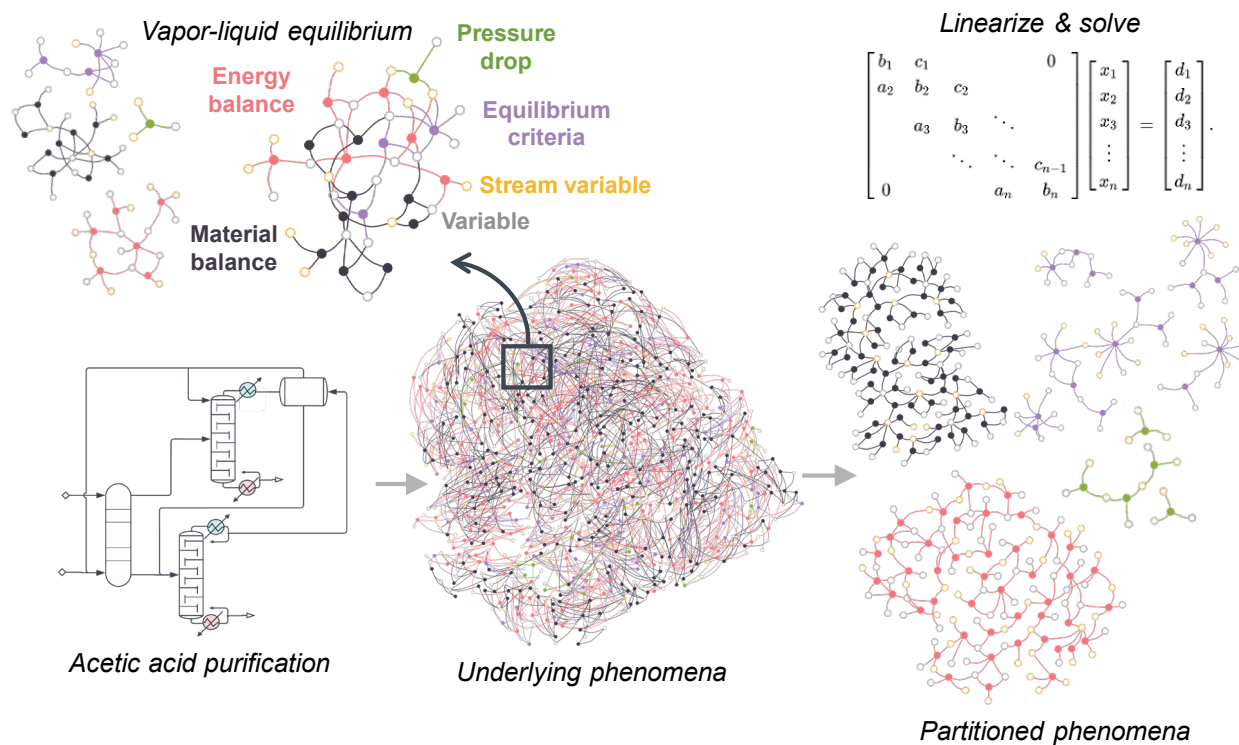
A classical chemical engineering method to solve an adiabatic flash vessel consists of a 3-step fixed-point iteration: (1) compute partition coefficients and temperature assuming the vapor and liquid compositions (e.g., by computing the bubble point), (2) estimate the vapor and

liquid compositions through material balances, and (3) solve for the vapor fraction using an energy balance [20]. In this iterative method, the equilibrium, energy, and material balances are decoupled. Given that non-linearities stem from the equilibrium criteria and the enthalpy of streams, this decomposition algorithm may offer greater stability than derivative-based numerical methods when employed for larger systems.

Partitioning the equations by equilibrium, energy, and material phenomena results in a more manageable problem formulation that is often used by published algorithms for multistage VLE and LLE. Assuming the temperature, pressure, and phase fraction (i.e., vapor or extract fraction) are held constant, the material balance for an arbitrary component becomes a linear combination of liquid (or vapor) flow rates and can be represented as a tridiagonal matrix that is convenient to solve. Similarly, if bulk liquid flow rates, compositions, and temperature are held constant, the energy balance becomes a linear combination of the boil-up ratio. The new phenomena-oriented simulation algorithm proposed in this study leverages this partitioning scheme towards the complete flowsheet to accelerate convergence.

Block convergence methods for distillation (e.g., Wang-Henke’s bubble point method) and liquid-liquid extraction —which iteratively solve equilibrium, material, and energy equations— are still subject to instabilities and may not converge. Russell’s inside-out method leverages approximate models for stage temperature and specific enthalpy to converge an inner loop of stage temperatures and phase flow rates [21]. Employing approximate models for highly coupled non-linear variables (i.e., partition coefficients, enthalpy, and temperature) allows for more robust convergence if the approximate form is better behaved than the strict form. Future work may





**Figure 5.** A phenomena graph representation of an industrial process for the separation of acetic acid from a dilute aqueous mixture to produce glacial acetic acid depicts how decoupling material (black), energy (red), and equilibrium (purple) equations may result in a more manageable formulation even for complex recycle systems. Equation nodes for repeated adiabatic stages in the partitioned phenomena are aggregated for clarity. Variable nodes are shaped as silver and gold rings to denote intermediate variables and stream variables.

seek to reformulate the system of equations with approximate models that can describe the complete system and propose more advanced numerical methods to accelerate flowsheet convergence.

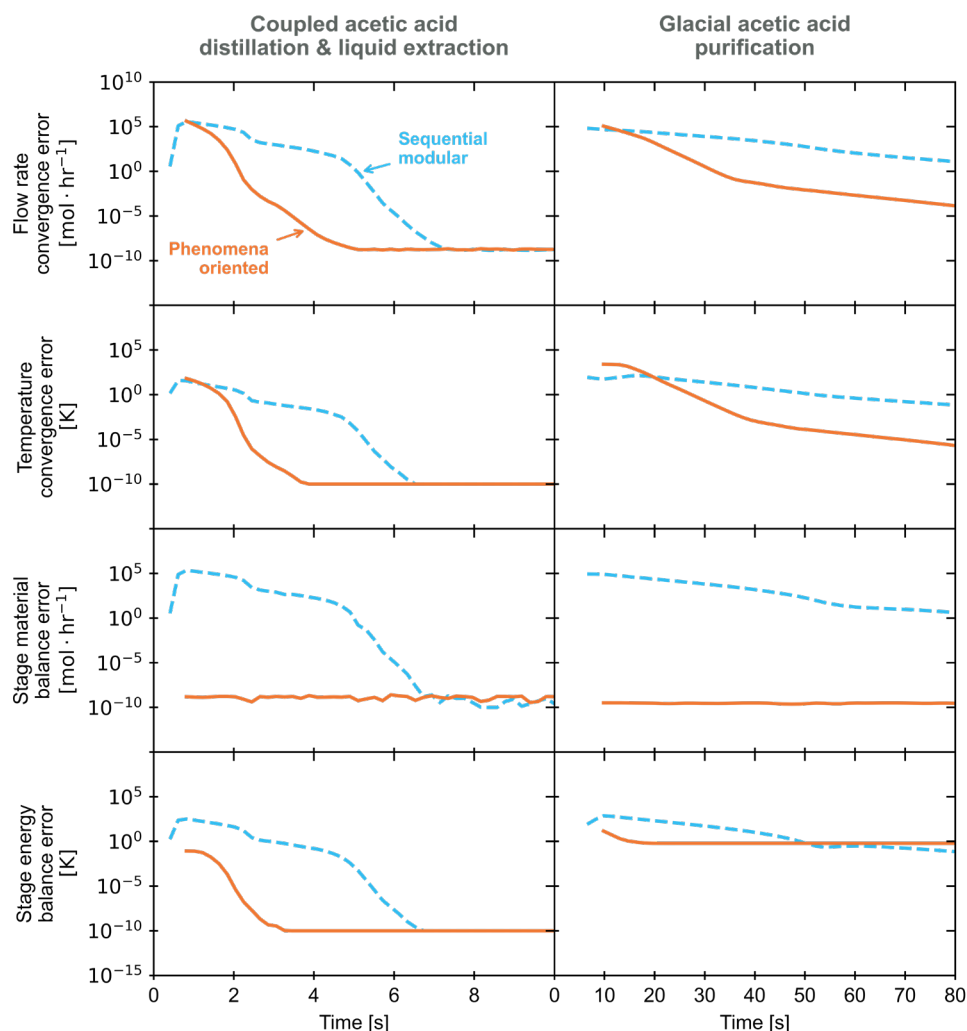
### Acetic Acid Separation Graph

A graph of an industrial separation process for separating glacial acetic acid from water suggests that it may be feasible to extend MESH equations as a decoupling strategy for integrated convergence of the complete flowsheet (Figure 5). While the graphical representation may seem more coupled due to the recycle loops, the resulting problem is essentially the same as in multistage equilibrium. In the proposed phenomena-oriented algorithm, equilibrium variables of each stage are calculated in parallel while energy and material balances result in sparse matrices that are linearly solved (steps 3–5). It is possible recycle loops between distillation columns and the liquid-liquid extraction destabilizes stage-wise material and energy balances and lead to ill-conditioned matrices and infeasible results (e.g., negative flows) which would break the model. To resolve such issues, we integrate sequential modular simulation as a robust method to rigorously estimate equilibrium criteria variables (step 2.1). The system-wide mass and energy balances (steps

2.2 and 2.3) helps propagate the information of how an individual unit simulation impacts the flowsheet, enforce mass and energy balances, and accelerate simulations.

### Phenomena-Oriented Simulation Speed

In both the benchmark cases, the preliminary phenomena-oriented simulation algorithm converged faster than the sequential modular approach (Figures 6 A, B, C, and D). The error in the mass and energy balance per stage was quickly minimized in the phenomena-oriented algorithm (Figures 6 E, F, G, and H) due to how mass and energy balances of the complete flowsheet are consolidated at the end of each iteration (steps 4 and 5). In comparison, sequential modular simulation has difficulties in closing the mass balance for the complex system even after 60 seconds of simulations (Figure 6H). By speeding up flowsheet convergence, phenomena-oriented simulation may enable large-scale simulations necessary for robust optimization and rigorous uncertainty and sensitivity analyses. It may be possible, however, that the proposed phenomena-oriented algorithm may not be as robust as sequential modular in other systems and more cases should be evaluated for more generalizable insight. Additionally, only one initial condition (starting with empty



**Figure 6.** The convergence error of the flow rate (A, B), temperature (C, D), as well as the error in the material balance (E, F), and the energy balance (G, H) was evaluated across computation time for a simple configuration with liquid-liquid extraction and distillation and a more complex system for acetic acid purification (left and right columns, respectively). The blue curve represents values using the sequential modular approach and the orange curve represents values using the new phenomena-oriented approach.

streams) was tested yet simulations used in optimization and uncertainty analyses may leverage more informed guesses. Future studies may seek to analyze the robustness of this new convergence algorithm more rigorously through Monte Carlo methods. It is possible certain algorithms scale better with large, sparse systems than with small, highly coupled systems.

It is important to note that this separation process is limited to only phase-based separations and does not encapsulate reactions and mass and energy transfer-based unit operations that are common in a production process. Still, the equations and decoupling strategies leveraged here can be extended to include other physical phenomena. For example, equilibrium approaches to reactive distillation extend Russell's inside-out method to include reaction terms and decouples the kinetic terms

from the equilibrium criteria [22]. We envision a family of phenomena-oriented simulation architectures that are optimized for different problems based on the relevant phenomena of the system, level of connectivity between unit operations, and chemical interactions that drive coupling between phenomenological equations. Ultimately, the phenomena graph representations introduced in this study may aid in the development of unified decoupling and linearization strategies that can allow for more rapid and robust flowsheet convergence.

## CONCLUSIONS

Efficiently solving a flowsheet remains a challenge due to the tightly coupled, non-linear nature of chemical processes. While a variety of methods exist for

converging a flowsheet, there is no clear winner in terms of speed, flexibility, and robustness. The phenomena graph representations developed in this study helped identify how decoupling equations into material, energy, and equilibrium blocks (and potentially other phenomena such as reactions) can be used to formulate a strategy to converge the complete flowsheet. While the proposed algorithm showed promising advantages in convergence speed composed to sequential modular, further analysis is needed to fully characterize the speed, robustness, and applicability to a broader set of chemical processes. The PhenomeNode library in Python can be used as an educational tool to help students visualize a chemical process not just as a flowsheet of unit operations, but also as a coupled network of physical phenomena.

## ACKNOWLEDGEMENTS

This material is based upon work supported by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Bioenergy Technologies Office under Award Number DE-EE0009285.

## REFERENCES

- Motard, R. L., Shacham, M. & Rosen, E. M. Steady state chemical process simulation. *AIChE Journal* **21**, 417–436 (1975).
- Mahalec, V., Kluzik, H. & Evans, L. B. Simultaneous modular algorithm for steady-state flowsheet simulation and design. *Computers & Chemical Engineering* **3**, 373 (1979).
- Bogle, I. D. L. & Perkins, J. D. Sparse newton-like methods in equation oriented flowsheeting. *Computers & Chemical Engineering* **12**, 791–805 (1988).
- Tsay, C. & Baldea, M. Fast and efficient chemical process flowsheet simulation by pseudo-transient continuation on inertial manifolds. *Computer Methods in Applied Mechanics and Engineering* **348**, 935–953 (2019).
- McBride, K. & Sundmacher, K. Overview of Surrogate Modeling in Chemical Process Engineering. *Chemie Ingenieur Technik* **91**, 228–239 (2019).
- Quirante, N., Javaloyes-Antón, J. & Caballero, J. A. Hybrid simulation-equation based synthesis of chemical processes. *Chemical Engineering Research and Design* **132**, 766–784 (2018).
- Ishii, Y. & Otto, F. D. Novel and fundamental strategies for equation-oriented process flowsheeting. *Computers & Chemical Engineering* **32**, 1842–1860 (2008).
- Ishii, Y. & Otto, F. D. An alternate computational architecture for advanced process engineering. *Computers & Chemical Engineering* **35**, 575–594 (2011).
- Monroy-Loperena, R. Simulation of Multicomponent Multistage Vapor–Liquid Separations. An Improved Algorithm Using the Wang–Henke Tridiagonal Matrix Method. *Ind. Eng. Chem. Res.* **42**, 175–182 (2003).
- Seader, J. D., Henley, E. J. & Roper, D. K. *Separation Process Principles, 3rd Edition*. (John Wiley & Sons, Inc., 2011).
- Cortes-Pena, Y. PhenomeNode: Graphical Representations of Process Phenomena.
- Ellson, J., Gansner, E., Koutsofios, L., North, S. C. & Woodhull, G. Graphviz— Open Source Graph Drawing Tools. in *Graph Drawing* (eds. Mutzel, P., Jünger, M. & Leipert, S.) vol. 2265 483–484 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2002).
- TSUBOKA, T. & KATAYAMA, T. General design algorithm based on pseudo-equilibrium concept for multistage multi-component liquid-liquid separation processes. *Journal of Chemical Engineering of Japan* **9**, 40–45 (1976).
- Cortes-Peña, Y., Kumar, D., Singh, V. & Guest, J. S. BioSTEAM: A Fast and Flexible Platform for the Design, Simulation, and Techno-Economic Analysis of Biorefineries under Uncertainty. *ACS Sustainable Chem. Eng.* **8**, 3302–3310 (2020).
- Cortés-Peña, Y. Thermosteam: BioSTEAM's Premier Thermodynamic Engine. *JOSS* **5**, 2814 (2020).
- Gmehling, J., Kleiber, M., Kolbe, B., Rarey, J., & WILEY-VCH. *Chemical Thermodynamics for Process Simulation*. (2019).
- Lemmon, E. W. & Tillner-Roth, R. A Helmholtz energy equation of state for calculating the thermodynamic properties of fluid mixtures. *Fluid Phase Equilibria* **165**, 1–21 (1999).
- Bell, I. H., Wronski, J., Quoilin, S. & Lemort, V. Pure and Pseudo-pure Fluid Thermophysical Property Evaluation and the Open-Source Thermophysical Property Library CoolProp. *Ind. Eng. Chem. Res.* **53**, 2498–2508 (2014).
- Zábranský, M., Kolská, Z., Růžička, V. & Domalski, E. S. Heat Capacity of Liquids: Critical Review and Recommended Values. Supplement II. *Journal of Physical and Chemical Reference Data* **39**, 013103 (2010).
- Boston, J. F. & Britt, H. I. A radically different formulation and solution of the single-stage flash problem. *Computers & Chemical Engineering* **2**, 109–122 (1978).
- Russel, R. A flexible and reliable method solves single-tower and crude-distillation-column problems. *Chem. Eng.* **90**, 53–59 (1983).
- Wang, L., Sun, X., Xia, L., Wang, J. & Xiang, S. Inside-Out Method for Simulating a Reactive Distillation Process. *Processes* **8**, 604 (2020).

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>

# Improved Design of Flushing Process for Multi-Product Pipelines

Barnabas Gao<sup>a</sup>, Swapana Jerpoth<sup>a</sup>, David Theuma<sup>a</sup>, Sean Curtis<sup>a</sup>, Steven Roth<sup>a</sup>, Michael Fracchiolla<sup>a</sup>, Robert Hesketh<sup>a</sup>, C. Stewart Slater<sup>a</sup>, Kirti M. Yenkie<sup>a\*</sup>

<sup>a</sup> Rowan University, Department of Chemical Engineering, Glassboro NJ 08028, United States

\* Corresponding Author: yenkie@rowan.edu

## ABSTRACT

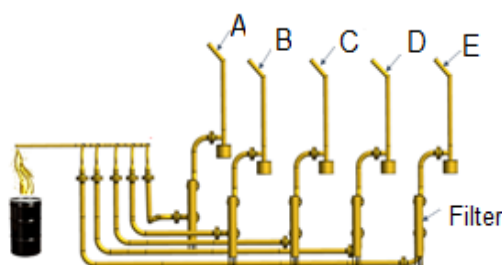
Maintaining product integrity in multi-product oil pipelines is crucial for efficiency and profit. This study presents a strategy combining design and process improvement to enhance flushing protocols, addressing the challenge of residual batch contamination. A pilot plant, mirroring industrial operations through dimensionless residence time distribution, was developed to identify and rectify bottlenecks during product transition. The pilot plant's success in replicating industrial operations paves the way for targeted experiments and modelling to enhance optimized flushing, ensuring product quality and operational excellence.

**Keywords:** Process Design, Optimization, Modeling, Flushing

## 1. INTRODUCTION

Multi-product pipelines are large-diameter lines used to transport different grades of the same product or different petroleum products. Such transport is performed by batching the products in a continuous succession [1]. Owing to this mode of transport, ensuring the quality and purity of each product processed in this network of pipelines is crucial for operational success [2]. Multi-product pipelines, tailored for various applications, are comprised of an array of components. These interconnected pipeline networks typically include straight segments, flow meters, valves, assorted fittings, bends, tees, and various other supplementary equipment.

Transporting different grades of products in multi-product pipelines often leads to residue accumulation, impacting the quality of subsequent products. Residues from previous batches adhere to the pipeline walls, compromising product yield, and integrity, and causing financial losses. Therefore, regular cleaning is essential during product changeovers in the multi-product pipeline industry. This study aims to develop an enhanced and comprehensive flushing process specifically for the generic lube oil industry.



**Figure 1:** Multi-Product Pipeline Configuration of the Generic Lube Oil Industry (The illustrated network details the sections of variable diameters and A, B, C, D, E represents sample different grades of products processed within this industry).

In this industry, various product grades are processed using multi-pipeline systems, as shown in Fig 1. These systems consist of straight and varying diameter sections, flow control mechanisms, U-bends, tees, fittings, and filters. To remove residues from previous products, pipelines are flushed between each product transition. This cleaning process involves two alternatives: using a pipeline inspection gauge (PIG) for straight sections and using the next product in line to flush the variable

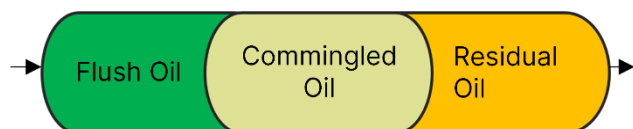
diameter sections. The PIG, a squeegee-like device, is 97% effective in cleaning the straight pipelines but less efficient for variable-diameter sections due to its size, shape limitations, and potential damage risks. Therefore, sections that cannot be cleaned by the PIG, known as "un-piggable sections," are flushed with the finished product (the next new oil to be packaged in the line).

This creates an optimization challenge to determine the volume of new oil needed for residual displacement. Therefore, we have designed and constructed a 1/5<sup>th</sup> scale pilot plant to replicate industrial processes, focusing on optimizing the flushing operation in the unpiggable sections of the multi-product pipeline system.

## 1.1 Flushing

Flushing in multiproduct pipelines is a crucial process involving the removal of residual products from the pipelines to prevent contamination when transitioning between several grades of products. This operation is particularly important considering systems where different grades of products are processed, as it ensures product quality and integrity [3]. Considering a changeover operation where a new product is to be processed through the pipeline, residues of the previous products need to be displaced. Referencing the generic lubricating oil industry with the pipeline configuration shown in Figure 1, entrapped residual oil within the unpiggable sections is flushed with the next product.

This process results in the creation of mixed or commingled oil systems, which are lower in value and considered downgraded. Consequently, large quantities of the subsequent product must be used to flush out the residuals and the commingled oil. As depicted in Figure 2, during the flushing operation, the tail ends of the pipeline contain the residuals, and the middle section is a blend of both oils, which are gradually and completely displaced over time. In sections of pipelines where pigging is not feasible, significant residue accumulation is often observed in U-bends, filter areas, and along the entire inner walls of the pipelines.



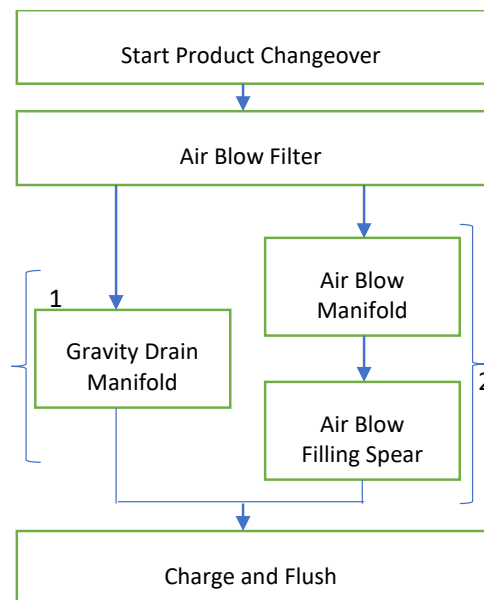
**Figure 2:** Illustration of Flushing Process During Product Changeover Operations

### 1.1.1 Review and Improvement of Industrial Flushing Operations

To enhance the flushing process in pipeline operations, it is crucial to analyze existing systems, identify inefficiencies, and devise improvements. This approach includes reviewing current practices, suggesting enhancements, and incorporating these into scaled pilot

plant designs to create effective flushing models. Before switching products in the pipeline, it's standard to remove the residue of the previous product. This is typically done by air blowing through filters and gravity draining the manifold line, effectively removing a significant portion of the residual oil. However, this method doesn't fully clear oil remnants in other parts of the pipeline, resulting in the use of excessive volumes of the next product to flush out these residuals. This leads to the formation of large quantities of mixed oil, highlighting a key area for process improvement.

To overcome this limitation, an enhanced process has been developed, combining experimental analysis with optimization techniques. This new approach aims to significantly lower the volume of residual materials, ensuring that only minimal amounts of finished products are employed in the flushing process. This reduction minimizes the creation of mixed oils. The improved methodology is depicted in Figure 3 contrasting it with the current operational mode, highlighting the key differences. This advanced process, tested, and refined through pilot plant studies, is designed to mirror and optimize industrial-scale operations.



**Figure 3:** Schematic Diagram of Flushing Operation. The existing operation is #1 and improved procedure is shown in #2

## 2. PILOT PLANT DESIGN AND ARCHITECTURE

The pilot plant aims to replicate the industrial scale operation to study the product changeover operations and optimize the flushing operation. Also, with this pilot plant the effects of changes in the process parameters like flowrate, viscosity of products, and temperature



is studied. The pilot plant is designed to mirror the generic lube oil industry on a smaller scale. The pilot plant features a detailed pipeline system where lube oil passes through the various components, including a filter, U-bends, manifold, and flow control devices, before being packaged. The process flow diagram of the designed pilot plant is illustrated in Figure 8 (the bottom figure).

## 2.1 Scale-Down Factors

An appropriate scale down factor must be used when designing and fabricating a pilot plant to replicate the characteristics of a much larger operational system.

### 2.1.2 Reynolds Number

The study concentrates on the mixing characteristics of viscous fluids in pipeline systems, using Reynolds Number as the scaling criterion for both systems. Matching the Reynolds Number ensures that the pilot plant accurately replicates the mixing behavior of an incompressible fluid. Considering an incompressible fluid,

$$N_{Re} = \frac{\rho v D}{\mu}$$

The Reynolds number,  $N_{Re}$  influenced by density, velocity, and viscosity, is a crucial factor. The pipeline at the lubricating industry is divided into two segments based on internal diameter (ID). The first segment, running from external tanks to the manifold, uses 3-inch schedule 40 carbon steel pipes with an ID of 3.068 inches. The shorter second segment comprises 2-inch schedule 40 carbon steel pipes with an ID of 2.067 inches. In contrast, the pilot plant's pipeline has been scaled down to a 1/2-inch schedule 40, with an ID of 0.602 inches. The product's viscosity plays a significant role in determining the dynamics during product change-over, affecting both the Reynolds number and the residual product amount in the system. Since the pilot plant operates at room temperature, it is necessary to adjust product viscosities to match the higher temperatures seen in full-scale operations. Additionally, the system's volumetric flow rate is crucial. Choosing a pump with the right flow properties is essential for accurately replicating the pilot plant's design at scale.

### 2.1.3 Volumetric ratio

The second metric used in the scale down process is the system volume ratio. This ratio is defined as the volume of the filter by the volume of the remaining system. Replicating this ratio for the pilot plant system was a crucial aspect for the design. The pilot plant system was modelled by theorizing the configuration as a combination of tanks (CSTR) and tubular (PFR) sections. The filter acts as the tanks and the pipe sections the tubular sections. The system volume at the industrial scale was determined to be approximately 32 gallons, with the filter containing 13 gallons of product. Thus, given the

expression for this ratio as follows;

$$V_R = \frac{V_{filter}}{V_{system}}$$

the ratio is calculated to be 0.4.

## 2.2 Designed Pilot Plant

The pilot plant features three product tanks, each with a 30-gallon capacity, serving distinct functions: a tank for residual oil, another for flush oil (the next product in line for processing), and a third for collecting the mix of these two oils. The tanks, made from high-density polyethylene for compatibility with petroleum lubricants and water, are set on a steel stand with rounded bottoms and bulkhead fittings for efficient drainage. They are linked to the main pipeline via flexible buna-nitrile hydraulic hoses. The plant also includes a suite of equipment: measuring balances placed under each tank for precise product tracking, spear gear pumps, a pressure relief valve, a compressed air system, a filter, various fittings and valves, an inline viscometer, and a sophisticated data acquisition system. This design ensures effective monitoring and management of the processing activities.



**Figure 4:** Pilot Plant CAD Design (Top) and Fabricated Pilot Plant Design at the Laboratory (Bottom)

When selecting an electronic balance for our system, key factors included its maximum weight capacity, size, response speed, and cost. Considering the density

of water and weight of tank stand, a balance capable of measuring up to 300 pounds was required for accurate safety testing. This was vital for monitoring the subtle, rapid weight changes during pump operation. Ultimately, we chose a balance with a 600-pound capacity and a width of 19.75 inches, ensuring it not only met our immediate needs but also accommodated larger loads for future experiments.

Due to the high viscosity of the fluid, a peristaltic (spear gear) pump is used for fluid flow in the pipeline, with a maximum capacity of 7 GPM. This pump features a cast iron interior and 3/4" NPT ports reduced to 1/2" outer diameter tubing. To ensure safety and prevent overpressure, a pressure relief valve (PRV) is installed, set at 70 PSI - above the normal operating pressure but below the filter's 100 PSI maximum. This PRV setting accommodates the system's needs, particularly considering the performance of lubricant oils up to 108 cSt, where the maximum pressure drop through the valve is 25 PSI, within the filter's safety threshold.

Viscosity is a critical factor in lubricating oil, essential for optimal machinery performance. Choosing an oil with appropriate viscosity for specific operating conditions is vital. Viscosity tests are key to ensuring product quality and the success of flushing operations. Traditional viscosity measurement involves labor-intensive sampling and lab analysis using glass capillary viscometers. However, this method can be time-consuming and error-prone. Automating the viscosity testing process can reduce manual labor and costs, while real-time viscosity measurement in product batches offers environmental and logistical advantages by minimizing repeated flushing operations. An inline viscometer, integrated into the pilot plant system, demonstrates the advantages of real-time measurement over traditional methods. This viscometer features a seamless bore, no moving parts, and requires minimal maintenance.

Integrating a data acquisition system into the pilot plant was essential for accurate, real-time recording of experimental data, crucial for understanding the process being studied. This system facilitates efficient collection, processing, and analysis of data from various equipment, including scales, inline viscometers, and thermocouples. It plays a key role in monitoring and controlling experimental variables, greatly aiding in system modeling and enabling swift decision-making. The selected data acquisition system features an eight-slot chassis, accommodating multiple input modules, chosen based on compatibility with the pilot plant's equipment, signal measurement types, maximum sampling rates, and the supported number of channels. The signal measurement type, the interacting equipment and the individual acquired modules are illustrated in Table 2.

**Table 1:** Pilot Plant Process Equipment Specifications

Equipment	Description	Capacity
<b>Spear Gear Pump</b>	Cast Iron, 3/4" NPT	7 GPM
<b>Inline Viscometer</b>	Flow-through viscometer, no internals and compactible with pigging systems	0 to 10,000 cP, accuracy +/- 0.1 cP, 100 bar, -50 °C – 100°C
<b>Pressure Relief Valve</b>	Buna-nitrile seal with stainless steel material	11 GPM, 60-175 PSIG

**Table 2:** Components of the Data Acquisition System, Measurement Signal Type and Interacting Process Equipment

Components	Measurement Signal	Interacting Equipment
<b>9212</b>	Current (4-20 mA)	Thermocouple
<b>9870</b>	Voltage (9 V DC)	Balance
<b>9203</b>	Current (4-20 mA)	Inline Viscometer

### 3. PILOT PLANT VALIDATION

In comparing two systems, which are of different scales, it is useful to assess the dimensionless residence time distribution between both systems. If the scaled down version (pilot plant) exhibits similar behavior and characteristics to the larger system (industrial plant), it indicates that the smaller system effectively replicates the larger system [4]. This comparison is crucial for validating and ensuring the pilot plant can reliably mimic the industrial system, providing a basis for scaling up processes with confidence.

#### 3.1 Residence Time Distribution

The prediction and modeling of residence time distributions (RTDs) are essential for understanding material flow in a process. Each component, such as filters and pipes, has its own RTD. By linking these RTDs using convolution integrals, the overall process RTD is calculated. This helps in estimating the average time materials spend in the process, understanding the system's response to material stream fluctuations, and devising process control strategies [5].

Residence time distribution modeling serves not only to characterize entire manufacturing lines but also to explain the intricate behaviors of individual units in a network. In reactor networks, the continuous stirred tank reactor (CSTR) and plug flow reactor (PFR) are the most prevalent types. While these reactors are overly idealized to accurately represent real reactors, combining these

basic models in a network allows for an effective depiction of actual unit operations, accounting for phenomena like dead zones and areas of insufficient mixing [6].

The residence time distribution (RTD) experiments in the pilot plant involved using diluted salt water and fresh water. These experiments are crucial as they bridge theoretical concepts with the pipeline system's practical application. The step change method was employed for these RTD experiments, reflecting the complexities of the flushing process. This approach effectively demonstrates the fluid dynamics and various mixing mechanisms operational within the pilot plant.

By comprehending the complex fluid flow dynamics in the pilot plant, we can fine-tune the system for a more precise emulation of industrial processes.

### 3.2 Methodology

The RTD experiment involved using diluted salt water as the flushing fluid and pure water as the residual fluid. The experiment began with a sudden switch from one fluid to the other at  $t = 0$ , and the outlet concentration was monitored. To simulate oil flushing operations where oils are miscible and compactible, water-soluble salts were used. NaCl was selected as the preferred salt, as indicated in Table 3. The conductivity of these diluted salt solutions was measured using a Vernier Conductivity Probe™, substituting the inline viscometer in the pilot plant. The recorded conductivities (S) were then converted to concentrations (C) using relations shown below.

$$C = aS^2 + bS + c$$

**Table 3:** Candidate salts and parameter values for conductance to concentration conversion

Salt	mS/cm	a	b	c
NaCl	0-55	1.95	609	0
KCl	0 - 19	0	631.32	0
	19 - 143	0.2816	708.07	-913.51
KI	0 - 11	0	1340.4	0
	11 - 113	0.7556	1449	-233.9
NaBr	0 - 16	5.6327	967.2	0
	16 - 100	2.4823	1067.2	44.46

### 3.3 Governing Equations

The first step in RTD modelling is obtaining the residence time distribution. The experimental setup introduces a conductive solution as a tracer and its concentration measured at the outlet. The concentration profile of the tracer is  $C(t)$ . The cumulative RTD,  $F(t)$  is obtained as below.

$$F(t) = \frac{C(t)}{C_{in}}$$

$$F(t) = \int_0^t E(t) dt$$

The  $E(t)$  curve describes the distribution of exit times. Its peak indicates the time where most of the tracer material is discharged.  $E(t)$  is known as the RTD function. Statistical indicators (mean residence time  $\tau$ , standard deviation  $\sigma$ ) can be evaluated directly from the  $E(t)$  curve.

$$E(t) = \frac{d(F(t))}{dt} = \frac{d}{dt} \left( \frac{C(t)}{C_{in}} \right)$$

The mean residence time,  $\tau$  is computed as below.

$$\tau = \frac{\int_0^\infty tE(t)dt}{\int_0^\infty E(t)dt} \cong 1$$

$$\tau = \int_0^\infty tE(t)dt$$

Normalised RTD function,  $E(\theta)$  is the metric used in the comparison of the pilot plant system with the industrial plant.

$$E(\theta) = \tau E(t)$$

### 3.4 Cubic Splines

Cubic Spline is piecewise defined curve that provides a balance between flexibility and smoothness. Splines utilizes interpolation and smoothing to fit a curve through data points [7]. Cubic splines are designed to have continuous first and second order derivatives, ensuring a smooth transition between each polynomial segment. The smoothness is particularly useful in requiring a smooth curve without abrupt changes in slope.

From the python SciPy Library, spline is defined as below.

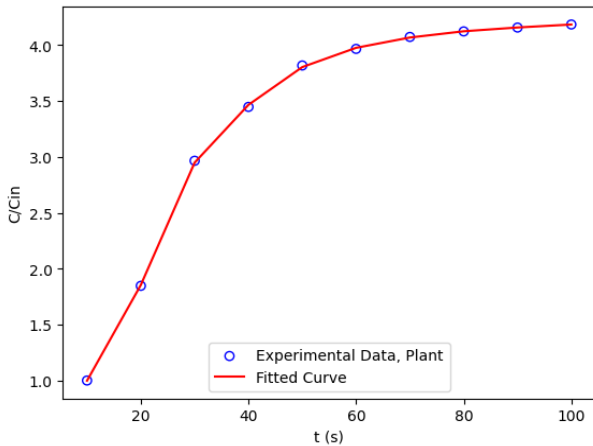
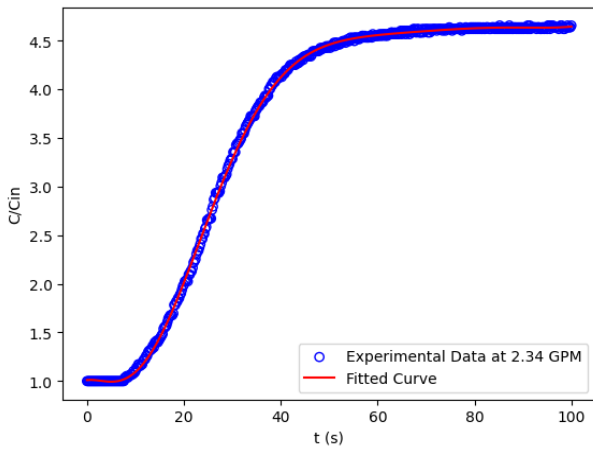
$$spline = UnivariateSpline(x, y, s, k)$$

Where  $x, y$  are the data points,  $s$  and  $k$  the smoothing factor and degree of fit respectively. The spline function was used in curve fitting and the finding the derivative of the data points for the RTD studies.  $E(t)$  can be calculated using the derivative of the spline where  $n$  represents the order of the derivative.

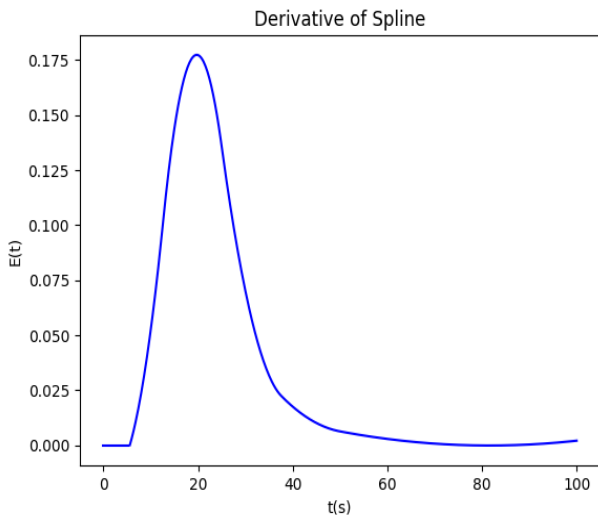
$$E(t) = spline(x, n)$$

## 4. RESULTS

The RTD studies for the pilot plant was evaluated looking at two extreme flowrates of 2.34 GPM, 1.15 GPM and also varying filter configurations of 0.5 L and 1L. The variation of filter configuration was necessary in satisfying the volumetric ratio scale down criterion. The industrial RTD studies were performed based on industrial data with oils as the simulant fluids as against the preliminary pilot plant studies involving conductive solutions. The step change tracer profile of  $F(t)$  for both the industrial and pilot plant system is shown in Figure 5.



**Figure 5:** Cumulative RTD  $F(t)$ : Pilot Plant (top) and Industrial Plant (bottom)



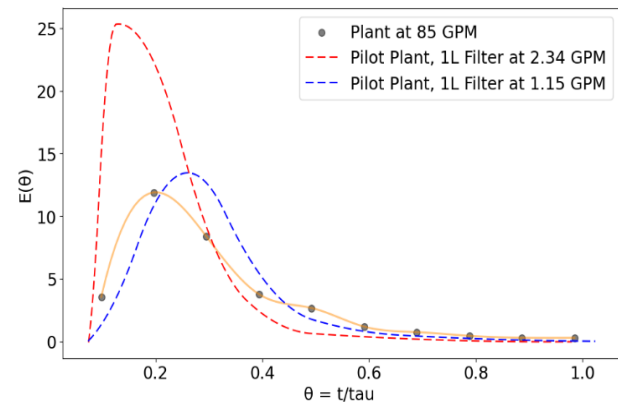
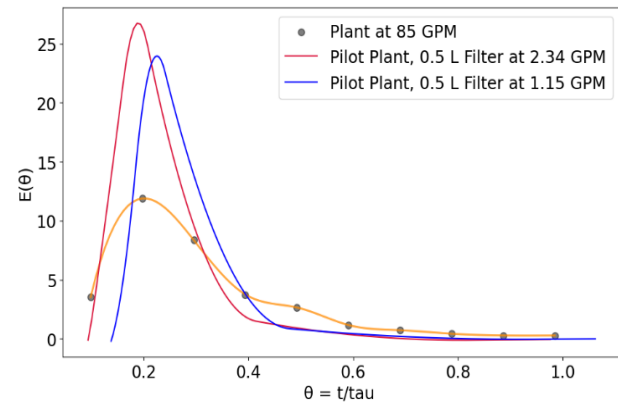
**Figure 6:** Sample RTD function  $E(t)$  of Pilot Plant Studies

The RTD function,  $E(t)$  of the pilot plant studies is illustrated in Figure 6. The peak of the profile indicates the most probable exit time for the conductive solution from the pipeline system.

Hence, providing insights into the point at which the highest concentration of the conductive solution (tracer)

exists the pipeline system. The mean residence time, representing the time average the tracer spends in the system is evaluated by inferring at the balance point of the area under the  $E(t)$  curve.

The Normalized RTD functions,  $E(\theta)$  is plotted against the dimensionless time,  $\theta$  shown in Figure 7. This plot is used as the metric for validating the pilot plant and the industrial plant. Inferring from the profile, we observe a sharp rise which suggests a rapid initial exit of the material, followed by a steadily decline after the peak, indicating a slower exit of the remaining material. Considering the pilot plant studies, at the highest flowrate, the peak is rapid and higher compared to the observed peak at the lower flowrate indicating high tracer concentrations exiting the system. Considering the two filter configurations, the distribution in both systems is symmetrical since the peak is closer to the mean of the distribution.



**Figure 7:** Dimensionless RTD of Industrial Plant Comparison with Pilot plant at 2.34 GPM and 1.15 GPM Considering 0.5 L Filter (top) and 1 L Filter (bottom)

The width of the peak reflects the spread in exit

times, which is related to the standard deviation. A narrow peak suggests that most of the tracer (residuals) exits at around the same time, indicating a small spread. Inferring from the pilot plant and the industrial plant, we observe a sharp and narrow peak, indicating approximately that a significant portion of the tracer exits the system. Hence, it can be said that low dispersion is observed and that both systems have a relatively uniform residence time. A measure of the spread of the distribution in both the pilot plant and industrial plant systems, can be inferred to be relatively small due to the sharpness of the peak, but not minimal to the noticeable tail. This indicates some variance in the residence time, however not a significant one.

The skewness of the curve is observed to be asymmetrical, with a longer tail to the right of the peak. This is an indication that, some of the tracer materials takes longer to exit the system than average, which could be due to areas of slower flow within the system.

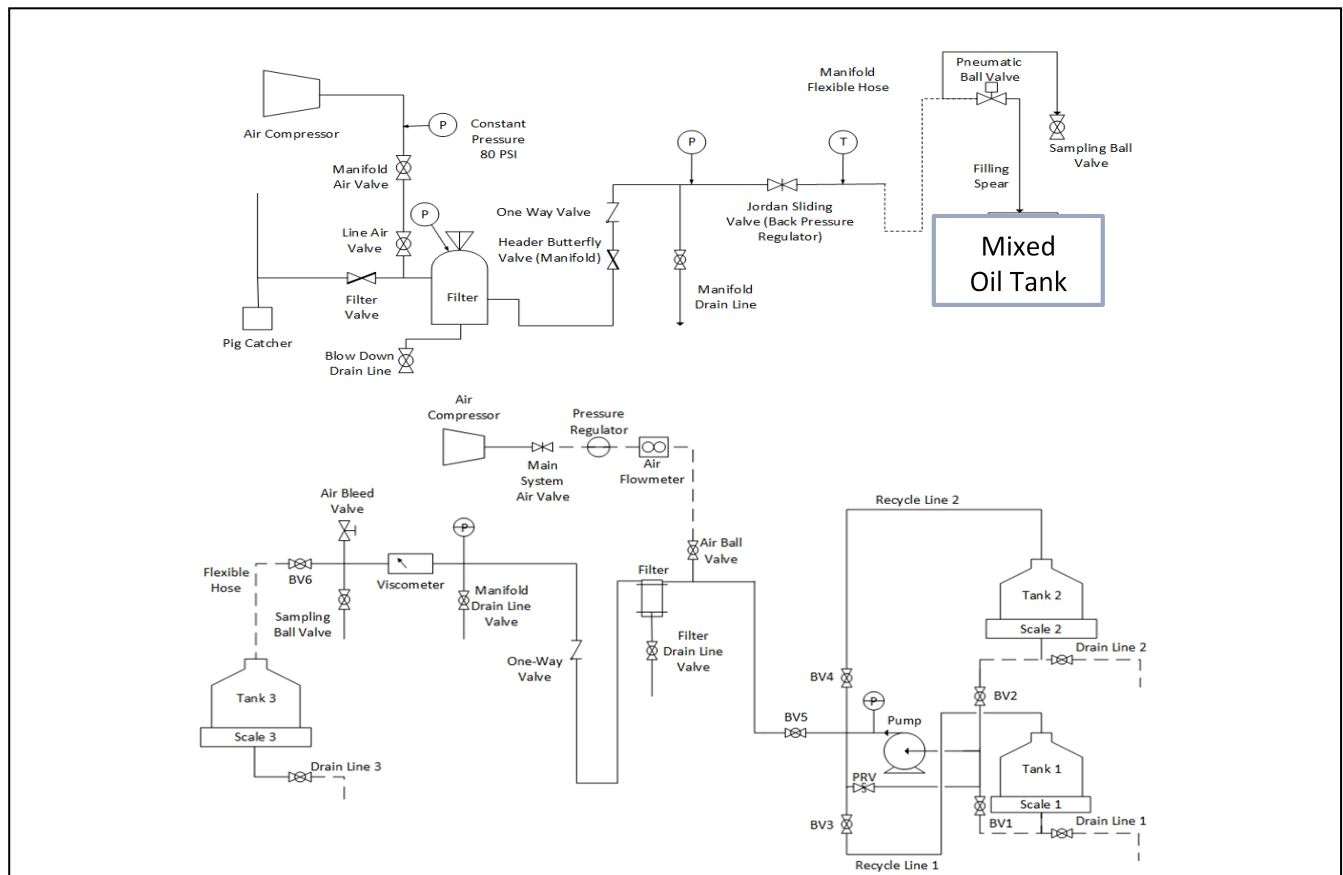
#### 4.1 Model Uncertainty

Model uncertainty in our RTD study primarily arises from the substitution of oil with salt water as a tracer, introducing discrepancies in flow behavior due to differing physical properties. Measurement precision of the

conductivity probe presents additional variability to the tracer concentration data obtained. Furthermore, the pilot plant's scale down from the industrial scale presents additional complexities; even with careful scale down factors to match industrial ratios, scale-dependent phenomena like flow regimes can behave unpredictably, impacting the RTD results. Thus, although our method offers a robust approximation in drawing similarities between the industrial plant and the pilot scale plant, acknowledging these uncertainties are crucial for the robust application and translation of our findings to the industrial scale operations.

#### 5. CONCLUSION

This study encapsulates the comprehensive process of enhancing operations through the design and construction of pilot plant, meticulously scaled down from an industrial scale. The analysis employs residence time distributions to draw parallels between the two systems. Aligning with the scaled dimensions, a 0.5 L filter was judiciously selected to maintain the requisite volumetric ratio of 0.4, mirroring the industrial setup. The RTD studies affirm that the pilot plant not only exhibits analogous dynamism but also effectively emulates the



**Figure 8:** Process Flow Diagram of Scaled-Down Pilot Plant (bottom) Replicating a Generic Lube Oil Industry (top)



industrial plant at a reduced geometry ratio of 1/5. This underscores the pilot plant's fidelity as a robust model for the industrial scale.

## ACKNOWLEDGEMENTS

This research work is funded by the U.S. Environmental Protection Agency (EPA). Grant number: NP-96248222.

## REFERENCES

1. R. M. Baptista, F. B. De Freitas Rachid, and J. H. Carneiro De Araujo, "Estimating Mixing Volumes Between Batches in Multiproduct Pipelines," in *Volume 2: Integrity and Corrosion; Offshore Issues; Pipeline Automation and Measurement; Rotating Equipment*, Calgary, Alberta, Canada: American Society of Mechanical Engineers, Oct. 2000, p. V002T08A008. doi: 10.1115/IPC2000-247.
2. R. Z. Sunagatullin, F. V. Timofeev, A. Kuznetsov, and Y. N. Oludina, "Relevant issues on quality evaluation of petroleum pipeline preparation for oil product transportation," *Oil & Gas Science and Technology – Revue D'IFP Energies Nouvelles*, vol. 74, 2019, doi: 10.2516/ogst/2018098.
3. F. K. Goudarzi, H. R. Maleki, and S. Niroomand, "Mathematical formulation and hybrid meta-heuristic algorithms for multiproduct oil pipeline scheduling problem with tardiness penalties," *Concurrency and Computation: Practice and Experience*, vol. 33, no. 17, 2021, doi: 10.1002/cpe.6299.
4. H. S. Fogler, *Elements of chemical reaction engineering*, Fifth edition. Boston: Prentice Hall, 2016.
5. M. C. Martinetz, A-P. Karttunen, S. Sacher, P. Wahl, J. Ketolainen, J.G. Khinast, and O. Korhonen, "RTD-based material tracking in a fully-continuous dry granulation tableting line," *International Journal of Pharmaceutics*, vol. 547, no. 1–2, pp. 469–479, Aug. 2018, doi: 10.1016/j.ijpharm.2018.06.011.
6. P. Toson, P. Doshi, and D. Jajcevic, "Explicit Residence Time Distribution of a Generalised Cascade of Continuous Stirred Tank Reactors for a Description of Short Recirculation Time (Bypassing)," *Processes*, vol. 7, no. 9, p. 615, Sep. 2019, doi: 10.3390/pr7090615.
7. "scipy.interpolate.CubicSpline — SciPy v1.11.4 Manual." Accessed: Nov. 29, 2023. [Online]. Available: <https://docs.scipy.org/doc/scipy/reference/generated/scipy.interpolate.CubicSpline.html#scipy.interpolate.CubicSpline>

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Advances in Process Synthesis: New Robust Formulations

Smitha Gopinath<sup>a</sup> and Claire S. Adjiman<sup>b,\*</sup>

<sup>a</sup> Department of Chemical and Biological Engineering, The University of Sheffield, Mappin Street, Sheffield, S1 3JD, United Kingdom

<sup>b</sup> Imperial College London, Department of Chemical Engineering, Sargent Centre for Process Systems Engineering and Institute for Molecular Science and Engineering, London, United Kingdom

\* Corresponding Author: c.adjiman@imperial.ac.uk

## ABSTRACT

We present new modifications to superstructure optimization paradigms to i) enable their robust solution and ii) extend their applicability. Superstructure optimization of chemical process flowsheets on the basis of rigorous and detailed models of the various unit operations, such as in the state operator network (SON) paradigm, is prone to non-convergence. A key challenge in this optimization-based approach is that when process units are deselected from a superstructure flowsheet, the constraints that represent the deselected process unit can be numerically singular (e.g., divide by zero, logarithm of zero and rank-deficient Jacobian). In this paper, we build upon the recently-proposed modified state operator network (MSON) that systematically eliminates singularities due to unit deselection and is equally applicable to the context of both simulation-based and equation-oriented optimization. A key drawback of the MSON is that it is only applicable to the design of isobaric flowsheets at a pressure fixed *a priori*. In this paper, as a first step towards the synthesis of general flowsheets with variable pressures, we extend the MSON to the synthesis of a gas-liquid absorption column at variable pressure (i.e., the pressure is a degree of freedom that may be optimized). We illustrate the use of the extended MSON on a carbon-capture process. The extended MSON is robust and enables the design of the column on the basis of detailed thermodynamic models and simulation-based optimization.

**Keywords:** Absorption, Algorithms, Carbon Dioxide Capture, Optimization, Process Synthesis

## INTRODUCTION

Process synthesis is central to the conceptualization of new chemical processes that can meet the manifold constraints of a circular economy. Process synthesis is the activity of identifying an optimal flowsheet which entails choosing a) process units (e.g., unit operations) from a set of alternatives, b) the connectivity of selected process units, and c) the degrees of freedom of selected units such that a design objective is optimized and all process constraints are satisfied.

Superstructure optimization is a mathematical programming approach to process synthesis. While there are several representations of a process superstructure [1], here, we focus our attention on the State Operator Network (SON) representation [2] of the process superstructure. A key feature of the SON is that each allowed process unit is described by its rigorous model which includes MESH equations and equipment sizing and costing correlations. The SON relies on a network of

conceptual mixers and splitters that enable up to full connectivity between the set of selected process units. A mixer and a splitter are located at each inlet and outlet, respectively, of each process unit.

The optimization of the SON is a challenging mixed-integer nonlinear programming problem (MINLP). A particular issue in the optimization of the SON, which is the subject of this paper, is the fate of a process unit that is deselected, that is, excluded from the flowsheet. Naturally, when a process unit is deselected, mass flowrates at each inlet of the unit must be set to zero. However, the models of many process units are well defined only at strictly positive mass flows. At zero-valued inlet flows, several numerical singularities (including undefined behaviour) in the constraint functions that describe the unit and/or in their derivatives can occur. For example, consider an isobaric-isenthalpic flash unit. At zero-valued flows, a two-phase solution to the phase-equilibrium problem does not exist. Further, the Jacobian of the mass-balance constraints of the unit is rank-deficient [3]

and costing and sizing correlations that depend on the flowrates may become numerically singular. As a result, the optimization of the SON may fail to converge.

To overcome this challenge, one may reformulate the SON MINLP using Generalized Disjunctive Programming (GDP) [4]. However, the application of GDP to simulation-based superstructure optimization is limited and computationally expensive [5]. Specifically, the initialization of the master problem in Logic-based Outer Approximation is computationally expensive when applied to simulation-based superstructure optimization [6]. Other reformulations include the Big-M reformulation of all the constraints that describe each unit, or multiplication of the constraints of each unit by the corresponding binary variable [7]. Not only do these modifications not fully eliminate singularities due to zero flows [6], but these also require modifications to the high-dimensional number of constraints that describe each process unit.

In this paper, we build upon the recently developed Modified State Operator Network (MSON) [6]. The MSON modifies mixers by introducing fictitious inlet streams that become active when a unit is deselected and take on strictly positive flowrates as well as intensive property values chosen to guarantee successful evaluation of the model of the corresponding process unit. The MSON modifies splitters to reject any flows at the outlets of the deselected process unit that arise due to these fictitious inlet flows. Further any quantities computed in the deselected unit that result in non-zero contributions to the flowsheet objective, design constraints and so on are also modified to take a zero value when the unit is deselected, thus resulting in an exact reformulation.

The SON and MSON are only applicable to the synthesis of isobaric flowsheets. The MSON is further limited as the pressure needs to be fixed *a priori*. In this paper, we present advances to the MSON towards the synthesis of general flowsheets in which: i) the pressure of any unit operation does not have to be fixed *a priori* and ii) unit operations can operate at different pressures. We address the first stipulation in the context of the synthesis of a counter-current separation column. We note that Smith (1996) [8] briefly outlined the use of pressure-driven flows between process units via a network of compressors and expanders to address the second stipulation.

A particular arena of superstructure optimization that has received much attention in the literature [9-10], including the pioneering work of Sargent and Gaminbandara (1976) [11], is the synthesis of separation columns (that is, the optimal design of number of stages and column degrees of freedom), especially distillation columns. The problem is of renewed importance today as separations are highly energy and capital intensive and improved designs are crucial to the success of emerging areas such as carbon capture and biomanufacturing [12].

The consideration of varying pressures is particularly important in this context, e.g., it can facilitate the design of separation solvents for carbon capture as the optimal choices of solvent, pressure and column configuration are intrinsically linked [13]. While we do not model any pressure drop in the column, the column pressure is a degree of freedom that we optimize, unlike our previous work [6]. The proposed extensions of the MSON are general and can be applied to any process unit in which the pressure is a degree of freedom. Additionally, the formulations presented here are amenable to both simulation-based optimization and equation-oriented optimization.

In the next section we present three different counter-current column synthesis formulations: (i) a counter-current column SON based on the R-graph decomposition of columns previously given by Farkas et al. (2008) [14] where we show how the standard formulation of the SON can suffer from numerical singularities, (ii) an exact MSON formulation of the same column that relies on a simple modification to the mixers and splitters to eliminate singularities due to deselection of stages, (iii) an extended MSON formulation so that the column pressure may be treated as a degree of freedom. We then present details of the implementation of the extended MSON and the application of the extended MSON to the design of carbon capture column. Lastly, we present results and conclusions.

## COLUMN SYNTHESIS FORMULATIONS

Consider a counter-current separation column such as a gas-liquid absorption column, liquid-liquid extraction column or the rectifying section of a distillation column. The column separation is driven by the contacting of two phases, labelled vapour and liquid here for convenience.

The column synthesis problem may be formulated as follows: Given a multi-component vapour feed to be separated, a solvent that is the mass separation agent and a maximum of  $N^U$  theoretical (equilibrium) stages, find the optimal number of stages and values of the column degrees of freedom such that the design objective is minimized and constraints on product purity and recovery are satisfied. The MSON formulation for column synthesis has previously been derived formally in full detail [5]. Here, we present the details of the SON and MSON in the context of column synthesis for completeness.

### SON

We decompose a column of  $N^U$  stages into equivalent conditional subsections [14]. Each subsection  $i$  has  $2^{i-1}$  identical equilibrium stages and an associated binary variable  $z_i$  that takes the value 1 when the subsection is selected, where  $i \in \mathcal{R}$ , where  $\mathcal{R}$  is the set of conditional subsections. For example, a column with a maximum of 15 stages is decomposed into four subsections of 1, 2, 4

and 8 stages, respectively as shown in Figure 1. As an alternative example, a column of 31 stages is decomposed into five subsections of 1, 2, 4, 8 and 15 stages, respectively. The number of subsections  $|\mathcal{R}|$  is equal to the minimum number of bits required to express integer  $N^U$  in binary notation. Each subsection has a vapour outlet, a vapour inlet, a liquid outlet and a liquid inlet as shown in Figure 1. We further include the following permanent units in the column superstructure: a vapour source, a liquid source, a vapour sink and liquid sink as shown in Figure 1. Each sink and source have one inlet and outlet each. The vapour feed to be separated and the fresh solvent enter the column via the vapour and liquid sources, respectively. The product and the spent solvent leave the column via the vapour and liquid sinks, respectively. We assume that all unit operations operate at pressure  $P$ .

### Mixer-splitter network

For each phase (vapour or liquid), we use a network of mixers and splitters to enable flows between the column subsections, and the corresponding source and sink. A mixer is placed at each of the inlets of the column subsections and at the inlet of the sink. A splitter is placed at each of the outlets of the column subsections and at the outlet of the corresponding source. The vapour mixer at inlet  $i \in \mathcal{J}^V$ , where  $\mathcal{J}^V$  is the set of vapour inlets, allows the mixing of vapour streams that leave from splitters at the outlets of the subsections  $i + 1, \dots, \mathcal{R}$  and the vapour source. Similarly, the liquid mixer at inlet  $i \in \mathcal{J}^L$ , where  $\mathcal{J}^L$  is the set of liquid inlets, allows the mixing of liquid streams that leave splitters at the outlets of subsections  $1, \dots, i - 1$  and the liquid source. An example column subsection with its mixers and splitters and their connections in detail is shown in Figure 2. Splitters and mixers are assumed to be isenthalpic and ideal. We consider a multi-component mixture with  $K$  components.  $f_i^{\text{in},V}$ ,  $\mathbf{q}_i^{\text{in},V}$  and  $T_i^{\text{in},V}$  represent the flowrate, composition and temperature of the vapour stream at inlet  $i$ .  $f_o^{\text{out},V}$ ,  $\mathbf{q}_o^{\text{out},V}$  and  $T_o^{\text{out},V}$  represent the flowrate, composition and temperature of the vapour stream at outlet  $o$ .  $f_{o,i}^V$ ,  $\mathbf{q}_{o,i}^V$  and  $T_{o,i}^V$  represent the mass flowrate, composition and temperature, respectively of the vapour stream that flows from splitter  $o$  to mixer  $i$ . All streams are assumed to be at constant pressure  $P$ . The following balances hold for vapour mixers:

$$f_i^{\text{in},V} = \sum_{o \in \mathcal{M}_i^V} f_{o,i}^V \quad (1)$$

$$f_i^{\text{in},V} \mathbf{q}_{i,c}^{\text{in},V} = \sum_{o \in \mathcal{M}_i^V} f_{o,i}^V \mathbf{q}_{o,i,c}^V \quad \forall c \in \{1, \dots, K\} \quad (2)$$

$$f_i^{\text{in},V} h_{en}^V(T_i^{\text{in},V}, P, \mathbf{q}_i^{\text{in},V}) = \sum_{o \in \mathcal{M}_i^V} f_{o,i}^V h_{en}^V(T_{o,i}^V, P, \mathbf{q}_{o,i}^V) \quad (3)$$

where  $\mathcal{M}_i^V$  denotes all the vapour splitters that a vapour mixer  $i$  may be connected to, and for vapour splitters:

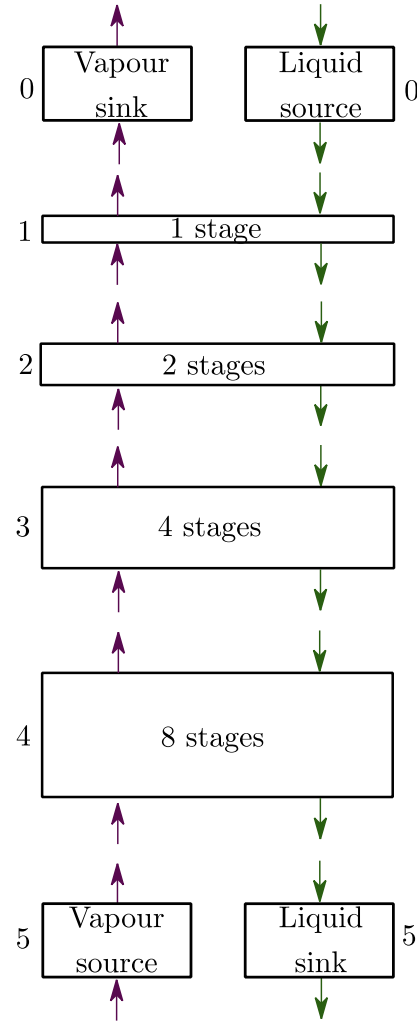
$$f_o^{\text{out},V} = \sum_{i \in \mathcal{S}_o^V} f_{o,i} \quad (4)$$

$$T_o^{\text{out},V} = T_{o,i}^V \quad \forall i \in \mathcal{S}_o^V \quad (5)$$

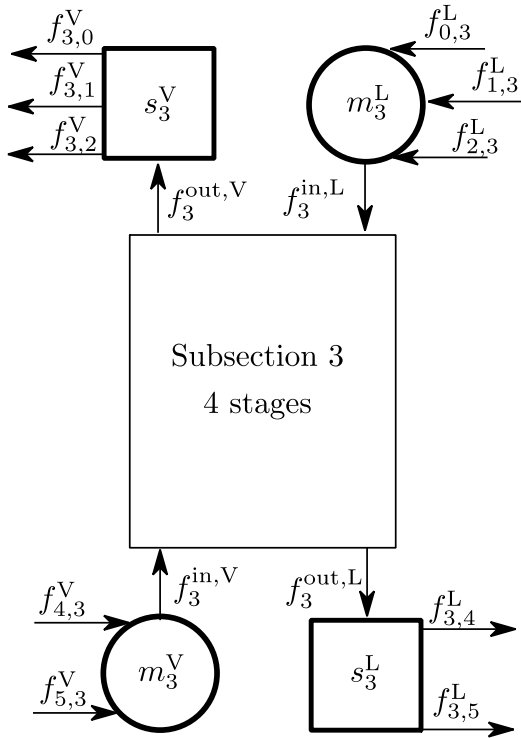
$$\mathbf{q}_o^{\text{out},V} = \mathbf{q}_{o,i}^V \quad \forall i \in \mathcal{S}_o^V \quad (6)$$

where  $\mathcal{S}_o^V$  denotes all the mixers that splitter  $o$  may be connected to. Analogous relationships may be written for the liquid mixers and splitters.

Note that the mixers and splitter alone cannot guarantee flow in the desired direction between two units at different pressures, and thus, limit the SON to the synthesis of isobaric flowsheets.



**Figure 1:** Schematic of the vapour source, liquid source, vapour sink, liquid sink, and the 4 column subsections with 1, 2, 4 and 8 stages in a column superstructure with at most 15 stages. The vapour and liquid sources and sinks have one inlet and outlet each. Each column subsection has one vapour inlet, one liquid inlet, one vapour outlet and one liquid outlet. Process units in the superstructure are labelled by numbers.



**Figure 2:** Detailed schematic of subsection 3 with 4 stages from Figure 1. The vapour mixer  $m_3^V$  allows the mixing of vapour streams that flow from subsection 4 and the vapour source. The liquid mixer  $m_3^L$  allows the mixing of vapour streams that flow from the liquid source, subsection 1 and subsection 2. The vapour splitter  $s_3^V$  allows the vapour that leaves subsection 3 to go to mixers at the inlets of subsections 1 and 2 and the vapour sink. The liquid splitter  $s_3^L$  allows the liquid that leaves subsection 3 to go to mixers at the inlets of subsection 4 and the liquid sink.

### Flow-validity constraints

Flow validity constraints ensure that the flows to a column subsection that is deselected,  $z_i = 0$ , are zero. Further constraints, given in [6], ensure that the flows do not bypass selected subsection.

### Process unit-level constraints

Each process unit (the sources, the sinks and the conditional subsections) is represented in the column superstructure via its rigorous model, which is referred to as process unit-level constraints here. Each source and sink may be represented by simple constraints that equate the state at the inlet to that at the outlet. Each column subsection is represented by a rigorous model for each of the equilibrium stages in the subsection. This includes the MESH equations. As the pressure drop is assumed to be zero, the pressure at the vapour and liquid outlets of each stage (and subsection  $i$  where  $i \in \mathcal{R}$ ) is set equal to the pressure at the inlet. Thus,

$$P_i^{\text{out},V} = P_i^{\text{out},L} \quad (7)$$

$$P_i^{\text{out},V} = P_i^{\text{in},V} \quad (8)$$

$$P_i^{\text{in},V} = P \quad (9)$$

For each subsection, we compute variables  $\widehat{\rho}_i^L$ ,  $\widehat{\rho}_i^V$  and  $\widehat{v}_i^V$ :

$$\widehat{\rho}_i^L = \sum_{j \in \mathcal{Z}^{i-1}} \rho_j^L \quad (10)$$

$$\widehat{\rho}_i^V = \sum_{j \in \mathcal{Z}^{i-1}} \rho_j^V \quad (11)$$

$$\widehat{v}_i^V = \sum_{j \in \mathcal{Z}^{i-1}} v_j^V \quad (12)$$

Where  $\rho_j^L$ ,  $\rho_j^V$  and  $v_j^V$  are the mass density of the liquid stream that exits stage  $j$ , the mass density of the vapour stream that exits stage  $j$  and volumetric flowrate of the vapour stream that exits stage  $j$ , respectively.

### Column-level constraints

We use column-level equations (also known as flowsheet-level equations in [6]) to compute the total dimensions and cost of the column. Some of these flowsheet-level constraints may depend on a few “output variables” whose values are obtained by solving the process unit-level constraints.  $D$ , the diameter of the column and  $H$ , its height, are computed using [15]:

$$N = \sum_{i \in \mathcal{R}} 2^{i-1} z_i \quad (13)$$

$$\overline{\rho}^L = \frac{\sum_{i \in \mathcal{R}} \widehat{\rho}_i^L}{N} \quad (14)$$

$$\overline{\rho}^V = \frac{\sum_{i \in \mathcal{R}} \widehat{\rho}_i^V}{N} \quad (15)$$

$$\overline{v}^V = \frac{\sum_{i \in \mathcal{R}} \widehat{v}_i^V}{N} \quad (16)$$

$$u^{\text{flood}} = (-0.1711t^2 + 0.271t - 0.047) \sqrt{\frac{\overline{\rho}^L - \overline{\rho}^V}{\overline{\rho}^V}} \quad (17)$$

$$D = \sqrt{\frac{4\overline{v}^V}{\pi u^{\text{flood}}}} \quad (18)$$

$$H = 1.15 \frac{lt}{E} N \quad (19)$$

where  $lt$  and  $E$  are the tray spacing and stage efficiency, respectively. We also introduce flowsheet-level constraints to compute the total capital investment ( $TCI$ ), the annual operating expenses ( $OPEX$ ) and the total annualized cost ( $TAC$ ) using costing correlations [6,16]. We also impose constraints on the minimum purity and flowrate of the vapour stream that exits the flowsheet.

### Numerical singularities

When a column subsection is deselected, all flows into the unit are driven to zero by the flow-validity constraints. Due to this, within each column subsection, a solution to the phase equilibrium equations at each stage does not exist. The mass-balance equations are also



rank-deficient. Sizing, costing or even mass transfer correlation that depends on the flowrates associated with a deselected stage can become singular (that is, the function or derivative may be numerically undefined).

## MSON

To overcome numerical singularities for any deselected process unit, we have developed the MSON formulation [6], which relies on the modification of mixers and splitters associated with the conditional subsections. The mixers and splitters associated with the sources and sinks remain unchanged and the pressure of all column subsections is a constant and fixed *a priori*.

### Modified mixer

We introduce a fictitious stream into each modified vapour mixer  $i$  (associated with column subsection  $i \in \mathcal{R}$ ) with mass flowrate  $f_i^{M,V}$ , composition  $\mathbf{q}_i^{A,V}$  and temperature  $T_i^{A,V}$ , as shown in Figure 3. Due to Equations (20) – (22), the flowrate, composition and temperature at the inlet of subsection  $i$  are  $f_i^{M,V}$ ,  $\mathbf{q}_i^{A,V}$  and  $T_i^{A,V}$ , respectively, when subsection  $i$  is deselected and all other flows into the mixer are zero. Further, due to Equation (23),  $f_i^{M,V}$  takes the constant value  $f_i^{A,V}$  when the conditional subsection is deselected and takes the value zero otherwise. Due to this, when the subsection is selected, the modified mixer equations are fully equivalent to a standard mixer. Analogous relationships are written for the modified liquid mixers.

$$f_i^{\text{in},V} = \sum_{o \in \mathcal{M}_i^V} f_{o,i}^V + f_i^{M,V} \quad (20)$$

$$f_i^{\text{in},V} \mathbf{q}_{i,c}^{\text{in},V} = \sum_{o \in \mathcal{M}_i^V} f_{o,i}^V \mathbf{q}_{o,i,c}^V + f_i^{M,V} \mathbf{q}_{i,c}^A \forall c \in \{1, \dots, K\} \quad (21)$$

$$f_i^{\text{in},V} h_{en}^V(T_i^{\text{in},V}, P, \mathbf{q}_i^{\text{in},V}) = \sum_{o \in \mathcal{M}_i^V} f_{o,i}^V h_{en}^V(T_{o,i}^V, P, \mathbf{q}_{o,i}^V) + f_i^{M,V} h_{en}^V(T_i^{A,V}, P, \mathbf{q}_i^{A,V}) \quad (22)$$

$$f_i^{M,V} = f_i^{A,V} (1 - z_i) \quad (23)$$

$f_i^{A,V}$ ,  $\mathbf{q}_i^{A,V}$  and  $T_i^{A,V}$  and the corresponding constants for the modified liquid mixer are chosen such that the two fictitious streams are in vapour-liquid equilibrium at pressure  $P$ . When a column subsection is not selected, thanks to the fictitious streams that enter via the modified vapour and liquid mixers, a two-phase solution exists in each stage. Thus, no singularities are encountered in the MESH equations and sizing correlations that describe the subsection.

### Modified splitter

We modify the splitters associated with the conditional column subsections as shown in Figure 4. With the modified splitters we can correct for the fictitious non-

zero flows at the outlets of deselected process units caused due to fictitious mixer streams. We describe each modified vapour splitter  $o$ , (associated with column subsection  $o \in \mathcal{R}$ ) using Equations (5) and (6) and,

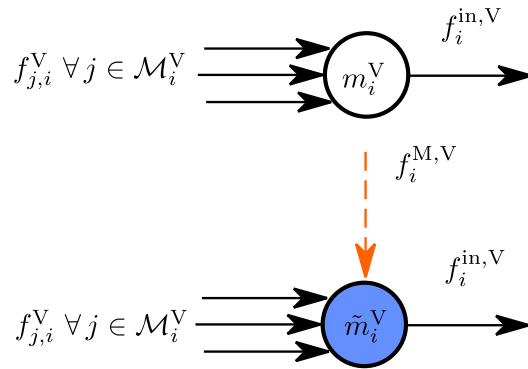
$$f_o^{\text{out},V} = \sum_{i \in \mathcal{S}_o^V} f_{o,i} + f_o^{S,V} \quad (24)$$

$$0 \leq f_o^{\text{out},V} - f_o^{S,V} \leq f^U z_o \quad (25)$$

$$0 \leq f_o^{S,V} \leq f^U (1 - z_o) \quad (26)$$

where  $f^U$  is an upper bound on the flowrates.

Equations (24)-(26), ensure that when the subsection is deselected, any vapour that leaves the subsection (due to the fictitious flows in the modified mixers) leaves the splitter  $o$  via a fictitious stream with flowrate  $f_o^{S,V}$ , composition  $\mathbf{q}_o^{\text{out},V}$  and temperature  $T_o^{\text{out},V}$  and is not propagated to the rest of the flowsheet. On the other hand, when the subsection is selected, the fictitious stream is constrained to have a zero flowrate, and hence has no effect on the rest of the flowsheet. As before, the composition and temperature of the streams that leave the splitters are set equal to that at the inlet of the splitter. A vapour splitter and a modified vapour splitter are shown in Figure 4. The modified liquid splitter is analogous.



**Figure 3:** A vapour mixer (denoted by an open circle) at inlet  $i$  and a modified vapour mixer (denoted by a shaded circle) at inlet  $i$ .

### Modified output variables

As a result of the fictitious flows into a deselected unit, several of the variables associated with the unit may take a spurious non-zero value. However, only a small subset of these variables, the output variables, is used in column-level computations. We introduce a “corrected output variable”  $x_j^S$  for any output variable  $x_j$  and enforce the following Big-M constraints:

$$-M(1 - z_i) \leq x_j - x_j^S \leq M(1 - z_i) \quad (27)$$

$$-Mz_i \leq x_j^S \leq Mz_i. \quad (28)$$

We further modify the column-level constraints that

depend on an output variable  $x_j$  to depend on the corrected output variable  $x_j^S$ . In the case of column synthesis, we reformulate constraints (10)-(12) in terms of these corrected output variables to obtain:

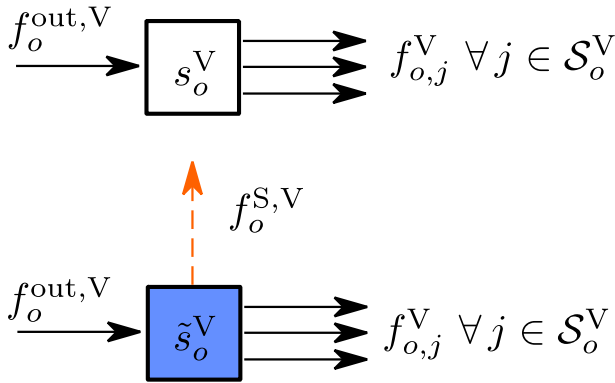
$$\overline{\rho^L} = \frac{\sum_{i \in \mathcal{R}} \rho_i^L S}{N} \quad (29)$$

$$\overline{\rho^V} = \frac{\sum_{i \in \mathcal{R}} \rho_i^V S}{N} \quad (30)$$

$$\overline{v^V} = \frac{\sum_{i \in \mathcal{R}} v_i^V S}{N} \quad (31)$$

### Process unit-level constraints

The process unit-level constraints and all column-level constraints that do not depend on any of the output variables are unchanged. Numerical singularities by the use of the MSON are entirely averted as shown in [6].



**Figure 4:** A vapour splitter (denoted by an open square) at outlet  $o$  and a modified vapour splitter (denoted by a shaded square) at outlet  $o$ . A fictitious stream (denoted by the dashed arrow) with flowrate  $f_o^{S,V}$  leaves the modified splitter.

### EXTENDED MSON (E-MSON)

In the derivation of the SON and MSON we assume that the pressure at the inlet of any unit (that is, the outlet of any mixer) is fixed at  $P$ . The operating pressure of any column subsection is also fixed at  $P$ , thanks to Equations (7)-(9). Further, we introduce fictitious streams in the modified mixers with states such so that when the unit is deselected, a two-phase solution is guaranteed for a column subsection at pressure  $P$ . However, if the deselected subsection is at any other pressure, the states of the fictitious stream may not necessarily result in a two-phase solution. Indeed, for any general process unit, the state of the fictitious stream in the mixer may not lead to the successful solution of the unit constraints when the operating pressure varies.

To overcome this limitation we first add the following constraint to each vapour splitter as well as each

modified vapour splitter  $o$  in the superstructure:

$$P_{o,i}^V = P_o^{\text{out},V} \quad \forall i \in \mathcal{S}_o^V \quad (32)$$

where  $P_{o,i}^V$  is the pressure of a vapour stream that flows from splitter  $o$  to mixer  $i$ . Analogous constraints are added to the liquid splitters.

We also modify the mixers in the flowsheet. The ‘‘mixing’’ of two streams that arise from two process units at unequal pressures may physically result in unintended flows, e.g., mass flows from the high-pressure unit into the low-pressure unit. Thus, in the extended MSON, only the mixing of streams that are at equal pressures is allowed. We note that this mixing rule is trivially satisfied for the column synthesis problem studied here (as all streams are at  $P$ ). The following pressure-mixing equation is introduced into the model of each vapour mixer  $i$  to describe the pressure at each vapour inlet:

$$f_i^{\text{in},V} P_i^{\text{in},V} = \sum_{o \in \mathcal{M}_i^V} f_{o,i}^V P_{o,i}^V \quad (33)$$

Due to the pressure-mixing constraint, the pressure at the inlet  $i$  is equal to that of stream(s) with strictly positive mass flowrates into unit  $i$ . The enthalpy balance for each mixer is modified to incorporate the pressure of each stream, yielding:

$$f_i^{\text{in},V} h_{en}^V(T_i^{\text{in},V}, P_i^{\text{in},V}, \mathbf{q}_i^{\text{in},V}) = \sum_{o \in \mathcal{M}_i^V} f_{o,i}^V h_{en}^V(T_{o,i}^V, P_{o,i}^V, \mathbf{q}_{o,i}^V) \quad (34)$$

A similar pressure-mixing equation is introduced into the modified vapour and liquid mixers. The pressure of the fictitious vapour and liquid streams in the modified mixer  $i$  is denoted by  $P_i^A$ .

$$f_i^{\text{in},V} P_i^{\text{in},V} = \sum_{o \in \mathcal{M}_i^V} f_{o,i}^V P_{o,i}^V + f_i^{\text{M},V} P_i^A \quad (35)$$

Due to Equation (35), the pressure at the inlet  $i$  is equal to that of stream(s) with strictly positive mass flows into unit  $i$ . When the unit  $i$  is deselected, the pressure at the inlet takes the value  $P_i^A$  that leads to successful evaluation of the process unit-level constraints that describe  $i$ . The enthalpy balance in modified mixers is also adapted, in an analogous matter to Equation (34).

The E-MSON only requires the modification to mixers and splitters. All process unit-level constraints and column-level constraints are the same as the MSON. The E-MSON can be used to model a process unit with variable pressure and ensures successful solution when the unit is deselected, irrespective of its operating pressure. Further, the extended mixers and splitters are also necessary to allow the E-MSON to be applied to the synthesis of flowsheets in which the unit operations are at different pressures. However, to fully address the latter case for a general flowsheet, additional logical constraints as well as compressors and expanders are needed and we leave this for future work.

## IMPLEMENTATION

The E-MSON problem formulation is a mixed-integer nonlinear programming problem (MINLP) that is solved using the outer approximation equality relaxation augmented penalty (OA-ER-AP) algorithm [17] that we have implemented in C++ [6]. All constraints of the primal problem are implemented and solved in gPROMS Model-Builder 7.0.7 [18]. Each stage in the column is modelled as an equilibrium stage and all thermodynamic properties are computed using the SAFT- $\gamma$  Mie equation of state [19,20]. The master problem is solved using Gurobi 10.0.2 [21] via its C++ application programming interface (API).

## CASE STUDY DESCRIPTION

We consider the synthesis of a gas-liquid absorption column with at most 15 theoretical stages. The column must recover methane from a carbon dioxide and methane stream, by physical absorption at high pressure. Given a feed of carbon dioxide and methane at flowrate  $1 \text{ kmol s}^{-1}$ , 298 K and pressure  $P$  with 20%  $\text{CO}_2$  and tetra(oxyethylene)dimethylether ( $\text{CH}_3\text{O}(\text{CH}_2\text{O})_4\text{CH}_3$ ) as a solvent at 298 K and pressure  $P$ , find the optimal theoretical number of stages  $N$ , flowrate of fresh solvent  $F$  as well as column pressure  $P$  such that the treated gas has a flowrate of at least  $0.66 \text{ kmol s}^{-1}$  and is at least 97% methane and the total annualized cost  $TAC$  is minimized. The  $TAC$  depends on the total capital investment  $TCI$  and the annual operating expenses  $OPEX$ . We assume  $E$  is 0.8,  $lt$  is 0.6m, the cost of capital is 15%, the column lifetime is 10 years and that  $2 \text{ MPa} \leq P \leq 7.5 \text{ MPa}$ . Details of the model can be found in [6, 16].

## RESULTS

The superstructure for the column to be designed has 4 conditional subsections with 1, 2, 4 and 8 equilibrium stages, respectively. The results of the case study are summarised in Table 1. The optimal column has 4 theoretical stages, obtained by selecting only subsection 3. The OA-ER-AP algorithm converges to a solution in 8 major iterations, 5 of which are found to be feasible. With the E-MSON, the primal problems are solved robustly, no singularities are detected, and a solution is obtained for the equations of the deselected subsections, despite the variation of the column pressure in the course of optimization. The primal problem in the E-MSON has 910 equality constraints (excluding the equations corresponding to SAFT- $\gamma$  Mie), 87 inequality constraints and 49 degrees of freedom to be optimized (including fictitious flowrates in the modified mixers and splitters). The runtime (wall clock time) of the primal problem solution is 65 s on average with the standard deviation of 22 s. Across the five feasible iterations, the

primal problems take 72.2 s to converge to a solution, with the standard deviation of 14.5 s.

**Table 1:** Results of the E-MSON for the synthesis of a counter-current carbon-capture column. The first row shows the results for the variable-pressure case with the use of the E-MSON. The second row shows the results of the same case study in the fixed-pressure case with the use of the MSON where all streams and the column are at an *a priori* fixed pressure of 7.5 MPa.  $F$  is in  $\text{kmol/s}$ ,  $P$  is in MPa and  $TAC$ ,  $TCI$  and  $OPEX$  are in Million USD.

Case	$P$	$F$	$N$	$TAC$	$TCI$	$OPEX$
E-MSON	2.6	1.9	4	0.37	0.58	0.26
MSON	7.5	0.6	6	0.59	1.37	0.32

We compare the results of the E-MSON with that of the MSON in which the column pressure and that of the feed and pure solvent are arbitrarily fixed at 7.5 MPa. The *a priori* fixed-pressure column has a minimum  $TAC$  that is 59% higher than the column in which pressure has been optimized, a  $TCI$  that is 136% higher and 2 more equilibrium stages. In the fixed pressure case, a column with 4 stages was found to be infeasible. The 6-stage column at 7.5 MPa has a 34% higher  $TAC$  than a 6-stage column at the optimal pressure of 2.6 MPa. The pressure heavily impacts the relative solubilities of  $\text{CO}_2$  and  $\text{CH}_4$ , product recovery, product quality and the capital cost of the absorber. The comparison across the two cases highlights the importance of making the pressure an additional degree of freedom in separation column synthesis both for the study of specific separation solvents as well as for solvent design. As we only use local optimization algorithms, the inferior performance of the fixed-pressure column could also be due to convergence of the MSON to low-quality local minima. However, to assuage discrepancies across formulations, we used the same initial guesses, solvers and solver parameters in both cases.

In our future work we shall extend the design envelope to also consider i) the compression/expansion of the feed, ii) the solvent regeneration and recycle and iii) the cost of make-up solvent.

## CONCLUSION

In this paper, we extended the state operator network paradigm of superstructure optimization that was previously implicitly limited to the synthesis of isobaric flowsheets. We built upon the recently-developed MSON that guarantees convergence of deselected units, thus, making the MSON robust and amenable to high-fidelity simulation-based optimization. We developed E-MSON, the extended MSON which retains the robustness of the MSON while making it applicable to the synthesis of columns as well as isobaric flowsheets in which the operating pressure is a variable.

By introducing pressure as an additional degree of freedom in absorption column synthesis, we found in our case study that overall costs decreased by 37%. The extended-MSON is a first step towards fully general superstructure optimization within the SON paradigm in which all unit operations are at variable pressures. The results of the paper indicate that modelling the effects of pressure rigorously can lead to superior designs. Furthermore this work expands the applicability of the state operator network to superstructure optimization problems with rigorous process models.

## REFERENCES

- Mencarelli L, Chen Q, Pagot A, Grossmann IE. A review on superstructure optimization approaches in process system engineering. *Comput. Chem. Eng.* 136:106808 (2020)
- Smith E, Pantelides C. Design of reaction/separation networks using detailed models. *Comput. Chem. Eng.* 19:83 – 88 (1995)
- Dowling AW, Biegler LT. A framework for efficient large scale equation-oriented flowsheet optimization. *Comput. Chem. Eng.* 72:3 – 20 (2015)
- Turkay M, Grossmann IE. Logic-based MINLP algorithms for the optimal synthesis of process networks. *Comput. Chem. Eng.* 20(8):959 – 978 (1996)
- Caballero J. Logic hybrid simulation-optimization algorithm for distillation design. *Comput. Chem. Eng.* 72:284 – 299 (2015).
- Gopinath S, Adjiman CS. Increasing the robustness of superstructure optimization with rigorous models via an exact reformulation. <https://zenodo.org/records/10946545>
- Burre J, Bongartz D, Mitsos A. Comparison of MINLP formulations for global superstructure optimization. *Optim. Eng.* 24:801–830 (2023)
- Smith E. On the Optimal Design of Continuous Processes. Imperial College London (1996).
- Yeomans H, Grossmann IE. Disjunctive programming models for the optimal design of distillation columns and separation sequences. *Ind. Eng. Chem. Res.* 39(6):1637–1648 (2000)
- Barttfeld M, Aguirre PA, Grossmann IE. Alternative representations and formulations for the economic optimization of multicomponent distillation columns. *Comput. Chem. Eng.* 27:363–383 (2003)
- Sargent R, Gaminibandara K. Optimum design of plate distillation columns. *Comput. Chem. Eng.* 19:83 – 88 (1976)
- Favre E. Specialty grand challenges in separation processes. *Front. Chem. Eng.* 2 (2020)
- Gopinath S, Jackson G, Galindo A, Adjiman, CS. Outer approximation algorithm with physical domain reduction for computer-aided molecular and separation process design. *AIChE J.* 62(9):3484–3504 (2016)
- Farkas T, Czuczai B, Rev E, Lelkes Z. New MINLP model and modified outer approximation algorithm for distillation column synthesis. *Ind. Eng. Chem. Res.* 47(9):3088–3103 (2008)
- Sinnott R, Towler G. Chapter 11 – separation columns (distillation, absorption and extraction). In: *Chemical Engineering Design (Sixth Edition)*. Ed: Sinnott R, Towler G. Butterworth-Heinemann (2020).
- Pereira FE, Keskes E, Galindo A, Jackson G, Adjiman CS. Integrated solvent and process design using a SAFT-VR thermodynamic description: High-pressure separation of carbon dioxide and methane. *Comput. Chem. Eng.* 35:474–491 (2011)
- Kocis GR, Grossmann IE. Relaxation strategy for the structural optimization of process flow sheets. *Ind. Eng. Chem. Res.* 26(9):1869–1880 (1987)
- Siemens. <https://www.siemens.com/>
- Papaioannou V, Lafitte T, Avendano C, Adjiman CS, Jackson G, Muller EA, Galindo A. Group contribution methodology based on the statistical associating fluid theory for heteronuclear molecules formed from Mie segments. *J. Chem. Phys.* 140(5):054107 (2014)
- Burger J, Papaioannou V, Gopinath S, Jackson G, Galindo A, Adjiman CS. A hierarchical method to integrated solvent and process design of physical CO<sub>2</sub> absorption using the SAFT- $\gamma$  Mie approach. *AIChE J.* 61:3249–3269 (2015)
- Gurobi Optimization, Inc. <https://www.gurobi.com>

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Integration of Design and Operation with Discretization Error Control

Christian Hoffmann<sup>a\*</sup>, Erik Esche<sup>a</sup>, and Jens-Uwe Repke<sup>a</sup>

<sup>a</sup> Technische Universität Berlin, Process Dynamics and Operations Group, Berlin, Germany

\* Corresponding Author: [c.hoffmann@tu-berlin.de](mailto:c.hoffmann@tu-berlin.de).

---

## ABSTRACT

Optimization-based process design is a central task of process systems engineering. However, solely relying on steady-state models may potentially lead to dynamic constraint violations, hinder robust performance, or simply reduce the controllability of a process. This has led to the consideration of process dynamics in the design phase, which is commonly termed integration of design and operation / control. Recently, we proposed a framework to carry out this integrative task by formulating a large-scale nonlinear programming problem that is solved simultaneously. To this end, the dynamic process model was discretized, and dynamic variability and parametric uncertainty were included. However, the proposed framework only operates on constant lengths of the finite elements. The discretization error was not assessed. Within this contribution, a method for quantifying this discretization error and adapting the number of finite elements accordingly is incorporated into the recently proposed framework and applied on the case study of a continuous tank reactor. The obtained results with and without discretization error control are compared and, based thereon, a more suitable way to apply the control variables on the process is proposed.

**Keywords:** Integration of design and operation, Process design, Grid refinement, Nonlinear programming

## INTRODUCTION

Process design is a central task of process systems engineering. This is conventionally achieved by using flowsheet simulators or other software to determine a suitable flowsheet. Based on this flowsheet, the process is designed using methods such as sensitivity analysis. A possible alternative is optimization-based process design in which the whole flowsheet is optimized for a given objective function. Both approaches typically rely on assumptions of stationarity. However, simply relying on steady-state models may potentially lead to dynamic constraint violations, hinder robust performance, or simply reduce the controllability of a process [1]. This has led to the consideration of process dynamics in the design, which is commonly termed integration of design and operation / control (IDO). Among the approaches to solving IDO problems is the dynamic optimization approach, which assesses the problem using rigorous dynamic models, and the robust approach, which ensures that no constraint violations occur at any point in time. For more information on the possible approaches to solving IDO

problems, the reader is referred to the literature [1,2]. In recent years, we have proposed a framework for the solution of IDO problems, which can be seen as a hybrid of the dynamic optimization approach and the robust approach. The concept is based on a full discretization of the dynamic problem via orthogonal collocation, which yields a large-scale nonlinear programming problem [3]. The framework was then extended to account for parametric uncertainty [4] and to a larger case study [5]. Note that we limit ourselves to deterministic solutions in this contribution. The extension to optimization under uncertainty is, however, straightforward.

So far, the proposed framework has considered a constant number and length for the finite elements in the discretized dynamic model, which may potentially introduce notable discretization errors. Therefore, this contribution addresses this drawback by adding a method for quantifying and controlling the integration error. The extended framework is applied on a case study containing a continuously stirred tank reactor (CSTR) for which previous results are available so that the former and the new approach can be compared fairly.



## STATUS QUO OF THE IDO FRAMEWORK

The current formulation of the IDO problem reads as follows [6]:

$$\min_{d, u_{cp,fe,sp}} \mathbb{E} [f(d, u_{cp,fe,sp}, x_{cp,fe,sp}, \xi_{sp}, v_{cp,fe})] \quad (1)$$

s.t.

$$0 = g(d, u_{cp,fe,sp}, x_{cp,fe,sp}, \xi_{sp}, v_{cp,fe}) \quad (2)$$

$$0 \leq h(d, u_{cp,fe,sp}, x_{cp,fe,sp}, \xi_{sp}, v_{cp,fe}) \quad (3)$$

$$0 = h_0(d, u_{0,sp}, x_{0,sp}, \xi_{sp}, v_0, t_0 = 0) \quad (4)$$

$$0 = h_t(d, u_{t,sp}, x_{t,sp}, \xi_{sp}, v_t, t_t) \quad (5)$$

$$d \in \mathcal{D} \quad (6)$$

$$u_{cp,fe,sp} \in \mathcal{U} \quad (7)$$

$$x_{cp,fe,sp} \in \mathcal{X} \quad (8)$$

$$v_{cp,fe} \in \text{APRBS} \quad (9)$$

$$v_{cp,fe} \sim \mathcal{N}(\mu, \Sigma) \quad (10)$$

Therein, the expected value of the objective function is minimized by varying the design  $d$  and the control variables  $u$  at all collocation roots  $cp$ , every finite element  $fe$ , and for all sigma points of the uncertainty space  $sp$ . The constraints include the dynamic process model  $g$ , path constraints, initial and terminal constraints  $h_0$  and  $h_t$ . The variability of dynamic inputs is represented by an amplitude-modulated pseudo-random binary sequence (APRBS) [7] with normally distributed amplitude whereas the uncertain parameters  $\xi$  and the resulting probability density function of the uncertain inequality constraints, i.e., chance constraints, are represented by a point estimation method, e.g., a cubature rule. All design, state, and control variables are defined on their respective domain. The problem is solved simultaneously as large-scale nonlinear problem in Python, which is interfaced with the model formulation in AMPL [8] and the optimization algorithm IPOPT [9] via AmplPy.

## METHODOLOGY

We now describe the theoretical basis of the error calculation. Afterwards, the solution of the optimization problem is sketched, including the initialization of the new variables that appear due to the increasing number of finite elements. Afterwards, the case study in this contribution is outlined.

### Error calculation

There are several options to assess the discretization error. In many situations, one can compare the error between a highly resolved solution and the solution obtained with the current mesh grid to find an appropriate

number of finite elements [10,11]. While this might be suitable for the solution of a standard boundary value problem, the additional challenge of this work are the relatively long time horizons and the additional degrees of freedom in terms of the design variables of the process. As a results, the highly resolved solution might require significant computational effort and careful initialization. To avoid that, we chose an error estimation that does not compare the difference between two solutions, but the residual between the left-hand and the right-hand side of the differential equations. This estimation of the discretization error is based on the work by Chen et al. [12,13]. In their first publication, they formulated the search for an appropriate number and length of the finite elements as a bilevel optimization problem. As they pointed out, this entails significant computational demand [12]. Hence, they proposed in their subsequent publication to estimate the discretization error by evaluating the residual of the orthogonal collocation  $T_{fe}$  [13]:

$$T_{i,fe} = \frac{dx_i}{dt}(t_{fe,nc}) - \Delta t_{fe} f(x, u, d, t_{fe,nc}). \quad (11)$$

Therein, the residual for the  $i$ -th state variable is evaluated at one (or several) non-collocation points (subscript  $nc$ ), i.e., points different from the used Radau roots. In their approach, Chen et al. also included the residuals of the algebraic equations. Here, we limit ourselves to the differential equations of the model although the extension to algebraic equations is straightforward. Based on the residual, the discretization error is estimated as:

$$err_{i,fe} = C \|T_{i,fe}\|. \quad (12)$$

The constant  $C$  is given by the following integral:

$$C = \frac{1}{A} \int_0^{\tau_{fe,nc}} \prod_{j=1}^K (s - \tau_j) ds, \quad A = \prod_{j=1}^K (\tau_{fe,nc} - \tau_j) \quad (13)$$

This constant only depends on the choice of the non-collocation point(s) and the type of collocation points (in this case Radau) and can thus be calculated once at the beginning. The discretization error must be smaller than a user-defined tolerance:

$$err_{i,fe} \leq \text{tol}_i \quad (14)$$

A similar scheme was recently proposed in [14].

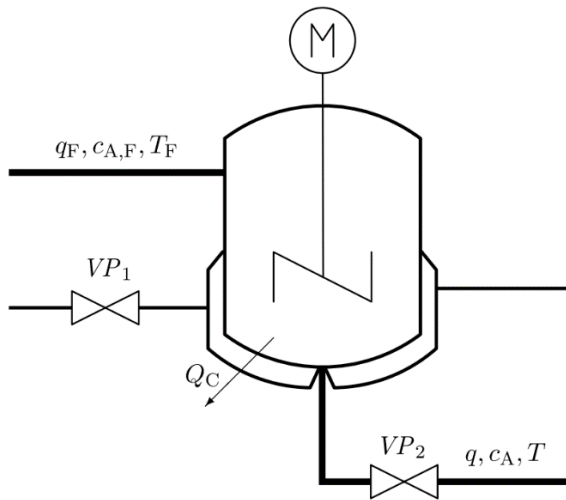
### Solution

The algorithmic approach of this contribution is as follows: First, the optimization problem above is initialized, e.g., via a steady-state optimization. Afterwards, the dynamic system is added, and dynamic variability of the inputs is introduced. Then, the optimization problem is solved by minimizing the objective using a nonlinear optimization algorithm, in this case IPOPT [9]. Based on the obtained results, the discretization error is calculated as described in Equations (12) and (13). If the error within one finite element is larger than the tolerance, the

respective finite element is split into two. These two finite elements are now initialized based on the current solution. This could be done via simple linear interpolation, but as we use collocation, the new initial guesses are calculated by evaluating the Lagrange interpolation polynomial, i.e., the sum of the Lagrangian basis polynomials of the collocation, at the new collocation points. This is repeated for all finite elements. Both approaches to initialize the finite elements have their advantages and disadvantages. For the simple case study below, both would probably work equally well. Finally, the optimization problem is solved again and the error calculation is repeated. This step continues until all errors of the finite elements are smaller than the set tolerance.

## Case study

The case study is a CSTR as shown in Figure 1. Component A enters the reactor via the volume flow  $q_F$  and reacts to B. The volume flow  $q$  then leaves the system with concentration  $c_A$  and temperature  $T$ . The feed flow and the feed temperature are assumed as variable, dynamic inputs. The cooling duty  $Q_C$  can be manipulated with cooling water whose stream is change via the valve position  $VP_1$ . The outlet flow can be manipulated using valve position  $VP_2$ .



**Figure 1.** Flowsheet of the CSTR in this case study [6].

For the process model, the following three assumptions are made [6]:

1. Density and heat capacity of the flows are constant and thus independent of temperature, pressure, or composition.
2. The liquid phase in the reactor is ideally mixed.
3. The pressure remains constant during operation.

This results in the following mass, component, and energy balances:

$$\frac{dV_{\text{CSTR}}^L}{dt} = q_F - q \quad (15)$$

$$\frac{dc_A}{dt} = \frac{q_F}{V_{\text{CSTR}}^L} (c_{A,F} - c_A) - k_0 \exp\left(-\frac{E}{RT}\right) c_A \quad (16)$$

$$\frac{dT}{dt} = \frac{q_F}{V_{\text{CSTR}}^L} (T_F - T) + k_0 \exp\left(-\frac{E}{RT}\right) c_A (-\Delta_R h) - \frac{\dot{Q}_C}{\rho c_p V_{\text{CSTR}}^L} \quad (17)$$

Therein,  $k_0$  and  $E$  are the pre-exponential factor and the activation energy of the reaction, respectively,  $\Delta_R h$  is the enthalpy of reaction,  $\rho$  is the density, and  $c_p$  is the heat capacity. Moreover, the cooling duty is a function of  $VP_1$ , whereas the liquid outlet depends on the liquid volume within the reactor and  $VP_2$ . The operating costs depend on the outgoing mole flow of component A, the investments costs depend on the diameter and the height of the CSTR. The objective function is the sum of both costs. The control variables can be changed from one collocation point to another. However, this change is bounded:

$$\|u(t_i) - u(t_{i+1})\| \leq \Delta u_{\text{max}} \quad (15)$$

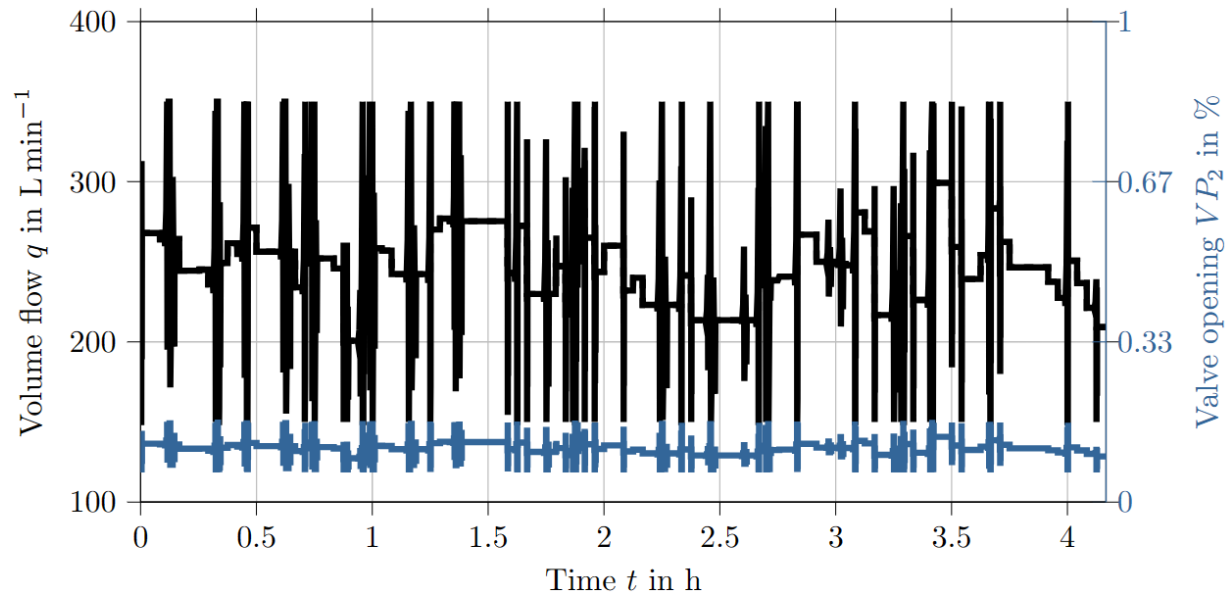
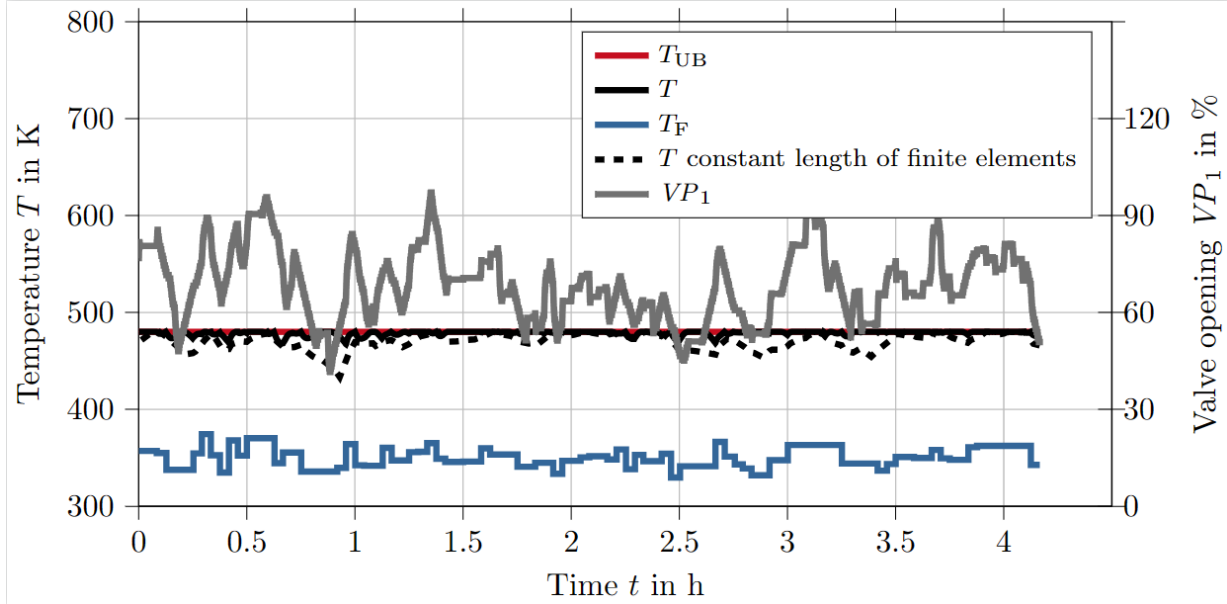
The maximum change of  $u$  is specified by the user. We call this control formulation 1. For the temperature, an upper bound of 480 K is specified. Not violating this upper bound must be ensured by the process design in combination with the allowable range of the control variables. More details on the equations and the specifications for mean and standard deviation of the dynamic variability of feed flow and feed temperature can be found in [6].

For the results shown below, 0.5 was chosen as non-collocation point based on the work by Chen et al. [13], and a tolerance of 0.1 was specified for the discretization error. This value will be critically discussed at the end of the following section.

## RESULTS AND DISCUSSION

The proposed optimization problem in combination with the method for quantifying the discretization error was solved for the presented case study. The obtained dynamic profiles are shown in Figure 2 (top). It can be observed that the temperature (black continuous line) is very close to but below the upper bound, meaning the IDO problem was solved successfully. In blue, the variation of the feed temperature is shown as a result of the APRBS. Lastly, the gray curve visualizes the change of the control variable  $VP_1$ . For comparison, the obtained temperature profile with fixed number and length of finite elements is shown [6]. We frequently observe situations in which the temperature lies below the upper bound (where conversion is the largest) to ensure that the process can be sufficiently cooled.

Figure 2 (bottom) shows, however, the negative impact of the refined discretization scheme. As more and more very small finite elements appear, these elements

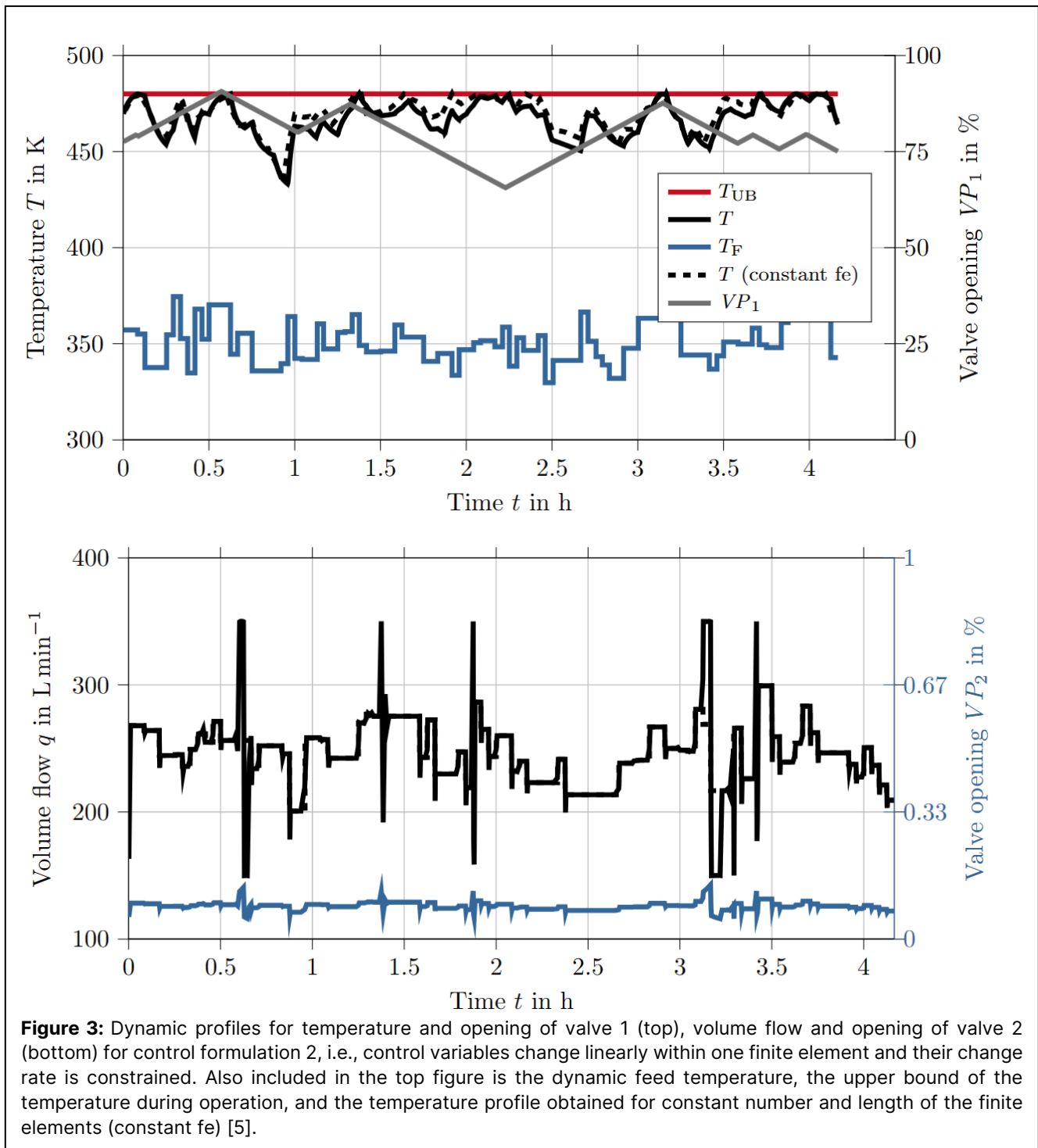


**Figure 2:** Dynamic profiles for temperature and opening of valve 1 (top), volume flow and opening of valve 2 (bottom) for control formulation 1, i.e., control variables change between collocation points and can change by a certain range. Also included in the top figure is the dynamic feed temperature, the upper bound of the temperature during operation, and the temperature profile obtained for constant number and length of the finite elements [6].

allow the same control changes as longer elements. This results in very undesirable bang-bang solutions for the volume flow from the reactor, which would be deemed unrealistic (at best) or would render the process inoperable in practice. As a result of these observations, an alternative formulation had to be found that allowed a varying length of finite elements but also a generically applicable bound on the control changes. For this reason, control formulation 2 is proposed:

$$\frac{du_{i,fe}}{dt} = u_{\text{ramp},i,fe} \text{ with } \|u_{\text{ramp},i,fe}\| \leq u_{\text{ramp,max},i} \quad (18)$$

This means that the control variables are linear functions in each finite element. In addition, the control profiles must be continuous, i.e., the endpoint of control  $i$  in finite element  $fe$  is equal to the starting point of control  $i$  in finite element  $fe+1$ . For this control formulation, the IDO



problem was solved again. The results are shown in Figure 3. The temperature profile is now much closer to the original solution with fixed number and length of the finite elements. The difference is, of course, that the number of finite elements is now much higher (here: 451 instead of 50) and that the profiles of the control variables are linear function in every finite element. In consequence, the IDO problem is successfully solved as well, but does

not show the bang-bang solutions of control formulation 1 (Figure 3, bottom). We conclude that the approach of linearizing the control profiles is suitable in the context of this contribution in which the process design can change to make the process resilient with respect to sudden changes in inputs, such as feed flows. We would necessarily recommend this approach for other applications, such as parameter estimation.

**Table 1:** Design specifications and cost functions for steady-state solution, dynamic solution with fixed number and length of finite elements ("previous solution", taken from [6]), and dynamic solutions with control formulation 1 and 2. The investment costs are given as annuity.

Variable	Unit	Steady-state solution [6]	Previous solution [6]	Control formulation 1	Control formulation 2
$V_{CSTR}$	$m^3$	$4.58 \cdot 10^{-1}$	$7.31 \cdot 10^{-1}$	$4.86 \cdot 10^{-1}$	$8.10 \cdot 10^{-1}$
$d_{CSTR}$	m	$5.26 \cdot 10^{-1}$	$6.15 \cdot 10^{-1}$	$5.37 \cdot 10^{-1}$	$6.37 \cdot 10^{-1}$
$L_{CSTR}$	m	$2.11 \cdot 10^0$	$2.46 \cdot 10^0$	$2.15 \cdot 10^0$	$2.55 \cdot 10^0$
$C_{inv}$	$\$ y^{-1}$	$2.21 \cdot 10^{+2}$	$4.70 \cdot 10^{+2}$	$3.65 \cdot 10^{+2}$	$5.01 \cdot 10^{+2}$
$C_{op}$	$\$ y^{-1}$	$3.51 \cdot 10^{+2}$	$2.29 \cdot 10^{+2}$	$2.17 \cdot 10^{+2}$	$2.38 \cdot 10^{+2}$
$f$	$\$ y^{-1}$	$5.72 \cdot 10^{+2}$	$6.99 \cdot 10^{+2}$	$5.82 \cdot 10^{+2}$	$7.39 \cdot 10^{+2}$

It should be noted here that the approach might benefit from techniques to combine finite elements again where the discretization error is significantly below the set tolerance as, for example, described by Liu et al. [11], to keep the computational effort as low as possible. However, repeatedly adding and removing finite elements could potentially slow down convergence of the overall design problem, which is why this has not yet been implemented.

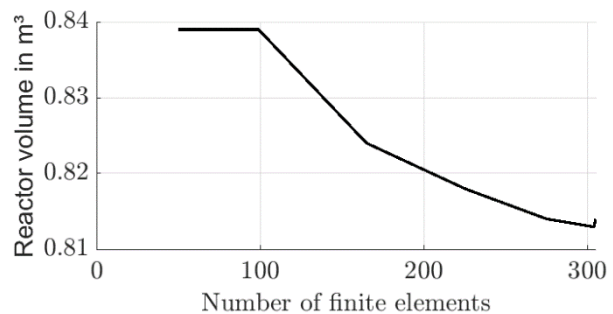
Table 1 compares the design results of our previous study [6] and the results of this work. The smallest reactor volume is still found for the steady-state case as this solution does not have to account for dynamic variability. The solution obtained for control formulation 1 results in a smaller reactor volume (and a smaller objective function) as the short-finite elements provide much more flexibility of the control variables at the cost of bang-band solutions. Control formulation 2 results in a larger reactor volume. This might be due to the combination of the linear profiles for the control variables, but also because constraint violations hidden in the previous solution are now removed because of the controlled discretization error. Lastly, the role of the non-collocation point used for quantifying the discretization error shall be discussed.

**Table 2:** Necessary number of finite elements under variation of the non-collocation point for control formulation 2. The case at 0.9 automatically satisfies the condition because C is negative and thus the error is automatically smaller than zero.

Non-collocation point	Constant C	Number of finite elements
0.1	0.2250	368
0.4	0.0000	50
0.5	0.1250	305
0.6	0.6000	355
0.9	-0.0592	50

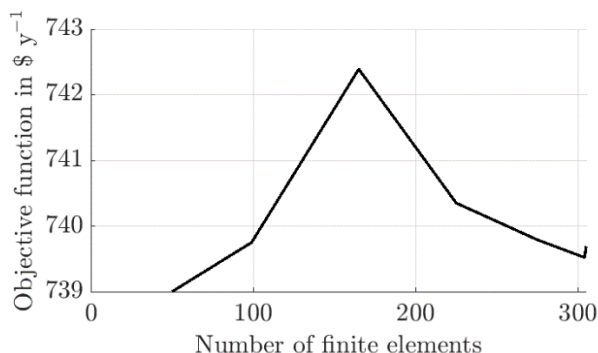
The very pronounced increase of the necessary finite elements by a factor of about 10 was very surprising, given that the system is not very complex. Hence, the specific location of the non-collocation point was varied to determine how the number of finite elements changes (Table 2). It is observed that the location of the non-collocation point has a notable effect on the necessary number of finite elements. Specifically, the constant C has roots in the interval from zero to one (so the discretization error estimate is automatically zero as well), and it can become negative for non-collocation points closer to one. Hence, better criteria for its selection must be found. Until now, this remains an open research question and has, to the author's knowledge, not been discussed so far in other contributions.

The impact of the continuously refined mesh is visualized in Figure 4. The design starts off with a reactor volume of  $0.84 m^3$  and declines to approximately  $0.81 m^3$  (see also Table 1). Although this impact is (in this case small) relatively small, it shows that it should still be investigated, especially for larger flowsheets with varying time constants. For the objective function, however, there is no clear trend observable (Figure 5).



**Figure 4.** Reactor volume with increasing number of finite elements for a non-collocation point of 0.5.





**Figure 5.** Objective function with increasing number of finite elements for a non-collocation point of 0.5.

## CONCLUSION AND OUTLOOK

This contribution extended our recently proposed framework for solving IDO problems by incorporating a measure of the discretization error to avoid incorrect profiles and, in consequence, possibly hidden constraint violations in these badly resolved regions. The extended framework was applied on a CSTR case study for which its geometry, i.e., volume, height, and diameter, was determined while satisfying operational constraints, such as an upper bound of the temperature, for all timepoints in the horizon. Based on the presented findings, a reformulation for control variables is recommended, i.e., the linearization of the profiles within one finite element, which prohibits strongly oscillating control profiles in highly resolved regions of the time horizon.

In the future, this error-controlled approach shall be extended to parametric uncertainty. This is of particular interest as the probability distribution of the uncertain parameters will influence the time constants of the individual scenarios, which will potentially result in completely different numbers and lengths of the finite elements for each uncertainty scenario.

## REFERENCES

1. Ricardez-Sandoval LA, Budman HM, Douglas PL. Integration of design and control for chemical processes: A review of the literature and some recent results. *Annu Rev Control* 33:158–171 (2009).
2. Vega P, Lamanna de Rocco R, Revollar S, Francisco M. Integrated design and control of chemical processes – Part I: Revision and classification. *Comput Chem Eng* 71:602–617 (2014).
3. Hoffmann C, Esche E, Repke J-U. Integration of design and control based on large-scale NLP formulations and an optimal economic NMPC. In: *Proceedings of the 9th International Conference on Foundations of Computer-Aided Process Design*.

- Ed: Garcia Munoz S, Laird CD, Realff MJ. Elsevier (2019).
4. Hoffmann C, Weigert J, Esche E, Repke J-U. Integration of design and operation using dynamic perturbation and chance constraints with unscented transform. In: *30th European Symposium on Computer Aided Process Engineering*. Ed: Pierucci S, Manenti F, Bozzano GL, Manca D. Elsevier (2020).
5. Hoffmann C, Esche E. Integration of design and operation for the CO<sub>2</sub>-based methanol synthesis. In: *14th International Symposium on Process Systems Engineering (PSE 2021+)*. Ed: Yamashita, Yoshiyuki, Kano, Manabu. Elsevier - Health Science (2022).
6. Hoffmann C. A Framework for the Integration of Design and Operation under Uncertainty Using Dynamic Perturbation and Chance Constraints. Dissertation. Technische Universität Berlin, Germany (2023).
7. Nelles O. *Nonlinear System Identification*. Springer (2001).
8. Fourer R., Gay D. M., Kernighan B. W. *AMPL: A Modeling Language for Mathematical Programming*. Duxbury Thomson (2002).
9. Wächter A, Biegler LT. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Program.* 106:25–57 (2006).
10. Ascher U. M., Petzold L. R. *Computer methods for ordinary differential equations and differential-algebraic equations*. Society for Industrial and Applied Mathematics (1998).
11. Liu F, Hager WW, Rao AV. Adaptive mesh refinement method for optimal control using nonsmoothness detection and mesh size reduction. *Journal of the Franklin Institute* 352:4081–4106 (2015).
12. Chen W, Biegler LT. Nested direct transcription optimization for singular optimal control problems. *AIChE J* 62:3611–3627 (2016).
13. Chen W, Ren Y, Zhang G, Biegler LT. A simultaneous approach for singular optimal control based on partial moving grid. *AIChE J* 65:e16584 (2019).
14. Palma-Flores O, Ricardez-Sandoval LA. Selection and refinement of finite elements for optimal design and control: A Hamiltonian function approach. *AIChE J* 69:e18009 (2023).

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Simultaneous Optimization of Design and Operating Conditions for RPB-based CO<sub>2</sub> Capture Process

Howoun Jung<sup>a</sup>, NohJin Park<sup>b</sup>, Jay H. Lee<sup>a\*</sup>

<sup>a</sup> University of Southern California, Mork Family Department of Chemical Engineering and Materials Science, Los Angeles, CA, USA

<sup>b</sup> GS Engineering and Construction Corp., Carbon Solution Research Team/RIF Tech, Seoul, Republic of Korea

\* Corresponding Author: [jlee4140@usc.edu](mailto:jlee4140@usc.edu).

---

## ABSTRACT

Although global efforts for CO<sub>2</sub> capture are underway, large-scale CO<sub>2</sub> capture projects still face economic risks and technical challenges. The Rotating Packed Bed (RPB) provides an alternative solution by mitigating location constraints and enabling a gradual increase in the scale of CO<sub>2</sub> capture through compact modular sizes. However, the main challenge in RPB-based CO<sub>2</sub> capture processes lies in the limited experience with implementing industrial-scale RPB processes. The intricate relationship between RPB unit design, operating conditions, and process performance further complicates the process-level analysis for scale-up. To address these challenges, we propose an optimization-based process design for RPB-based CO<sub>2</sub> capture. Leveraging rigorous process modeling and simulation, we aim to make simultaneous decisions on RPB unit design and operating conditions. Ultimately, our goal is to develop a cost-effective and optimal RPB-based CO<sub>2</sub> capture process, supported by comprehensive cost evaluations. This modularized and cost-effective approach is expected to facilitate rapid implementation and gradual scale-up, thereby reducing entry barriers to CO<sub>2</sub> capture technology for industries.

---

**Keywords:** Carbon Dioxide Capture, Process Intensification, Modelling and Simulations, Process Design, Technoeconomic Analysis

## INTRODUCTION

In the ongoing efforts to combat climate change, CO<sub>2</sub> capture is expected to play a pivotal role for the foreseeable future. Among the available capture techniques, amine-based absorption stands out as a mature technology, with several industrial-scale CO<sub>2</sub> capture facilities in operation worldwide [1]. However, experiences such as the Petra Nova project, one of the largest industrial-scale CO<sub>2</sub> capture projects, have revealed the economic risks and technical complexities associated with building and operating such large-scale CO<sub>2</sub> capture processes. While economies of scale can benefit large-scale processes, the substantial initial cost and space requirement for vast CO<sub>2</sub> capture facilities make them less favorable investments [2]. Process intensification, particularly through technologies like the Rotating Packed Bed (RPB), offers a potential solution to address this hurdle.

RPB, as a form of process intensification, enhances mass transfer and unit throughput with the rotation of a packed bed. The increased centrifugal force widens the

selection window for packing materials and facilitates the use of highly viscous solvents. This improved throughput can significantly reduce the required volume for column units, which typically account for approximately 50% of the capital expenditures (CAPEX), by up to 65% [3].

Driven by the potential benefits of RPB in shrinking mass transfer units, there have been numerous lab-scale experiments [4-6]. On the other hand, its application to industrial-scale CO<sub>2</sub> capture has been limited. Recent studies have explored RPB-based CO<sub>2</sub> capture on a scale of approximately 2200 tons per day (TPD)[7]. However, these studies have overlooked certain important factors, such as the pressure drop, in the RPB design, hindering the provision of realistic insights. Moreover, the overall lack of large-scale RPB-based process implementations poses challenges in extrapolating findings to larger and more practical scales. Notwithstanding the potential economies of scale, the practical viability of large-scale RPB units raises questions, especially considering that their characteristics are more aligned with small-to-medium scale applications.

The central question addressed in this research revolves around establishing a cost-effective design and optimal operating conditions for the RPB-based CO<sub>2</sub> capture process. The efficacy of applying RPB for CO<sub>2</sub> capture at an industrial scale remains uncertain, primarily due to a lack of experience in establishing such processes. Furthermore, limited insights exist regarding unit design guidelines and operating conditions for an industrial-scale RPB-based CO<sub>2</sub> capture process. While Agarwal et al. [8] presented a systematic RPB column design procedure, it is heuristics-based and can only address minimal design requirements to prevent flooding of the liquid or jet ejection of the vapor. Furthermore, the RPB column's design is contingent upon the initial assumptions about some process operating variables, e.g., the rotation speed of packing. The design of the RPB column influences mass transfer and hydraulic phenomena, impacting optimal conditions and overall performance. The inherent characteristics of the RPB demand iterative design and operating condition decisions, pointing to the limited effectiveness of the conventional sequential process design approach.

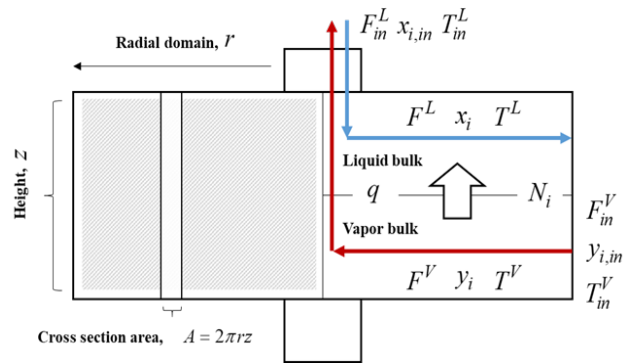
To address these limitations, we adopt a simultaneous optimization-based approach for both RPB design and operating conditions. Utilizing a comprehensive model for an RPB-based CO<sub>2</sub> capture process with a reference MEA solvent, we seek cost-effective RPB design and operating conditions for the compact CO<sub>2</sub> capture process. The envisioned modularized RPB-based CO<sub>2</sub> capture process aims to expedite implementation, enabling a phased scale-up of the CO<sub>2</sub> capture process and thus reducing entry barriers for the industry.

## SYSTEM DESCRIPTION

### RPB Column and Process Model

A process model for the RPB-based CO<sub>2</sub> capture process is constructed using the gPROMS custom modeling environment, building upon the foundation established in our prior publication [9]. The enhancement factor and effective surface area models have been updated for improved accuracy, incorporating a wider range of pilot plant operation data in this study. To ensure consistent insights into design and energy consumption, we selected the widely used MEA amine as the reference solvent. The increased centrifugal force in RPB units facilitates the application of high-viscosity solvents, prompting exploration into concentrated amine solvents, often enhanced with anti-degradation additives. MEA concentrations ranging from 30 to 75wt% MEA solvents are typical for RPB-based CO<sub>2</sub> capture, and we employed the eNRTL (electrolyte Non-Random Two-Liquid) model with updated parameters for this broad range of MEA concentrations [10] as our thermodynamic model. The RPB column model adopts the two-film theory for the

calculation of mass and heat transfer rates, complemented by an enhancement factor model.



**Figure 1.** Conceptual scheme of RPB column model.

**Figure 1** illustrates the conceptual scheme of the developed RPB column model, where vapor and liquid phases flow counter-currently along the radial axis, resulting in the formulation of the partial differential equation (PDE) with radial distribution:

$$\varepsilon^L \frac{\partial C_i^L}{\partial t} = -\frac{1}{2\pi rz} \frac{\partial(F^L x_i)}{\partial r} + a^L N_i \quad (1)$$

$$\varepsilon^V \frac{\partial C_i^V}{\partial t} = \frac{1}{2\pi rz} \frac{\partial(F^V y_i)}{\partial r} - a^L N_i \quad (2)$$

$$\varepsilon^L C_{tot}^L C_p^L \frac{\partial T^L}{\partial t} = -\frac{F^L C_p^L}{2\pi rz} \frac{\partial T^L}{\partial r} + a^L (h^L (T^V - T^L) + N_{H_2O} \Delta H_{H_2O}^{vap} + N_{CO_2} \Delta H_{CO_2}^{abs}) \quad (3)$$

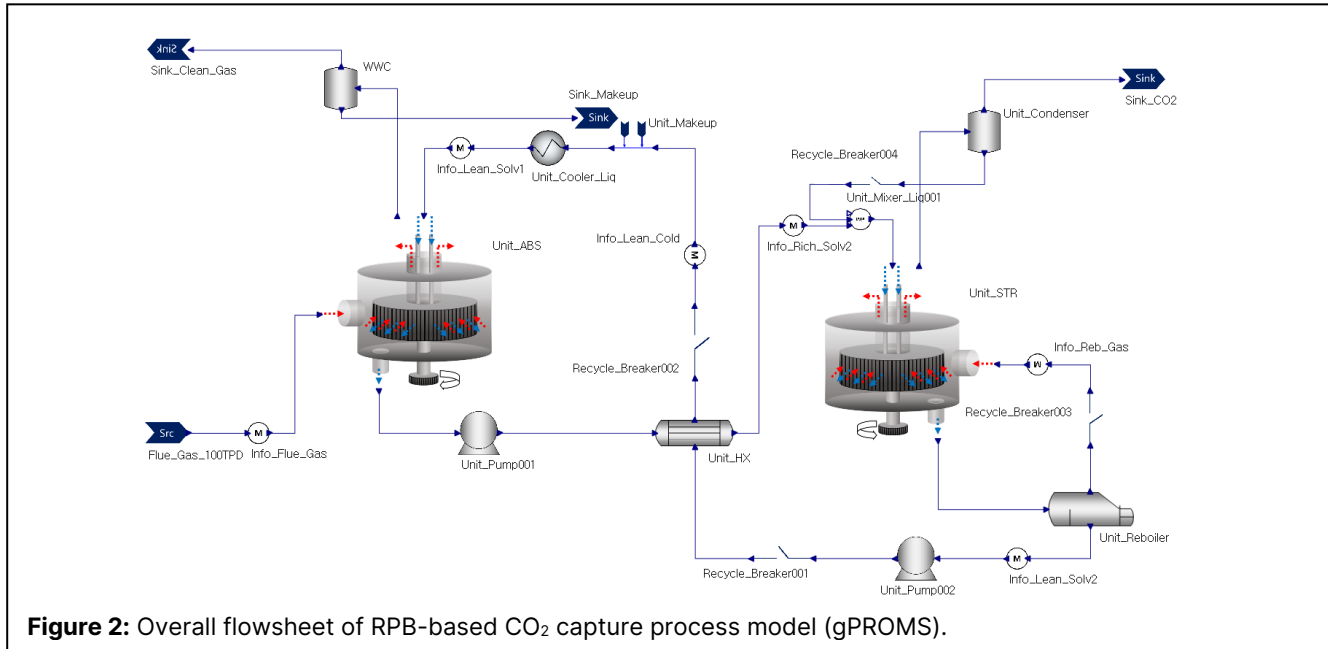
$$\varepsilon^V C_{tot}^V C_p^V \frac{\partial T^V}{\partial t} = \frac{F^V C_p^V}{2\pi rz} \frac{\partial T^V}{\partial r} - a^L h^L (T^L - T^V) \quad (4)$$

Here,  $r_{inner}$ ,  $r_{outer}$  and  $z$  represent the inner radius, outer radius, and height of the RPB unit, respectively.  $C$  and  $F$  denote concentration and molar flowrate, respectively. As the control volume changes along the radial axis, the continuity equations  $F = 2\pi rz C_{tot} u$  hold for both vapor and liquid phases. The mass transfer rate of CO<sub>2</sub>,  $N_{CO_2}$ , is calculated using the two-film theory with an enhancement factor as follows:

$$N_{CO_2} = K_{CO_2}^{overall} (P_{CO_2} - P_{CO_2}^*) \quad (5)$$

$$K_{CO_2}^{overall} = \frac{1}{\frac{RT^V}{k_{CO_2}^V} + \frac{H_{e_i}}{E_{CO_2} k_{CO_2}^L}} \quad (6)$$

Here,  $K_{CO_2}^{overall}$ ,  $P_{CO_2}$  and  $P_{CO_2}^*$  represent the overall mass transfer coefficient, CO<sub>2</sub> partial pressure, and equilibrium CO<sub>2</sub> partial pressure, respectively. **Table 2** provides the transfer correlation and reaction models used for the column model. **Figure 2** illustrates the overall flowsheet of the RPB-based CO<sub>2</sub> capture process, encompassing the developed RPB model and additional process units within



**Figure 2:** Overall flowsheet of RPB-based CO<sub>2</sub> capture process model (gPROMS).

the gPROMS simulation environment.

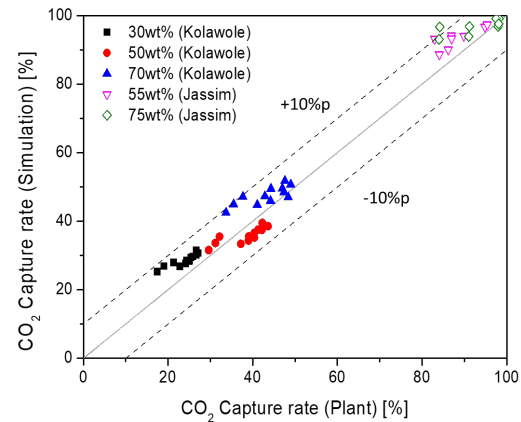
**Table 2:** Transfer correlation and reaction models

Variables	Model
Effective surface area ( $a'$ )	Xie et al. [11]
Liquid transfer coefficient ( $k^L$ )	Tung et al. [12]
Vapor transfer coefficient ( $k^V$ )	Onda et al. [13]
Heat transfer coefficient ( $h'$ )	Chilton-Colburn analogy
Enhancement factor ( $E_{CO_2}$ )	Wellek et al. [14]
Reaction kinetics ( $k_{app}$ )	Luo et al. [15]

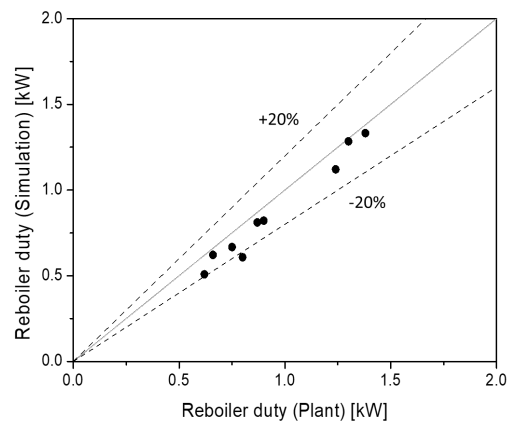
### Process Model Validation

Given the lack of comprehensive operational data for the overall RPB-based CO<sub>2</sub> capture process in existing literature, we subjected our developed process model to a validation process using operation data from separate absorber and stripper pilot plants. We utilized pilot plant operation data from Jassim's (Run 1 to 16) [4] and Kolawole's study (Run 1 to 36) [16] for absorber column validation and Cheng's study (Run 1 to 12) for stripper column validation.

As shown in **Figure 3** and **Figure 4**, our validation results affirm the satisfactory fidelity of the developed RPB column model. The RPB model demonstrates robust predictive capabilities, with an error range of about  $\pm 10\%p$  for CO<sub>2</sub> capture rate and  $\pm 20\%$  for energy consumption. Even under a range of MEA concentration and various operating conditions, the RPB model maintains a mean absolute relative error (MARE) of 11.4 % and 9.2% for CO<sub>2</sub> capture rate and energy consumption, respectively.



**Figure 3.** Comparison of pilot plant and process model (CO<sub>2</sub> capture rate).



**Figure 4.** Comparison of pilot plant and process model (Energy consumption).

## ANALYSIS OF LARGE-SCALE PROCESS

### Target Scale and System

The implementation of an excessively large-scale RPB-based CO<sub>2</sub> capture process faces inherent challenges, given the quadratic increases in both rotational energy and pressure drop. The rotational nature of the RPB introduces escalating demands for momentum and centrifugal energy as the process unit size expands. In addition, the heightened angular acceleration and torque as the plant scale increase can require additional maintenance and safety concerns due to the intricate system control, increased mechanical strain on components, impact on the lifespan of process units, elevated risk of failures, and potential discomfort for operators [17, 18]. Consequently, the preference leans towards favoring an industrial-scale RPB-based process in the small-to-medium range, deviating from the conventional CO<sub>2</sub> capture process with fixed packed beds.

In this study, we explore a medium-scale process at 100 TPD CO<sub>2</sub> capture, specifically targeting flue gas from a coal-fired power plant. This choice aligns with a pragmatic compromise, as the flue gas flow rate for the 100 TPD scale equates to the flue gas from an approximately 6 MW scale power plant. **Table 3** shows the considered inlet flue gas stream condition in this study.

**Table 3:** Inlet flue gas stream condition

Variables	Value
Scale (TPD CO <sub>2</sub> capture)	100
Flow rate (kg/s)	5.94
Temperature (°C)	40
Pressure (bar)	1.1
Mole fraction, CO <sub>2</sub> (%)	14.5
Mole fraction, H <sub>2</sub> O (%)	6.8
Mole fraction, N <sub>2</sub> (%)	76.6
Mole fraction, O <sub>2</sub> (%)	2.1

### Simultaneous Optimization of RPB Design and Operating Condition

The highly coupled nature of RPB design and operating conditions often leads to local optima when a heuristic-based and sequential process design approach is applied. Consequently, the simultaneous determination of process unit design and operating conditions can offer a more cost-effective solution.

In our simultaneous design approach, an optimization problem is formulated to minimize the total annual cost (TAC) per the amount of captured CO<sub>2</sub>, utilizing both RPB design parameters ( $d$ ) and operating variables of the overall process ( $x$ ) as decision variables. The formulation is as follows:

$$\min_{d,x} \frac{TAC}{\dot{m}_{CO_2,Captured}} = \frac{ACC+AOC}{\dot{m}_{CO_2,Captured}} \quad (7)$$

$$d \in [r_{inner,j}, r_{outer,j}, H_j] \quad (8)$$

$$x \in [F_{Solv}, T_{Reb}, P_{Str}, \omega_j] \quad (9)$$

$$\text{s.t. } \eta_{Cap} \geq 90\% \quad (10)$$

$$T_{Reb} \leq 120^\circ\text{C} \quad (11)$$

$$0\% \leq \phi_{flood,j} \leq 80\% \quad (12)$$

$$r_{min,j} \leq r_{inner,j} \leq r_{outer,j} \quad (13)$$

$$H_j/r_{outer,j} \leq 1 \quad (14)$$

In the above formulation,  $ACC$  and  $AOC$  represent the annualized capital and operating cost, respectively. The index  $j$  represents the RPB units, encompassing absorber ( $Abs$ ) and stripper ( $Str$ ) columns. We employ the Lang factor method, based on the free-on-board (FOB) cost with a value of 5.93 for continuous process [19], to estimate the capital cost. Meanwhile the operating cost is calculated from the energy consumption using rigorous process simulation. The details on decision variables as summarized in **Table 4**.

**Table 4:** Considered decision variables

Process variables	Symbols
RPB inner radius (m)	$r_{inner,j}$
RPB outer radius (m)	$r_{outer,j}$
RPB height (m)	$H_j$
Solvent flowrate (kg/s)	$F_{Solv}$
Reboiler temp. (°C)	$T_{Reb}$
Stripping pressure (bar)	$P_{Str}$
RPB rotation speed (RPM)	$\omega_j$

Constraints include a lower bound on 90% capture rate, an upper bound of 120°C on reboiler temperature to prevent thermal degradation of solvent, and a constraint on the 80% flooding condition to govern the RPB design parameters. Additionally, constraints on the minimum RPB inner radii [8] and mechanical recommendations for RPB design [20] are incorporated to ensure the reliability of RPB design.

Due to the absence of a specialized capital cost estimation model for industrial-scale RPB units in existing literature, we adopt a cost estimation model designed for centrifuges. The FOB purchase cost for the RPB unit is assumed to be the aggregate of costs associated with the rotation components (motor and rotor) and the packing bed. Leveraging cost models developed for centrifuges, the cost model unfolds as follows:

$$C_{RPB}^{FOB} = C_{Centrifuge}^{FOB} + V_{RPB}^{Packing} C_{RPB}^{Packing} \quad (15)$$

$$C_{Centrifuge}^{FOB} = \$6180 \cdot (D_{outer})^{0.94} \quad (16)$$

In the above equations,  $V_{RPB}^{Packing}$  and  $C_{RPB}^{Packing}$  represent the volume and cost of packing, respectively, with a



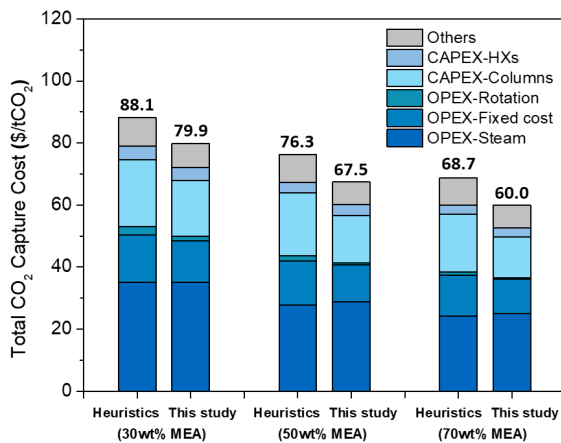
packing price of \$285/ft<sup>3</sup>.  $D_{outer}$  denotes the diameter of centrifuges (RPB in this study) in inches. For utility cost calculations, \$8/GJ, \$19.2/GJ, and \$0.015/GJ are applied for steam, electricity, and cooling water, respectively.

## RESULT AND DISCUSSION

The cost model developed for the simultaneous design approach enables a comparative analysis of CO<sub>2</sub> capture costs against the sequential heuristics-based design. In the heuristics-based method, RPB units were designed using Agarwal's RPB design procedure [8], and operating conditions were determined through an optimization problem aimed at minimizing energy consumption per captured CO<sub>2</sub> as follows:

$$\min_x E_{cap} = E_{steam} + C_{conversion} E_{Elec} \quad (17)$$

The decision variables in eq (9) were employed, and a conversion factor of 1/0.4 was applied to align energy types.



**Figure 5.** Comparison of CO<sub>2</sub> capture cost with heuristic and simultaneous design approach according to varying MEA concentrations.

**Figure 5** illustrates the total CO<sub>2</sub> capture cost at varying concentrations of MEA solvent for both heuristics-based and simultaneous design approaches. The estimated total capture costs with RPB units range from \$60.0 to 88.1/tCO<sub>2</sub>, with the primary contributors being utility cost from steam, RPB column CAPEX, and fixed operation costs. Notably, the energy consumption by packing rotation does not constitute a significant portion of the overall cost. Given that the RPB CAPEX constitutes a large portion of the CO<sub>2</sub> capture cost, obtaining empirical data and developing appropriate cost estimation models will be a crucial for future scale-up studies of RPB-based processes.

The results reveal that employing the simultaneous

decision results in an 9.4–12.7% reduction in CO<sub>2</sub> capture cost. This reduction is predominantly attributed to a decrease in column CAPEX, highlighting the effectiveness of incorporating RPB design parameters into the optimization problem as decision variables. The limitations of the sequential and heuristic design approach become more evident through the optimal design parameters and process variables shown in **Table 5** and **Table 6**. The optimized RPB design parameters and rotational speed exhibit significant differences in both approaches. The optimization results align with our understanding that maintaining the inner radius at a minimum is advantageous. However, the RPB design from the simultaneous solution shows the increased RPB height and reduced outer radius. Such a tall and stout RPB design is expected to contribute not only to reducing RPB CAPEX but also rotation energy and pressure drop, while maintaining the same unit throughput.

**Table 5:** RPB design and optimized process variables (Heuristics-based approach)

Variables	30wt%	50wt%	70wt%
ABS-Inner radius (m)	0.18	0.18	0.18
ABS-Outer radius (m)	1.31	1.28	1.19
ABS-Height (m)	0.36	0.30	0.28
STR-Inner radius (m)	0.10	0.10	0.09
STR-Outer radius (m)	0.35	0.29	0.23
STR-Height (m)	0.43	0.36	0.31
Solvent flowrate (kg/s)	21.2	12.5	9.1
Reboiler temp. (°C)	120*	120*	120*
Stripping pressure (bar)	1.86	1.58	1.08
ABS-RPM (rpm)	491	498	502
STR-RPM (rpm)	503	415	425

\*Constraint active

**Table 6:** Optimized RPB design and process variables (Simultaneous optimization approach)

Variables	30wt%	50wt%	70wt%
ABS-Inner radius (m)	0.18*	0.18*	0.18*
ABS-Outer radius (m)	1.03	0.91	0.80
ABS-Height (m)	1.03	0.91	0.80
STR-Inner radius (m)	0.10*	0.10*	0.09*
STR-Outer radius (m)	0.27	0.21	0.18
STR-Height (m)	0.27	0.21	0.18
Solvent flowrate (kg/s)	21.7	13.8	10.0
Reboiler temp. (°C)	120*	120*	120*
Stripping pressure (bar)	1.88	1.62	1.13
ABS-RPM (rpm)	269	279	292
STR-RPM (rpm)	786	756	888

\*Constraint active

Variations in MEA concentration also exert a significant influence on process performance and total capture costs. Increasing the MEA concentration from 30wt% to 70wt% results in notable savings in energy consumption

and capture cost, ranging from 28.6 to 31.3% and 22.1 to 25.0%, respectively. This observed improvement is mainly attributed to the reduction in the amount of water evaporation heat in the reboiler. Furthermore, there is a notable reduction in the size of the RPB units, which can be attributed to the accelerated CO<sub>2</sub> absorption reaction rate associated with the increased amine concentration. While the widespread adoption of highly concentrated MEA solvent on an industrial scale is constrained by concerns about solvent degradation, the expanded solvent selection window offered by the RPB unit proves advantageous in saving energy consumption and reducing the total CO<sub>2</sub> capture cost.

The interdependence of rotational speed and RPB design is evident, as alterations in the RPB unit design necessitate adjustments in the rotation speed. Particularly in larger-scale processes, lowering the rotation speed becomes advantageous to counteract the quadratic rise in energy consumption with packing rotation.

## CONCLUSION

The RPB-based CO<sub>2</sub> capture process is anticipated to find application across diverse industries owing to its enhanced processing capacity and reduced space requirements. To address the lack of process-level cost evaluation and a systematic design procedure for scaling up on this system, we undertook an exploration of a cost-effective capture process through modeling, simulation, and optimization. Initially, a first principle-based process model was developed and validated. By simultaneously optimizing RPB design and process operating conditions, we realized cost savings of 9.4-12.7% compared to sequential and heuristics-based design approaches. The estimated total CO<sub>2</sub> capture cost with RPB units ranged from \$60.0 to 88.1 per ton of CO<sub>2</sub>, varying with MEA concentration. The optimal RPB design found in this study, characterized by a taller height and shorter diameter, deviates from the heuristics-based approach. Given the strong correlation among RPB design, operating conditions, and process performance, our simultaneous design approach for RPB-based process holds promise for insightful process analysis across various scales in the future.

## ACKNOWLEDGEMENTS

This work was supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) funded by the Ministry of Trade, Industry and Energy (MOTIE) (No. 20224C10300050).

## REFERENCES

1. P. H. M. Feron, A. Cousins, K. Jiang, R. Zhai, and M.

- Garcia, "An update of the benchmark post-combustion CO<sub>2</sub>-capture technology," *Fuel*, vol. 273, pp. 117776, 2020/08/01/, 2020.
2. S. Chen, J. Liu, Q. Zhang, F. Teng, and B. C. McLellan, "A critical review on deployment planning and risk analysis of carbon capture, utilization, and storage (CCUS) toward carbon neutrality," *Renewable and Sustainable Energy Reviews*, vol. 167, pp. 112537, 2022/10/01/, 2022.
3. N. Chamchan, J.-Y. Chang, H.-C. Hsu, J.-L. Kang, D. S. H. Wong, S.-S. Jang, and J.-F. Shen, "Comparison of rotating packed bed and packed bed absorber in pilot plant and model simulation for CO<sub>2</sub> capture," *Journal of the Taiwan Institute of Chemical Engineers*, vol. 73, pp. 20-26, 2017/04/01/, 2017.
4. M. S. Jassim, G. Rochelle, D. Eimer, and C. Ramshaw, "Carbon Dioxide Absorption and Desorption in Aqueous Monoethanolamine Solutions in a Rotating Packed Bed," *Industrial & Engineering Chemistry Research*, vol. 46, no. 9, pp. 2823-2833, 2007/04/01, 2007.
5. C.-H. Yu, H.-H. Cheng, and C.-S. Tan, "CO<sub>2</sub> capture by alkanolamine solutions containing diethylenetriamine and piperazine in a rotating packed bed," *International Journal of Greenhouse Gas Control*, vol. 9, pp. 136-147, 2012/07/01/, 2012.
6. T.-W. Wu, Y.-T. Hung, M.-T. Chen, and C.-S. Tan, "CO<sub>2</sub> capture from natural gas power plants by aqueous PZ/DETA in rotating packed bed," *Separation and Purification Technology*, vol. 186, pp. 309-317, 2017/10/02/, 2017.
7. O. Otitoju, E. Oko, and M. Wang, "Modelling, scale-up and techno-economic assessment of rotating packed bed absorber for CO<sub>2</sub> capture from a 250 MWe combined cycle gas turbine power plant," *Applied Energy*, vol. 335, pp. 120747, 2023/04/01/, 2023.
8. L. Agarwal, V. Pavani, D. P. Rao, and N. Kaistha, "Process Intensification in HiGee Absorption and Distillation: Design Procedure and Applications," *Industrial & Engineering Chemistry Research*, vol. 49, no. 20, pp. 10046-10058, 2010/10/20, 2010.
9. [9] D. Im, H. Jung, and J. H. Lee, "Modeling, simulation and optimization of the rotating packed bed (RPB) absorber and stripper for MEA-based carbon capture," *Computers & Chemical Engineering*, vol. 143, pp. 107102, 2020/12/05/, 2020.
10. T. Nakagaki, H. Isogai, H. Sato, and J. Arakawa, "Updated e-NRTL model for high-concentration MEA aqueous solution by regressing thermodynamic experimental data at high temperatures," *International Journal of Greenhouse Gas Control*, vol. 82, pp. 117-126, 2019/03/01/,

- 2019.
11. P. Xie, X. Lu, H. Ding, X. Yang, D. Ingham, L. Ma, and M. Pourkashanian, "A mesoscale 3D CFD analysis of the liquid flow in a rotating packed bed," *Chemical Engineering Science*, vol. 199, pp. 528-545, 2019/05/18/, 2019.
  12. H.-H. Tung, and R. S. H. Mah, "MODELING LIQUID MASS TRANSFER IN HIGEE SEPARATION PROCESS," *Chemical Engineering Communications*, vol. 39, no. 1-6, pp. 147-153, 1985/12/01, 1985.
  13. K. Onda, H. Takeuchi, and Y. Okumoto, "Mass transfer coefficients between gas and liquid phases in packed columns," *Journal of Chemical Engineering of Japan*, vol. 1, no. 1, pp. 56-62, 1968.
  14. R. M. Wellek, R. J. Brunson, and F. H. Law, "Enhancement factors for gas-absorption with second-order irreversible chemical reaction," *The Canadian Journal of Chemical Engineering*, vol. 56, no. 2, pp. 181-186, 1978.
  15. X. Luo, A. Hartono, S. Hussain, and H. F. Svendsen, "Mass transfer and kinetics of carbon dioxide absorption into loaded aqueous monoethanolamine solutions," *Chemical Engineering Science*, vol. 123, pp. 57-69, 2015/02/17/, 2015.
  16. T. O. Kolawole, "Intensified post-combustion carbon capture using a pilot scale rotating packed bed and monoethanolamine solutions," Newcastle University, 2019.
  17. X. Hai-jun, and S. Jian, "Failure analysis and optimization design of a centrifuge rotor," *Engineering Failure Analysis*, vol. 14, no. 1, pp. 101-109, 2007/01/01/, 2007.
  18. K. Neumann, K. Gladyszewski, K. Groß, H. Qammar, D. Wenzel, A. Górak, and M. Skiborowski, "A guide on the industrial application of rotating packed beds," *Chemical Engineering Research and Design*, vol. 134, pp. 443-462, 2018/06/01/, 2018.
  19. W. D. Seider, D. R. Lewin, J. Seader, S. Widagdo, R. Gani, and K. M. Ng, *Product and process design principles: synthesis, analysis, and evaluation*. John Wiley & Sons, 2017.
  20. D. L. Trent, "Chemical processing in high-gravity fields," *Re-Engineering the Chemical Processing Plant*, pp. 43-75: CRC Press, 2018.

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Beyond Yield: Assessing Reaction System Performance using Economics

Mary A. Katebah<sup>a</sup>, Ma'moun Al-Rawashdeh<sup>a</sup>, and Patrick Linke<sup>a\*</sup>

<sup>a</sup> Texas A&M University, Department of Chemical Engineering, Doha, Qatar

\* Corresponding Author: patrick.linke@qatar.tamu.edu.

---

## ABSTRACT

Early stage exploration of reaction systems, including catalyst selection, operating conditions' specifications, reactor design, and optimization, is critical in the engineering field. It is general practice in the reaction engineering field to explore systems against certain performance metrics, of which yield is one of the most commonly utilized objectives. While the yield provides a quantitative measure of how efficiently reactants are converted into target product(s), its definition is ambiguous, particularly in the presence of side/ incomplete reactions, and multiple products. Most of the yield definitions focus on a specific target product; however, conditions within the reactor search space that provide a maximum yield for one product may not be the same as those for another. Moreover, the presence of other undesired products that are not considered may reduce the overall efficiency of the system. This necessitates the utilization of a more holistic metric that encompasses the value of all the generated products. Attempts to address this consider lumping components into a total yield metric. However, this assumes equal weights on all components without adequately capturing their individual significance on the actual performance. This study proposes the utilization of an "economic-value yield" objective that captures all the products' value by using the market price as a weight factor. The traditional yield metric for the various products is contrasted against the economic one to highlight its ability of providing insight into regions within the reactor search space that are associated with high-value products that are otherwise not observed in the conventional definition. This is illustrated with a case study utilizing propane as a feedstock in the novel piston reactor technology.

---

**Keywords:** Reaction, Technoeconomic Analysis, Reaction Engineering, Propane

## INTRODUCTION

The development of optimized and novel reactors is key for sustainable technology design and development [1]. Early-stage exploration and design of reaction systems is of paramount importance in reaction and chemical engineering for a plethora of reasons including the identification and screening of promising reactions, selection of raw materials and catalysts, understanding reaction mechanisms, optimization of reaction conditions, and assessing the overall feasibility of a reaction system. It is general practice in the catalysis and reaction engineering field to study reaction systems against certain performance metrics that assist in evaluating the efficiency of reactors and chemical reactions. Various objectives are utilized to evaluate reactors, depending on the

characteristics of the reaction and nature of the process. Some of the most commonly used metrics include conversion, selectivity, and yield.

While the yield is a fundamental and widely used metric in chemical engineering and reaction optimization, there is no universally agreed upon definition for such a term. It remains associated with limitations including the ambiguity of its definition, as it can refer to conversion yield, selectivity yield, total yield, or other variations that may result in misinterpretations. Moreover, most of the yield definitions focus primarily on the desired product without accounting for the formation of by-products or undesired side reactions that can be crucial for assessing the system's overall performance. For instance, a reaction with high yield may be deemed undesirable if it results in the production of substantial amounts of

unwanted by-products. Maximizing the yield for one product may not result in the conditions that reflect optimum performance. A more holistic approach is required to account for the contribution of multiple products in a reaction system. Attempts to address this issue consider lumping the components into a total yield term [2]. However, this assumes equal weights on all the generated products. A proper weighted yield term is required to account for the contribution of multiple products in a reaction system based on their significance.

The concept of economics can serve as a powerful tool to practically evaluate chemical reactions. This study proposes the utilization of an “economic-value yield” (EVY) metric, by assigning weights to the various components based on their market price, and contrasts it against traditional yield metric definitions to highlight its ability of providing insight into regions within the reactor search space that are associated with high-value products. To the authors’ knowledge, there are no peer-reviewed publications utilizing metrics that capture the economics of the reaction products when assessing reaction chemistries. Such a metric is particularly helpful in the assessment of novel reactor technologies.

One of the emerging novel reactor concepts is the piston reactor technology, which is essentially a repurposed internal combustion engine that focuses primarily on the production of chemicals, in addition to the power and heat associated with the conventional combustion engine. Its advantages include high temperature and pressure conditions at short residence times, thereby enhancing the performance of numerous chemical reactions. The cyclic nature of the reaction system offers a novel mode of carrying out chemical reactions, where the fast temperature quench enables a target non-equilibrium state where the desired product could be generated at optimal yield and efficiency [3].

In this study, the EVY objective is utilized to assess the performance of the piston reactor technology to screen chemistries using propane as a feedstock to determine economically attractive regions within the reactor search space. Results are contrasted against the traditional yield metric, highlighting the importance of utilizing economics when assessing the performance of reactor systems.

## CONVENTIONAL YIELD METRICS

There is no universally agreed upon definition for the “yield” concept [4] as it is applied in numerous contexts, and its definition varies depending on the application, or process. Table 1 depicts a summary of the most common definitions of yield used in the literature. The presence of multiple descriptions for the yield metric introduces uncertainty, particularly when studies fail to explicitly define the parameter utilized in their

evaluations. This lack of terminology hinders the comparability of research results and findings within the reaction and engineering fields. Most of the published studies use the “overall reactor yield” term, on a molar basis, to define the reactor’s yield [5]. In this case, the yield is defined as the ratio of the target product formed to the reactant fed. This is observed in numerous studies including methanol formation from syngas [6], naphthalene oxidation to phthalic anhydride [7], and hydrogen production from methane reforming [8], to name a few.

The issue with such a definition, and its variations in Table 1, is highlighted when there are multiple products formed, as these common definitions neglect the impact of formation of by-products and undesired products. Conditions representing a high yield of one product are not the same as those for another product, leading to a misrepresentation of the actual conditions that pertain to the reactor’s optimal performance. Moreover, a high yield of a certain component does not necessarily guarantee a high overall performance efficiency particularly if substantial amounts of undesired components are formed.

To alleviate this issue, some studies lump certain components together to formulate a total yield term, as in the case of oxidative coupling of methane (OCM). In OCM, methane, co-fed with oxygen, react to produce primarily ethane and ethylene [2]. Since both ethane and ethylene are the main products, a “C<sub>2</sub> Yield” term is utilized to account for the formation of the two products by summing up the yield of each component. In other studies, a C<sub>2+</sub> yield term is utilized, corresponding to the total yield of ethane and ethylene, in addition to propane and propylene [9]. While this takes into consideration the formulation of other products, it automatically assigns equal weights to them, potentially overlooking the variations in their significance, thereby also leading to a misrepresentation of the overall results.

Moreover, such definitions cannot be applied in situations where the target product is not known from the offset. This is particularly highlighted in the technology assessment and chemistry screening cases phases of early-stage reaction exploration, such as the case of the piston reactor.

## ILLUSTRATION: PISTON REACTOR

As an illustration, the concept of yield is applied on screening chemistries within the piston reactor. The piston reactor is essentially a repurposed internal combustion engine (ICE) that is utilized primarily as a reactor to produce chemicals, in addition to the electricity and heat associated with the conventional ICE [10]. The piston reactor operates in a cyclical manner by



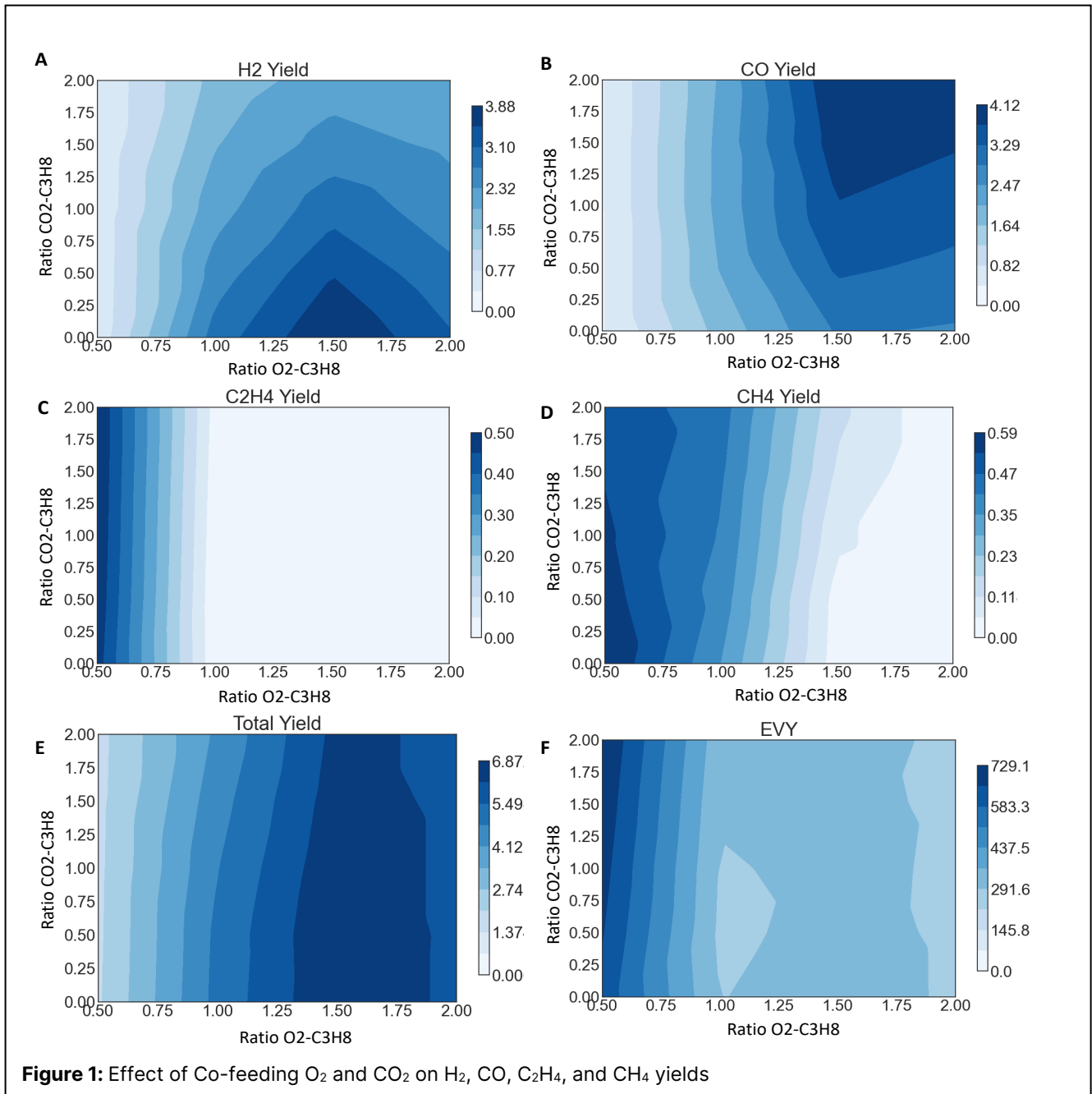
**Table 1:** Common Definitions for the Yield Metric

	Definition	Formula	Ref.
Overall reactor yield	Overall reactor efficiency in converting the reactants to target products	$Y = \frac{\text{Target product formed (mol)}}{\text{Reactant fed (mol)}}$	[17]
Reaction Yield	Ratio of actual amount of product formed to the maximum amount that could be generated (theoretical yield)	$Y_{th} = \frac{\text{Actual amount of product made (mol)}}{\text{Theoretical yield (mol)}}$	[18]
Yield	Amount of desired product generated to the amount formed if there were no by-products and main reaction reached completion	$Y = \frac{\text{Desired product generated (mol)}}{\text{Products generated (no byproducts) (mol)}}$	[19]
Overall fractional Yield	Ratio of all the products formed to the reactants that have reacted	$\Phi = \frac{\text{All product formed (mol)}}{\text{All reacted reactant (mol)}}$	[20]
Instantaneous Fractional Yield	Rate of generation of a desired product relative to rate of consumption of a key reactant	$\varphi = \frac{\text{Rate of product generation}}{\text{Rate of reactant consumption}}$	[7]
Weight Yield	Mass of target product as percentage of total reactant mass	$\text{Weight Yield} = \frac{\text{Target product (mass)}}{\text{Total reactant (mass)}}$	[21]
Total Yield	Ratio of total products formed to the reactor fed	$\text{Total Yield} = \frac{\text{Total products formed}}{\text{Reactant fed}}$	[9]

compressing the feed gas and expanding within short time-frames in the order of milli-seconds [11]. The compression phase results in elevated temperatures and pressures that enhance the efficiency of certain reactions, whereas the expansion results in a rapid quenching of the mixtures enabling the formation of products at optimal conditions [12].

We illustrate the metrics for the case of exploring propane conversion in the piston reactor, based on the piston reactor results of Abousrafa et al. [13]. There are multiple variables that can be explored for a piston reactor applied to propane conversion such as the impact of co-feeding the propane with both O<sub>2</sub> and CO<sub>2</sub> as performed by Abousrafa et al. [13]. Adding O<sub>2</sub> to the reaction promotes the autoignition of propane, resulting in heat generation that can be utilized to co-generate power from the piston reactor. As for CO<sub>2</sub>, in addition to the environmental benefit of utilizing it as a feedstock, CO<sub>2</sub> addition facilitates the occurrence of several reactions such as dry reforming and oxidative dehydrogenation. This analyzes data for an intake temperature and pressure of 773K, 1 bar, and an RPM of 3000 [13]. Products from the reaction include H<sub>2</sub>, CO, C<sub>2</sub>H<sub>4</sub>, CH<sub>4</sub>, C<sub>3</sub>H<sub>6</sub>, C<sub>2</sub>H<sub>2</sub>, C<sub>6</sub>H<sub>6</sub>, and C<sub>4</sub>H<sub>6</sub>. At complete propane conversion, the impact of co-

feeding O<sub>2</sub> and CO<sub>2</sub> on the overall reactor yield of the major generated species is shown in Figure 1 (a-d). For the four main components formed in the piston reactor, H<sub>2</sub>, CO, C<sub>2</sub>H<sub>4</sub>, and CH<sub>4</sub>, maximum yields are obtained at CO<sub>2</sub>/C<sub>3</sub>H<sub>8</sub> and O<sub>2</sub>/C<sub>3</sub>H<sub>8</sub> ratios of 0 and 1.5, 2 and 1.5, 2 and 0.5, 0 and 0.5, respectively. To account for all the components, a total yield metric, similar to the one utilized in the OCM studies [9], is calculated, and the results are illustrated in Figure 1e. The figures show that the yield maxima for all four species and also the total yield are different, highlighting the difficulty in using yield as a performance metric for the optimal search of reactions systems with multiple potentially desired species formed as products. The total yield metric considers all formed species; however, the metric assumes each product with equal weight.



**Table 2:** Price assumptions for the EVY calculation

Component	Formula	Price(\$/t)	Ref.
Hydrogen	$H_2$	1,200	[14]
Carbon Monoxide	$CO$	55	[15]
Methane	$CH_4$	278	[16]
Ethylene	$C_2H_4$	882	[16]
Acetylene	$C_2H_2$	1313	[16]
Benzene	$C_6H_6$	865	[16]
Propene	$C_3H_6$	1488	[16]
Butadiene	$C_4H_6$	1357	[16]

### Economic-Value Yield

Rather than assuming equal weighted total yield or individual component yields as objectives, a differently weighted yield metric that considers the formation of multiple valuable reaction products would be preferred. Considering that the implementation of most production systems is to a large extent based on the value of products formed from raw materials, this work proposes the utilization of an economic-value yield metric for the reactor performance that uses the product prices as a weight

factor to obtain an overall reactor-yield metric:

$$EVY = \sum_j P_j \times WY_j \quad (1)$$

Where:

- EVY is the economic-value yield (\$/t feed)
- $P_j$  is the market value, or price, for product  $j$  (\$/t product)
- $WY_j$  is the weight yield of the product  $j$  as percentage of the total reactant weight (t product/ t reactant)

Table 2 contains the prices of the chemicals produced in the piston reactor. Figure 1f shows the results for EVY at various  $O_2/C_3H_8$  and  $CO_2/C_3H_8$  ratios. Results show that the EVY profile does not match any of the individual yield components. It also shows opposite trends to those observed in the total yield metric (Figure 1e). The maximum total yield is observed at  $CO_2/C_3H_8$  and  $O_2/C_3H_8$  ratios of 0.5 and 1.5, whereas maximum EVY is at  $CO_2/C_3H_8$  and  $O_2/C_3H_8$  ratios of 2 and 0.5. However, these results assume equal weights for all those four main components. Although the maximum EVY is in line with maximum  $C_2H_4$  yield, there are regions within the  $C_2H_4$  figure that show zero yield, yet are associated with substantial value, highlighting the importance of utilizing a properly weighted total yield metric to obtain more representative conclusions. Another important observation is the presence of trace components including  $C_3H_6$ ,  $C_2H_2$ ,  $C_6H_6$ , and  $C_4H_6$  that have a substantial contribution to the EVY. Although they are present in concentrations <2 mol%, and are associated with overall reaction yields <0.1, there are cases where they result in double the EVY. These results further emphasize the importance of adding appropriate weight, in terms of product value, to the conventionally utilized yield terms.

## CONCLUSIONS

It is general practice in the reaction engineering field to evaluate reactions based on certain performance metrics, of which reaction yield is one of the most utilized objectives. There is no universal definition for the yield metric, resulting in challenges including lack of consistency, and results comparability. Moreover, most of the definitions focus on a single desired product without taking into consideration the formation of other undesired or by-products. This leads to difficulty in pinpointing regions of optimal reactor performance, as the yields of one component are associated with different conditions than another. Moreover, the presence of undesired products that are not considered could lead to an inefficient overall performance. A more comprehensive metric is required to account for the contribution of all the generated products in a reaction system. Some studies

attempt to alleviate these issues by lumping the components into a total yield term. However, the assumption of equal weights of components may not be representative of actual performance. This necessitates the utilization of a proper weighted yield term to account for the contribution of all the generated products in a reaction system. This study proposes an “economic-value yield, EVY” metric, where the prices are utilized as weight factors. In addition to its simplicity, using chemical prices reflects actual market value, thereby indicating the relative contribution, or importance, of each generated component. The piston reactor technology is used as an illustration to contrast the EVY results against the traditional yield metric. Results showed that regions within the reactor search space that are associated with high-values do not match the individual yield results, nor the conventional total yield metric. Moreover, the contribution of elements that are associated with low-yield, yet high value, are not shown in the traditional calculations. This highlights the importance of utilizing economics when assessing the performance of reactor systems.

## ACKNOWLEDGEMENTS

This work was made possible by funding from Qatar National Research Fund (QNRF) project number NPRP12S-0304-190222 and co-funding by Qatar Shell Research and Technology Center (QSRTC). The statements made herein are solely the responsibility of the author(s).

## REFERENCES

1. Murtaza G. Chemical reaction engineering: An essential approach towards sustainable technology development. *Ann. Materials Sci. Eng.* 6: 1046 (2022)
2. G. Dimitrakopoulos, B. Koo, B. Yildiz, and A. F. Ghoniem. Highly durable C2 hydrocarbon production via the oxidative coupling of methane using a  $BaFe_{0.9}Zr_{0.1}O_{3-\delta}$  mixed ionic and electronic conducting membrane and  $La_2O_3$  catalyst. *ACS Catal.* 11: 3638–3661 (2021)
3. M. A. Katebah, A. Abusrafa, M. Al-rawashdeh, and P. Linke. Design and analysis of a process for methane to hydrogen conversion using piston reactor technology. *Chem. Eng. Trans.* 88: 829–834 (2021)
4. D. Himmerblau, J. Riggs. Basic Principles and Calculation in Chemical Engineering. Pearson Education. (1989)
5. Pirola Carlo, Rosetti Ilenia, and Ragaini Vittorio. Are conversion, selectivity and yield terms unambiguously defined in chemical and chemical engineering terminology? *La Chimica & L'Industria* (2013)
6. G. Leonzio. Analysis and optimization of a methanol reactor with the adsorption of carbon monoxide and

- water. *Renew. Energy* 146: 2744–2757 (2020)
7. J. J. Carberry. Yield in chemical reactor engineering. *Ind. Eng. Chem.* 5: 40–53 (1966)
  8. H. Messaoudi et al. Hydrogen production over partial oxidation of methane using Ni-Mg-Al spinel catalysts: A kinetic approach. *Comptes Rendus. Chimie* 20: 738–746 (2017)
  9. C. Karakaya, H. Zhu, C. Loebick, J. G. Weissman, and R. J. Kee. A detailed reaction mechanism for oxidative coupling of methane over Mn/Na<sub>2</sub>WO<sub>4</sub>/sio<sub>2</sub> catalyst for non-isothermal conditions. *Catal. today* 312:10–22 (2018)
  10. M. Katebah, A. Abousrafa, M. Al-Rawashdeh, and P. Linke. Hydrogen production using piston reactor technology: Process design and integration for CO<sub>2</sub> emission reduction *Energy J.* 259: 124999 (2022)
  11. A. Ashok, M. Katebah, P. Linke, D. Kumar, D. Arora, K. Fischer, T. Jacobs, M. Al-Rawashdeh. Review of piston reactors for the production of chemicals. *Rev. Chem. Eng.* 39: 1–30 (2023)
  12. B. Atakan et al. Flexible energy conversion and storage via high-temperature gas-phase reactions: The piston engine as a polygeneration reactor,” *Renew. Sust. Energ. Rev.* 133 (2020)
  13. A. Abousrafa, M. Al-Rawashdeh, and P. Linke, “Evaluation of an Electro-Mechanical Piston Reactor for Propane Chemistry,” in AIChE Annual Meeting , Florida (2023).
  14. M. Katebah, M. Al-Rawashdeh, and P. Linke, “Analysis of hydrogen production costs in Steam-Methane Reforming considering integration with electrolysis and CO<sub>2</sub> capture,” *Clean Eng. Technol.* 10:100552 (2022)
  15. J. G. M. S. Monteiro, O. De Queiroz Fernandes Araújo, and J. L. De Medeiros, “Sustainability metrics for eco-technologies assessment, part I: Preliminary screening,” *Clean Technol. Environ. Policy.* 11: 209–214 (2009)
  16. Intratec. <https://www.intratec.us/products/indexes-and-data/utility-cost-database>
  17. R. Smith, *Chemical Process Design and Integration*. Wiley (2008)
  18. E. B. Naumann, *Chemical reactor design, Optimization and Scale-up*. McGraw-Hill (2002)
  19. P. Harriott, *Chemical Reactor Design*. Marcel Dekker. (2003)
  20. O. Levenspiel, *Chemical Reaction Engineering*. Wiley. (1999)
  21. A. McNaught and A. Wilkinson, *Compendium of Chemical Terminology*. Wiley. (1997)

under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared

# Adsorption-based Atmospheric Water Extraction Process: Kinetic Analysis and Stochastic Optimization

Jinsu Kim<sup>a</sup>, Shubham Jamdade<sup>a</sup>, Yanhui Yuan<sup>b</sup>, and Matthew J. Realff<sup>a\*</sup>

<sup>a</sup> Georgia Institute of Technology, Department of Chemical and Biomolecular Engineering, Atlanta, GA, 30332, USA

<sup>b</sup> Avnos, Inc., Los Angeles, CA, 90046, USA

\* Corresponding Author: [matthew.realff@chbe.gatech.edu](mailto:matthew.realff@chbe.gatech.edu).

## ABSTRACT

Adsorption-based Atmospheric Water Extraction (AWE) is an energy-efficient distributed fresh-water supply method. This research focuses on AWE's kinetic analysis and stochastic optimization, investigating the impact of ambient conditions, kinetics, and weather variability. A one-dimensional fixed-bed system was numerically analyzed using the validated isotherm of MIL-100 (Fe), assuming different kinetic parameters within the linear driving force model. Stochastic optimization, based on annual weather data from Georgia (GA), illustrates the influence of weather conditions on AWE process performance, operation, and cost. Our study offers valuable insights for future research, including site selection, adsorbent material development, and process design. We outline three critical areas for further exploration: experimental verification, material screening, and meteorological site selection.

**Keywords:** Atmospheric Water Extraction (AWE), Metal-Organic Framework (MOF), Kinetic Analysis, Meteorological Analysis, Two-stage Stochastic Programming (TSSP)

## 1. INTRODUCTION

Global warming significantly threatens our climate, rendering it increasingly unpredictable. In particular, it changes the natural water cycle, leading to a growing regional freshwater scarcity problem.

The Atmospheric Water Extraction (AWE) process offers a promising solution for sustainable freshwater production (Wang et al., 2022). Among the various AWE techniques, adsorption-based AWE stands out due to its scalability, energy efficiency, and environmental resilience advantages (Xu & Yaghi, 2020).

This study aims to establish a connection between the characteristics of adsorbent materials, kinetic performance, and the variability of weather conditions, contributing to the development of the AWE process from a process systems engineering perspective. In this brief 6-page paper, the AWE process description is very brief; the analysis methods in Sections 2.1 on the S-Shaped Isotherm and Section 2.2 on the Kinetics Analysis are abbreviated, as are many aspects of the process analyses for their temperature and vacuum-swing, adsorption-desorption processes.

## 2. METHOD

### 2.1 S-shaped isotherm

MIL-100 (Fe) was selected as the reference MOF material because of its high capacity for moisture uptake ( $\sim 0.75 \text{ g}_{\text{H}_2\text{O}} \text{ g}_{\text{ads}}^{-1}$ ), and the isotherm model was mathematically formulated in Eqn (1)-(2) by using four isotherm parameters that were previously validated (Kim et al., 2024).  $q_{0,max}$  is the theoretical maximum that the MOF can take moisture inside of the structure,  $K$  is the distributional factor, which is the function of temperature, and  $m$  is the shape factor determining the steepness of the S-shaped isotherm (Sun & Chakraborty, 2014), and  $h_{\text{evap}}$  is the evaporation heat of moisture.

$$q_{eq} = q_{0,max} \frac{KRH^m}{\{1+(K-1)RH^m\}} \quad (1)$$

$$K = \alpha \exp\left(\frac{m(Q_{\text{sat},0} - h_{\text{evap}})}{RT}\right) \quad (2)$$

### 2.2 Kinetic analysis

A one-dimensional model of the AWE contactor is proposed in Eqn (3), and a linear-driving force (LDF)



model for the kinetic analysis is defined in Eqn (4). The simulation conditions, including  $K_{LDF}$  parameter uncertainty, are shown in **Table 1**.

$$\frac{\partial C_i}{\partial t} + \frac{\partial(uC_i)}{\partial z} + \rho_{ads}(1 - \epsilon) S = 0 \quad (3)$$

$$S \equiv \frac{dq}{dt} = k_{LDF}(q_{eq} - q) \quad (4)$$

**Table 1:** Simulation conditions for the numerical analysis.

Physical dimension	Value
Length (L)	200 mm
Area (A)	0.5 m <sup>2</sup>
Charged mass	2.91 kg <sub>MOF</sub> /m <sup>3</sup> <sub>Bed</sub>
Ambient conditions	
Air velocity (u)	1.0 m/s
Temperature (T)	0, 10, 20, 30 °C
Humidity (RH)	20, 40, 60, 80 %
Kinetic parameter	
$K_{LDF}$	0.5, 0.05, 0.005, 0.005 1/s

## 2.3 Key Performance Indicators

For the intrinsic performance measures of the AWE process, the average bed utilization was analyzed in Eqn (5). From an economic perspective, we introduced the cost of moisture capture (CoMC) as an objective function for stochastic programming. The detailed formulation of cost and energy was taken from the previous work (Kim et al., 2024).

$$Bed_{util}(\%) = \frac{q}{q_{eq}} \times 100 \quad (5)$$

$$CoMC (\text{¢ kg}_{H_2O}^{-1}) = \frac{\text{Annualized capital+operating}}{\text{The total captured moisture}} \quad (6)$$

## 2.4 Two-stage stochastic programming

Two-stage stochastic programming (TSSP) was developed to address the variability of ambient weather conditions and ensure the operability of the process.

To solve the 1<sup>st</sup> stage, a grid search was performed over a sufficiently wide range of capital costs. A conservative weather condition was selected to decide the upper bound on capital cost. Capital cost was evenly gridded out in the range of +/-20% of base cost as shown in Eqn (7). After defining the grid, the minimization problem in the 2<sup>nd</sup> stage was iteratively solved in Eqn (8). The binary decision variables represented whether or not to operate at a given weather condition. The  $i^{\text{th}}$  objective function was the sum of the  $i^{\text{th}}$  annualized capital cost and the expectation of the operating cost divided by the expectation of the harvesting performance. Here, the operating cost and the swing capacity were the functions of capital cost,  $C_i$ .

The more capital invested the better the performance in specific weather scenarios but with reduced effectiveness as the capital investment increases further.

The performance decreases at lower capital and eventually this reduced performance increases the overall average cost. The minimum of the overall cost was located within the grid search for all examples we have studied. Should it appear at the extreme of a range a further expansion of the grid should be performed.

By separating the decision-making process from weather uncertainties within the two-stage framework, we achieved scalable solutions for screening potential MOF candidates and identifying a favorable site.

$$C_i = \text{Grid of } [0.8 \min(C_{cons}), 1.2 \max(C_{cons}), 8] \quad (7)$$

$$\text{minimize } Obj_i = \frac{C_i \times f_{an} + E[B^T OPEX(C_i)]}{E[B^T q_{swing}(C_i)]} \quad (8)$$

## 3. RESULTS AND DISCUSSION

### 3.1 Numerical code verification

The original PDE was semi-discretized and then numerically analyzed in MATLAB using an ode15s solver. The simulation results are consistent with those simulated by the commercial software (Oh et al., 2020), gPROMS, and the relative numerical error of the mass balance in Eqn (9)-(11) was under 0.05%.

$$m_1 = Au \int_0^t (C_{init} - C_{exit}) dt \quad (9)$$

$$m_2 = (1 - \epsilon_{bed})(1 - \epsilon_s)\rho_s AL \frac{10^3}{18} \int_0^z q dz \quad (10)$$

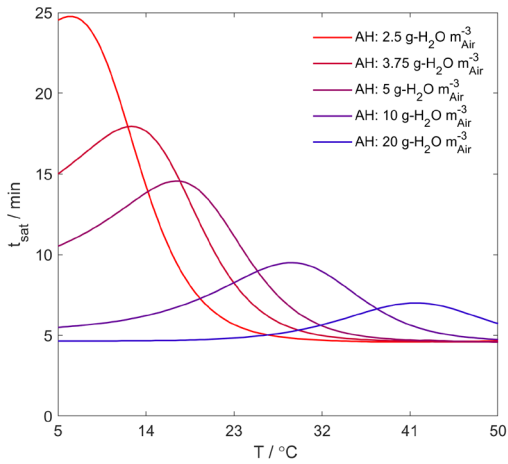
$$Err_{rel} = 100 \times \frac{|m_1 - m_2|}{m_1} \quad (11)$$

The following section illustrates the impacts of ambient conditions and hypothetical kinetic parameters. Finally, we provide TSSP results for the example of the state of Georgia, USA, using weather data.

### 3.2 Saturation time

Saturation time gives direct information for the process design and operation variables such as the dimension of the bed, adsorption time, and required amount of MOF, etc. We showed how the saturation time changes alongside the value of absolute humidity, as shown in **Fig. 1**, by simplifying that the value is relatively unchanged during the diurnal cycle.

This fluctuation implied that the operating variable, adsorption time should be adaptively designed considering the ambient condition. Each variable was parametrically studied, focusing on the breakthrough behavior and bed saturation as follows.



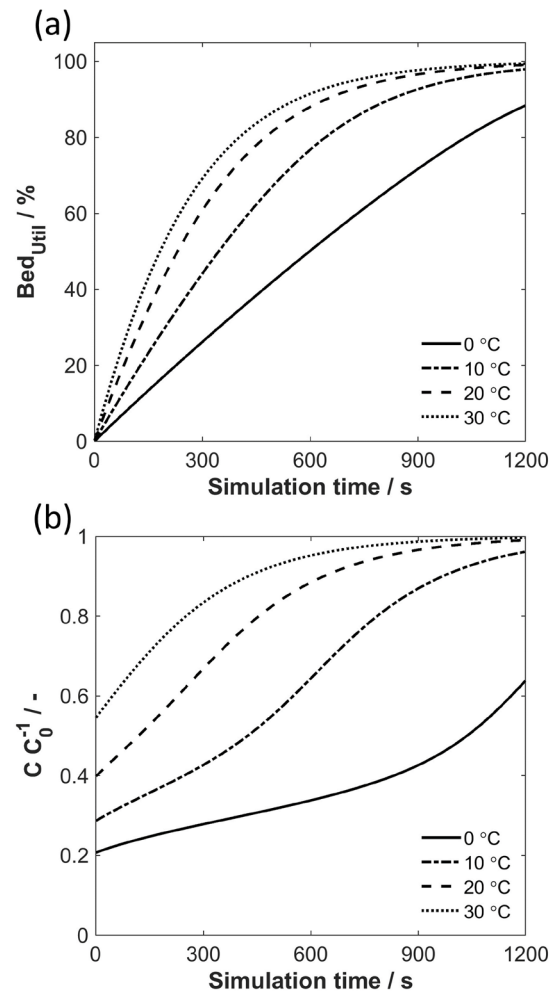
**Figure 1:** Saturation time variation according to the level of absolute humidity.

### 3.3 Effects of ambient temperature

The ambient temperature directly affects the bed utilization ratio, and the high-temperature conditions enable the bed to reach a faster saturation state, as shown in **Fig. 2a**. This was determined by the concentration difference given the fixed kinetic parameter in every condition. In **Fig. 2b**, the normalized exit concentration profiles showed that breakthrough occurred slower in the low-temperature condition, consistent with **Fig. 2a**. The experiments on temperature sensitivity should be supported to support the simulation results, which should be the future work.

### 3.4 Effects of ambient humidity

The effects of relative humidity are shown in **Fig. 3** by fixing the ambient temperature. The RH and the bed utilization were not directly related, as shown in **Fig. 3a**. Firstly,  $q_{eq}$  increased as RH increased following the S-shape. Secondly, the moisture concentration increased as RH increased linearly. These two compensated each other, causing a minimal impact on performance. Although the bed saturation behavior showed similar results, the absolute amount of moisture and the regenerability directly affect the performance, which should be discussed further. As seen in **Fig. 3b**, RH had apparent effects on breakthrough behavior. Because of the low concentration and the low  $q_{eq}$ , the driving force of moisture capture was low in low-temperature conditions.

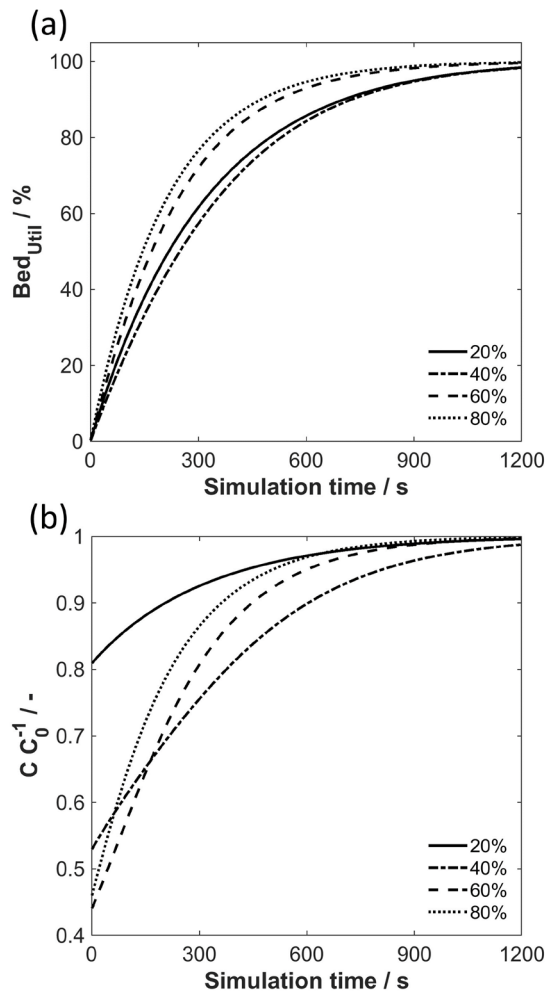


**Figure 2:** The effects of ambient temperature on (a) bed utilization ratio and (b) breakthrough curve on normalized moisture concentration at RH 50% condition ( $k_{LDF}$  assumed as  $0.005 \text{ s}^{-1}$ ).

### 3.5 Effects of kinetic parameter

Finally, the sensitivity of the hypothetical kinetic parameters was analyzed in **Fig. 4**. By taking the ranges of LDF parameters reported in the literature, the expected profiles on the bed utilization and the breakthrough were predicted. Regarding bed utilization, the impact of the kinetic parameters was asymmetric in **Fig. 4a**. As the kinetic parameters became smaller, utilization decreased monotonically. However, the utilization curve converged when the kinetic parameters became more prominent. It failed to observe the breakthrough curve in the slowest case, and the breakthrough behavior in the fastest case was observed in **Fig. 4b**.

For the stochastic optimization, we fixed the  $k_{LDF}$  as  $0.005 \text{ s}^{-1}$ , taken from the moisture uptake systems over Silica gel (El-Sharkawy, 2011).



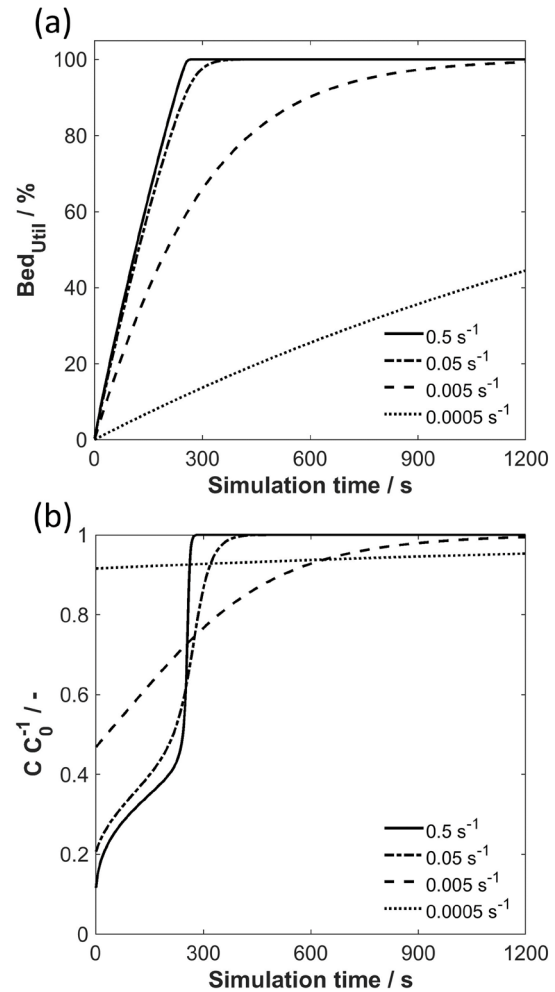
**Figure 3:** The effects of ambient relative humidity on (a) bed utilization ratio and (b) breakthrough curve on normalized moisture concentration at temperature 25°C condition ( $k_{LDF}$  assumed as  $0.005 \text{ s}^{-1}$ ).

### 3.6 Optimization results

The original weather data was plotted as a function of temperature and RH. The data source originated from the Automated Surface Observing Systems (ASOS) of daily measurements during 2022. In **Fig. 5**, probability distribution was plotted, and the 100 random samples were marked in 'x,' representing the original distribution well. The source and sampling algorithm was provided in previous work (Kim et al., 2024). Stochastic optimization was conducted utilizing the random samples.

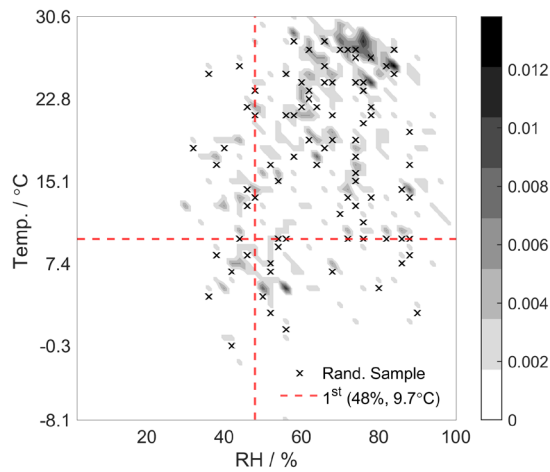
**Fig. 6a** shows stochastic optimization results by dividing the economically feasible and infeasible regions from the programming algorithm. It was observed that the feasible region was clustered in the mild RH range (>40 %) and mild temperature range (2~25°C). **Fig. 6b** shows how the capital cost was selected in the design grid. The programming results were repeated three

times, and the average and standard deviation were plotted to consider the randomness of the sampling. The 5<sup>th</sup> grid point was selected for the first-stage decision because it balanced capital against performance and hence achieved the minimum cost without sacrificing the swing capacity. These results showed the tailored optimal solution based on selecting the optimal capital investment, operating degrees of freedom and conditions in which to operate simultaneously.



**Figure 4:** The effects of kinetic parameter on (a) bed utilization ratio and (b) breakthrough curve on normalized moisture concentration at temperature 25°C and RH 50% condition.

**Fig. 7** demonstrates the relationship between performance and cost. The color shows the economic feasibility, red or blue, for infeasible and feasible respectively. Blue has an average performance of  $0.67 \text{ kg}_{\text{H}_2\text{O}} \text{ kg}_{\text{Ads}}^{-1}$  with  $4.04 \text{ ¢ kg}_{\text{H}_2\text{O}}^{-1}$ . Red has a performance of  $0.36 \text{ kg}_{\text{H}_2\text{O}} \text{ kg}_{\text{Ads}}^{-1}$  with  $5.38 \text{ ¢ kg}_{\text{H}_2\text{O}}^{-1}$ .



**Figure 5:** Weather variability as a joint probability function for Georgia, USA, and random sampling ( $n = 100$ ) for the stochastic optimization.

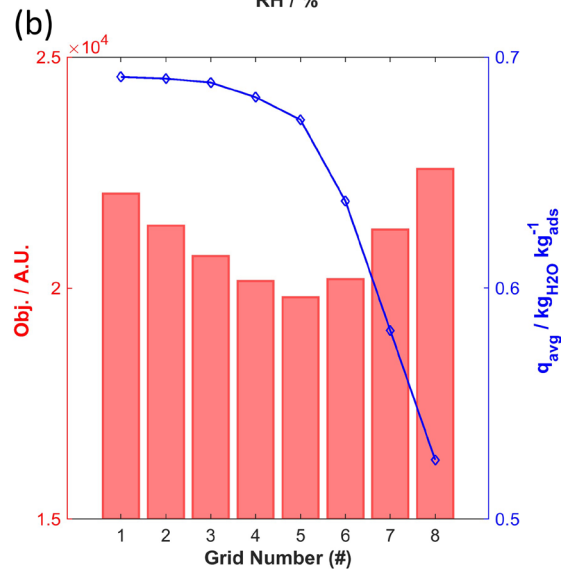
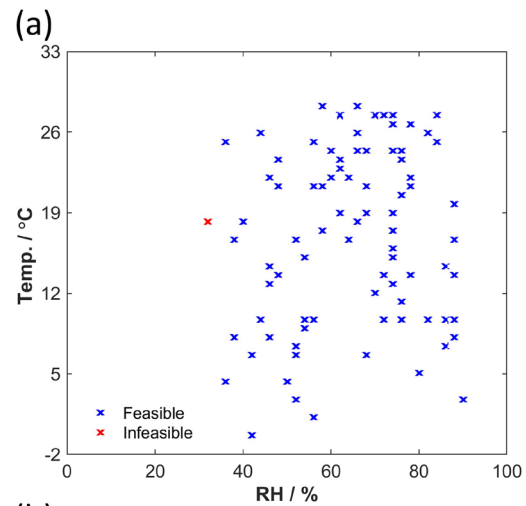
### 3.7 Future work

Three areas should be expanded on to better understand the adsorption-based AWE process. First, experiments are required to refine the kinetic parameter value from the hypothetical range given here. Different temperature and RH conditions should be experimentally measured together for rigorous validation. Second, different types of material should be compared to derive the main factors that influence design and operation, including isotherm shape and moisture uptake. Molecular simulation will be appropriate for the massive screening (Zhang et al., 2023). Lastly, various meteorological sites could be investigated further, including varying the operation based on the time of day.

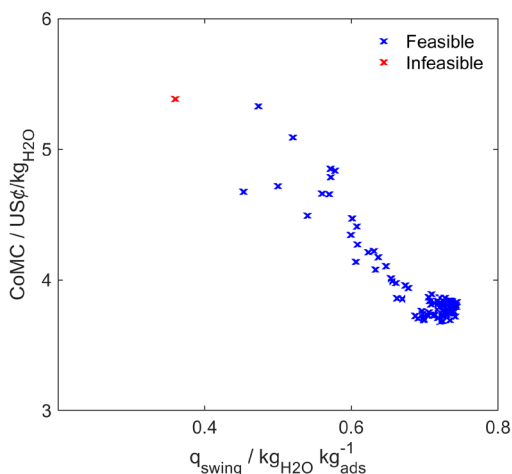
## 4. CONCLUSION

Based on the critical issues summarized herein, the AWE process is promising for producing fresh water. Process design and weather uncertainties are discussed together, starting from the isotherm features at the material level. This work aims to bridge the gap between material and process development to realize robust design and operation. It directs the future research of adsorption-based AWE systems in perspectives on process systems engineering. From this initial work, a generalized framework for simultaneous material-process selection in different climates accounting for weather variability can be developed in the future.

Through this work, we contribute to two areas (i) understanding the system's dynamics and (ii) quantifying the meteorological characteristics, thereby giving insights into process systems engineering (PSE) aspects of the AWE process.



**Figure 6:** Two-stage stochastic programming results for (a) the identification of the feasible weather conditions and (b) the decision of conservative capital cost from average swing capacity.



**Figure 7:** The relationship between water extraction performance and cost of moisture capture according to the weather feasibility in terms of economics.

## ACKNOWLEDGEMENTS

Jinsu Kim was supported by a sub-award to Georgia Tech from ONR N00014-23-C-1011. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Office of Naval Research.

## REFERENCES

1. El-Sharkawy, I. I. (2011). On the linear driving force approximation for adsorption cooling applications. *International Journal of Refrigeration*, 34(3), 667–673.
2. Kim, J., Jamdade, S., Yuan, Y., & Realff, M. J. (2024). System-level analysis of atmospheric water extraction with MIL-100 (Fe) for design and optimal site selection using meteorological characteristics. *Energy*, 299, 131376.
3. Oh, H., Beum, H. T., Yoon, Y.-S., Kim, J., Han, Y., Kim, J., Lee, I.-B., Lee, S.-Y., & Han, S. S. (2020). Experiment and Modeling of Adsorption of CO from Blast Furnace Gas onto CuCl/Boehmite. *Industrial & Engineering Chemistry Research*, 59(26), 12176–12185. <https://doi.org/10.1021/acs.iecr.0c01752>
4. Sun, B., & Chakraborty, A. (2014). Thermodynamic formalism of water uptakes on solid porous adsorbents for adsorption cooling applications. *Applied Physics Letters*, 104(20), 201901.
5. Wang, J., Hua, L., Li, C., & Wang, R. (2022). Atmospheric water harvesting: Critical metrics and challenges. *Energy & Environmental Science*, 15(12), 4867–4871.
6. Xu, W., & Yaghi, O. M. (2020). Metal–organic frameworks for water harvesting from air,

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Optimal Process Synthesis Implementing Phenomena-based Building Blocks and Structural Screening

David Krone<sup>a</sup>, Erik Esche<sup>a\*</sup>, Mirko Skiborowski<sup>b</sup>, and Jens-Uwe Repke<sup>a</sup>

<sup>a</sup> Technische Universität Berlin, Process Dynamics and Operations Group, Berlin, Germany

<sup>b</sup> Hamburg University of Technology, Institute of Process Systems Engineering, Hamburg, Germany

\* Corresponding Author: [erik.esche@tu-berlin.de](mailto:erik.esche@tu-berlin.de).

---

## ABSTRACT

Superstructure optimization for process synthesis is a challenging endeavour typically leading to large scale MINLP formulations. By the combination of phenomena-based building blocks, accurate thermodynamics, and structural screening we obtain a new framework for optimal process synthesis, which overcomes prior limitations regarding solution by deterministic MINLP solvers in combination with accurate thermodynamics. This is facilitated by MOSAICmodeling's generic formulation of models in MathML / XML and subsequent decomposition and code export to GAMS and C++. A branch & bound algorithm is implemented to solve the overall MINLP problem, wherein the structural screening penalizes instances, which are deemed nonsensical and should not be further pursued. The general capabilities of this approach are shown for the distillation-based separation of a ternary system.

---

**Keywords:** Process Synthesis, Optimization, Distillation, Phase Equilibria, Phenomena Building Block

## MOTIVATION & INTRODUCTION

### Optimal Process Synthesis

Thermal separation processes make up a large part of energy consumption in the US and worldwide [1]. Given the drive to reduce greenhouse gas emissions, it is imperative to explore more energy efficient solutions, e.g., through heat integration, novel separation process concepts, or novel equipment, etc. Exploring these options is complex and a continuing challenge. Mathematical methods with mechanistic models for process synthesis help overcome this hurdle. However, process synthesis is quite a challenging field, for which Chen et al. [2] emphasize the delicate trade-off between generality, fidelity, and tractability in process synthesis methods. So far, methods for process synthesis try to reduce the search space either by focusing on individual synthesis tasks (lower generality) or by simplifying models (lower fidelity).

Synthesis methods based on rigorous superstructure optimization promise to overcome these shortcomings. Formulating and solving these process synthesis problems as large-scale mixed-integer nonlinear programming (MINLP) or generalized disjunctive

programming (GDP) problems is tough. While, it is well understood how these should be formulated, tractability is the main challenge here.

Three approaches for superstructure optimization build on phenomena-based formulations. The group of Pistikopoulos developed the generalized modular framework (GMF), which builds on a multipurpose mass / heat transfer module. GMF has, e.g., been applied for synthesis of distillation processes [3, 4] and even reactive distillation [5].

Similarly, the group of Hasan has derived abstract building blocks (ABB), which are arranged in a chessboard-like two-dimensional block superstructure. The boundaries to adjacent blocks and physical attributes of the blocks can be modified during optimization [6, 7]. The ABB formulation is highly versatile. However, application to synthesis of more complex processes so far requires simplification of the ABB model [8] or specialized iterative solution sequences [9].

Finally, Kuhlmann and Skiborowski [10] developed a Phenomena-based Building Blocks (PBB) approach. These building blocks consist of thermodynamically accurate equilibrium stage models and kinetics to form a general state-space superstructure. The capabilities of

this approach have been demonstrated for the synthesis of a membrane reactor [11] and membrane-assisted reactive distillation [12]. So far, solving the formulated MINLPs relied on an evolutionary strategy for the mixed-integer side and local solution of the nonlinear sub-problems in Aspen Custom Modeler. This heavily limits the capabilities of the solvers and poses challenges for the formulation, initialization, and solution of larger examples.

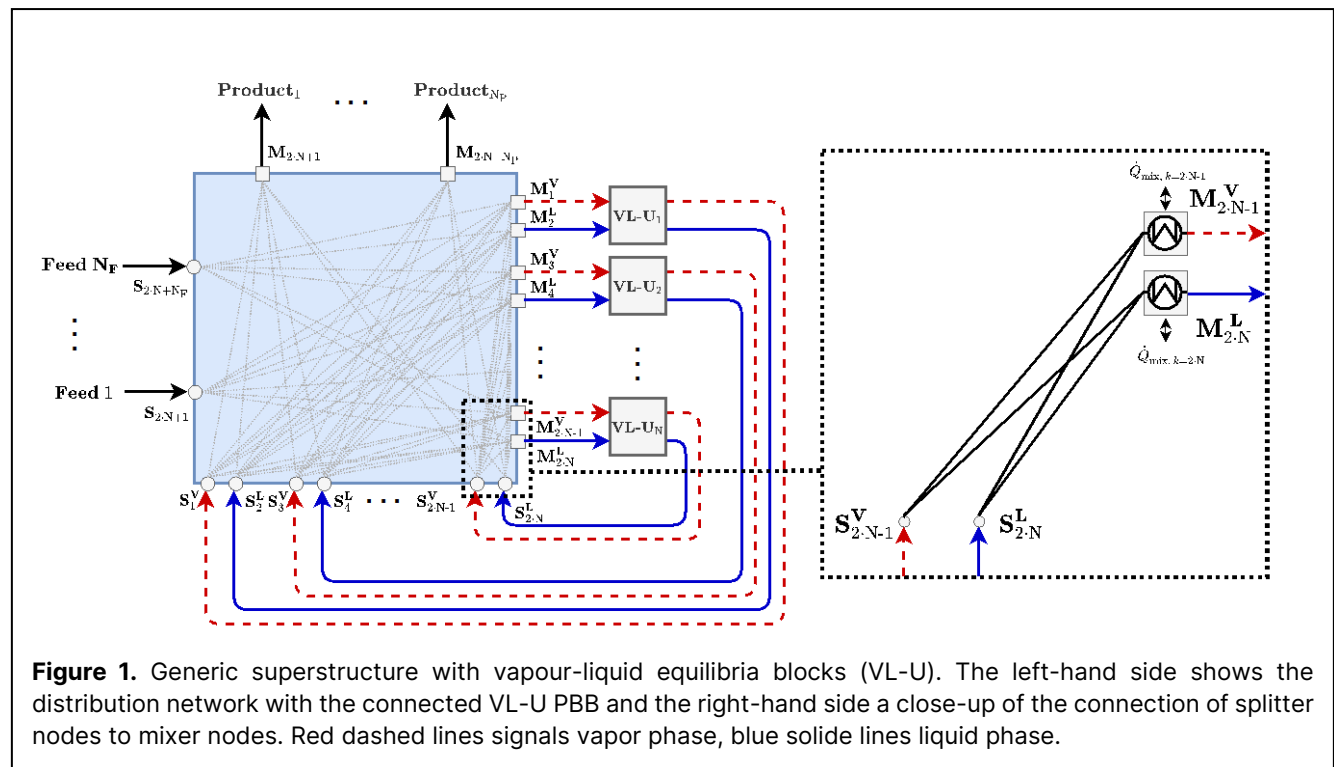
## Challenges & Approach

In the current work, the limitations of the approach by Kuhlmann and Skiborowski [10] will be addressed by suggesting a new framework for formulation and solution of the MINLP problems. In this contribution, we advance the prior work in two directions: (1) formulating and generating large-scale superstructures and (2) facilitating the solution through structural screening. Towards (1) the modeling environment MOSAICmodeling [13, 14] is here extended for the formulation of superstructures of phenomena-based building blocks and their decomposition and code generation for a target language, i.e., the modeling language for the actual solution, e.g., GAMS, AMPL, Python, etc. The implementation in MOSAICmodeling involves formulation of the superstructure problem as a meta model in MathML / XML. With this formulation, executable code for different program components can be obtained automatically, including a mathematical programming platform, an external CAPE-OPEN thermodynamic engine, and a platform for structural screening. This is achieved through the utilization of separate User-defined Language Specifiers (UDLS) [15] within the

web-based modeling and optimization tool MOSAICmodeling, developed at TU Berlin [16]. All thermodynamic property calculations are outsourced to a CAPE-OPEN-compliant property package, while the MINLP is exported to GAMS (version 40.4.0).

Towards the structural screening (2), a number of pruning and screening techniques have been previously suggested, e.g., in [10, 17]. These typically operate at a local level, i.e., the connections surrounding a module or PBB. In our contribution, we deviate from this approach and add towards screening / pruning techniques, which also analyse the entire superstructure instance. The wider perspective aligns with the axioms for generating feasible superstructures within the P-graph framework [18].

Here, a branch and bound-type algorithm is implemented in GAMS and augmented with a middle layer, which pre-screens based on binary decision variables of the superstructure. The middle layer performs a set of graph- and rule-based analyses of a structural instance, which are further detailed in section "Structural Screening". In case of violations of these rules, i.e., an instance will not lead to a physically sensible solution, this structure is penalized without solution of the underlying MINLP. While a similar concept is also present in the approach by Kuhlmann and Skiborowski [10], the MINLP framework extends the screening process with additional rules and algorithms from graph analysis.



**Figure 1.** Generic superstructure with vapour-liquid equilibria blocks (VL-U). The left-hand side shows the distribution network with the connected VL-U PBB and the right-hand side a close-up of the connection of splitter nodes to mixer nodes. Red dashed lines signals vapor phase, blue solide lines liquid phase.

# METHODOLOGY & IMPLEMENTATION

## Generic Superstructure Model

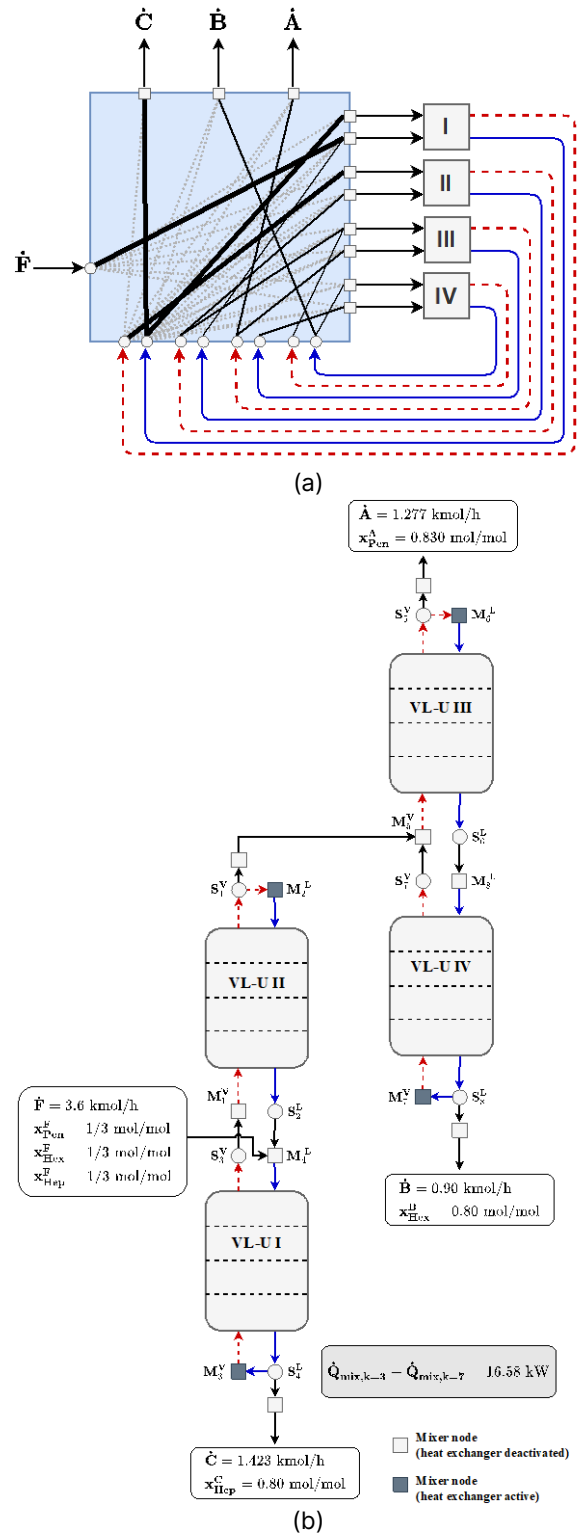
The superstructure formulated here is based on the work of Kuhlmann and Skiborowski [10], which combines a distribution network connecting feed and product streams (blue box in Fig. 1) with a fixed number of PBBs (light grey boxes in Fig. 1). The following extensions are made compared to prior work:

In the distribution network, separate energy flows for pressure manipulation and heat exchange are present. This should allow for deactivation of equipment based on (in-)existence of preceding or subsequent blocks. For the current contribution, we focus on PBBs featuring vapor-liquid equilibrium units (VL-U). In the distribution network, splitter nodes connect the recycles and inlets to mixer nodes and product nodes. Restrictions are implemented on the mixer nodes based on each connected VL-U PBB: By default, the outlet must either be in boiling liquid state or saturated vapor. Binary variables govern the distribution of splitter outlets to mixer nodes. A stream is (for now) split into at most two streams to limit computational complexity. This shall be relaxed later but is deemed for now of low importance. The distribution network changes temperatures before streams enter a PBB.

Here, the formulation for the distribution network is implemented as given on the right-hand side of Fig. 1 (dashed box). Each mixed stream flows through a heat exchanger. The temperature is adjusted according to the input requirements of the connected VL-U, i.e., heated or cooled to the respective boiling or dewpoint temperature of the mixture. Compression and expansion are implemented as isentropic state changes with fixed efficiencies. Each VL-U block is a countercurrent multi-stage contactor in accordance with Kuhlmann and Skiborowski [10]. On each stage, mixing of liquid and vapor and separation take place.

As stated before, phase constraints are enforced on both inlets of a VL-U PBB. This implies that the preceding heat exchangers must be active to ensure that the inlet is either in vapor or liquid state. Nevertheless, as a results of the superstructure optimization, combinations of the VL-U should be able to form conventional equipment, e.g., distillation towers. For these, the heat exchangers at the connecting point of two VL-U sitting "on top of" one another, should be deactivated (see, e.g., VL-U I and VL-U II in Fig. 2).

For this purpose, additional binary variables are included in the model, which (de-)active these heat exchangers and hence also the phase constraints at the respective inlet of the VL-U PBB. This allows for increased flexibility of the overall superstructure. An inactive heat exchanger is bypassed, and the fluid is supplied to the VL-U in whatever state it might currently be in.



**Figure 2.** Optimization results for the separation of a feed stream of n-pentane, n-hexane, and n-heptane by 4 VL-U: (a) resulting connections in the superstructure, (b) interpretation as a sequence of two columns.

Further details regarding the model formulation, e.g., activation / deactivation of heat exchangers, can be

found in [19].

As a result of deactivated heat exchangers, the superstructure also allows, e.g., a vapor stream to enter a VL-U by its dedicated liquid inlet. In addition to this, further uncommon situations within each VL-U may arise during superstructure optimization. Therefore, the VLE formulation on each stage of the VL-U is augmented by relaxation through complementarity constraints [20, 21] as shown in Krone et al. [13] to account for liquid-only or vapor-only scenarios on each stage.

For a given synthesis task, we implement an objective function, which minimizes the total annualized costs consisting of investment costs for heat exchangers, and operation costs for heating utilities. A penalty term is added to the objective for regularization of the complementarity constraints. The penalty converges to zero at a feasible solution.

### Thermodynamic Properties

Within the MathML / XML model in MOSAICmodeling, all thermodynamic properties are denoted as external function calls, i.e., enthalpies, entropies, equilibrium coefficients, temperatures of boiling or dew point. For these, only the appropriate input variables are assigned to compute the desired output, e.g., enthalpy as a function of temperature, pressure, and composition. These are linked in accordance with the CAPE-OPEN standard for thermodynamic engines using COBIA (version 1.2.0.8) as architecture. TEA (version 3.5) provided by AmsterChem is used as thermodynamic engine.

### Code Generation and MINLP Framework

The novel MINLP framework is illustrated in Fig. 3. Starting from MOSAICmodeling, code is automatically

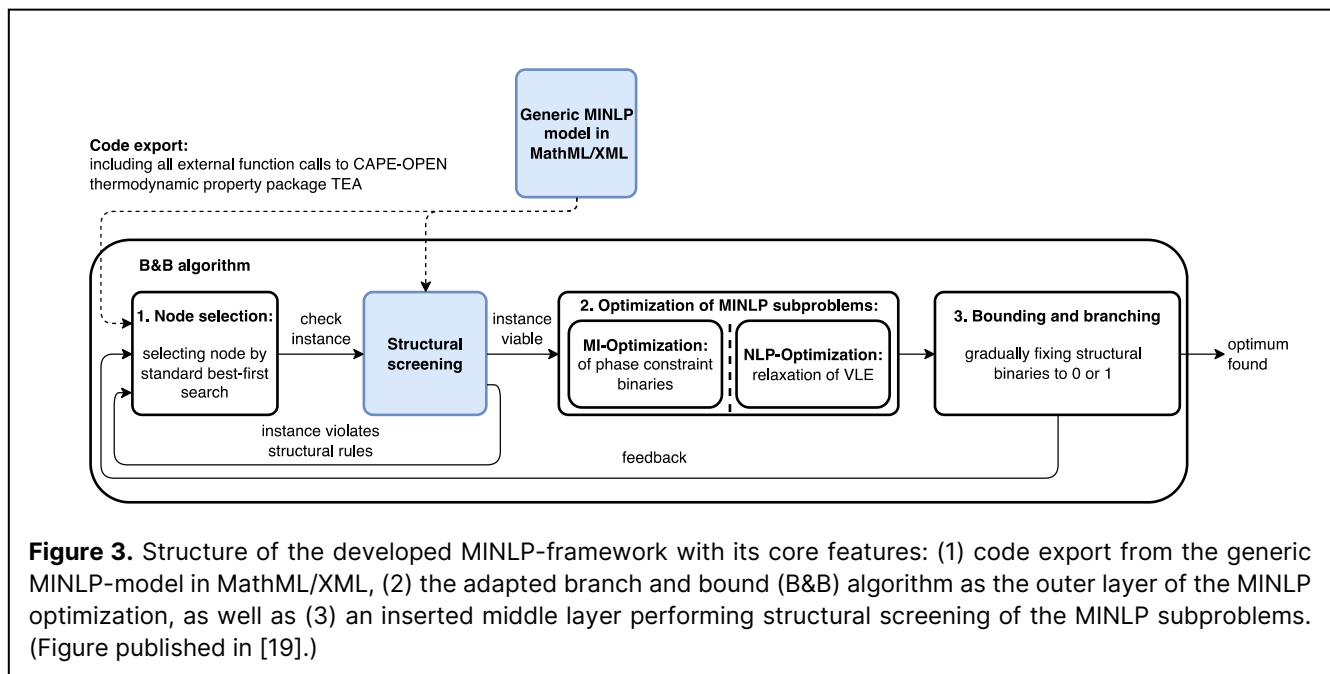
generated for GAMS and C++ (see Krone et al. [13]). The GAMS code then contains the entire MINLP including the state-space superstructure. The thermodynamic properties are marked as external equations and further detailed in the linked C++ code. Therein, the prerequisites for CAPE-OPEN function calls and the COBIA architecture are implemented, function values and derivatives are obtained and returned to GAMS. The overall system is a mixed integer nonlinear programming (MINLP) problem. There are two types of binary variables for the superstructure and for activation or deactivation of phase constraints of the mixers in the network. The superstructure variables by themselves cause a very high computational complexity.

Tractability of the problem is achieved through an augmented solution strategy: The MINLP is solved by a branch and bound algorithm combined with a middle layer that analyzes individual structures. For details on the branch and bound implementation see [19] and [22]. This is implemented in GAMS. The middle layer with the structural screening, however, is formulated within MATLAB (version R2023b).

In each iteration of the branch and bound solver, the current structural instance is passed on to MATLAB, where the screening is carried out regarding the rules specified in the section on “Structural Screening” below. These subproblems are still MINLP as there are further binaries, which are not elemental to the superstructure but form further options within the model formulations, e.g., phase constraints (see above).

### Structural Screening

As a novelty of our contribution, we employ graph- or network-based constraints in addition to algebraic



constraints. At the moment, we assume that the graph-based constraints can only be resolved (efficiently) algorithmically and not implemented as algebraic constraints in a straightforward fashion. This remains to be analyzed in future work.

For graph-based analysis, the current superstructure is exported as a directed graph. The following seven rules are implemented to screen structures at intermediate iterations of the branch and bound algorithm; three of the rules implemented for structural screening are formulated similarly to [10]:

Rule 1: An active VL-U PBB without a connected stream to the inlet mixers needs to be avoided. This PBB will be considered degenerate and the according structure is discarded.

Rule 2 (cf. [10]): During superstructure optimization, it could occur that a subsection of a structural instance is not connected to any feed or product streams – not even indirectly, i.e., a structure which is fully closed to the outside. Such a substructure might cause a host of numerical issues, definitely does not contain any added value, and will hence be discarded.

Rule 3: Within the superstructure several equivalent formulations exist. One example is the interchange of two split streams. To limit the number of equivalencies, rule 3 is implemented to ensure that a connected product node is always served by the first split stream coming from a connected splitter node.

Rule 4: Connections to the product outlets, which appear nonsensical, should be ruled out. For example, this includes connection of two splitter nodes in the distribution network to a product node. This is deemed to be unnecessary entropy generation.

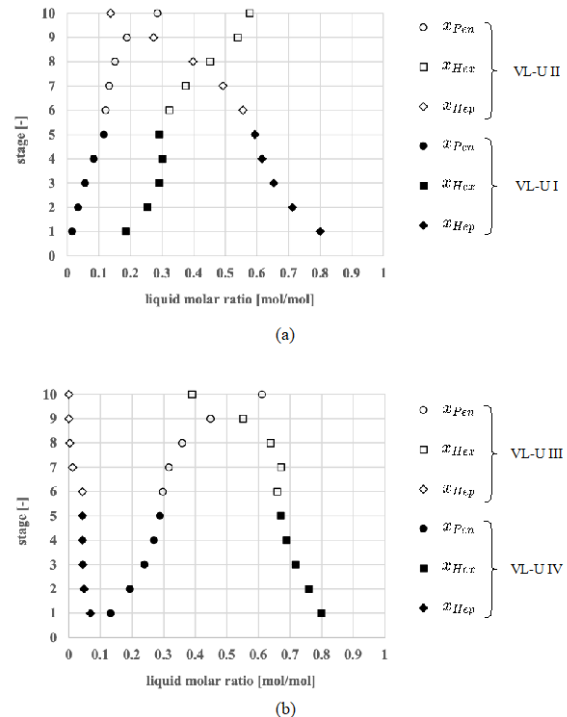
Rule 5 (cf. [10]): A stream that leaves a VL-U PBB should not be fed back to an inlet of the same VL-U PBB. This is considered a “direct recycle”, which also does not appear to be logical from a thermodynamic point of view.

Rule 6: The two inlets of a VL-U PBB should not originate from the same splitter node. Yet again, this rule should be well-founded in thermodynamic considerations and always apply to PBBs of type VL-U.

Rule 7 (cf. [10]): Two outlet streams of a VL-U PBB should not be fed back to the same mixer node. This would counteract any previously achieved separation and hence also amounts to entropy generation.

## CASE STUDY

The formulation described above is applied on a process synthesis task to separate a mixture of n-pentane, n-hexane, and n-heptane by a network of up to six building blocks each representing a counter-current cascade of five vapor-liquid equilibrium stages (VL-U).



**Figure 4.** Liquid composition profiles inside of the optimal structure determined for the separation of an n-pentane, n-hexane, and n-heptane mixture by 4 VL-U: (a) liquid molar ratios on the stages of VL-U I & II, (b) liquid molar ratios on the stages of VL-U III & IV.

This leads to an optimization problem with 1303 (in-)equalities, 62 binary, and 1123 continuous variables. The problem is implemented in MOSAICmodeling and solved in GAMS combining the branch and bound solver with DICOPT [23] for MINLP subproblems, CONOPT3 (version 3.17) [24] for relaxed MINLPs, and IPOPT (version 3.14) [25] for NLP subproblems. The structural screening in Matlab is performed in each iteration of the branch and bound solver.

For now, the system pressure is fixed at 100 kPa and the desired product qualities are specified at purities of 80 mol/mol % each for n-pentane, n-hexane, and n-heptane. The feed stream is specified at 3.6 kmol/h as saturated vapor and features equal amounts of all three components.

## Results

Fig. 2 shows the results for four VL-U PBB in the network (a) and an interpretation as a sequence of two columns (b) and Fig. 4 shows the respective liquid mole fractions of all three components. The results appear to be consistent and are cost-optimal compared to alternative configurations.

This result is obtained after 94 iterations of the Branch & Bound algorithm with roughly 172 h of CPU time. Computation is done on an Intel® Core™ i7-2600



processor (3.8 GHz and 4 cores) with 16 GB of RAM. This is still quite computationally expensive but shows the general feasibility of the overall approach. It should be noted at this point that no particular initialization strategy is applied for the continuous nonlinear variables during the superstructure optimization. Each MINLP and NLP subproblem is solved with the currently available prior values. We assume that a considerable speedup could be achieved by implementing some kind of rigorous initialization strategy or homotopy to guide the NLP solver to the solution.

We would like to stress at this point that all these results are obtained using no specific / manual initialization of the state variables. At the moment, all variables are initialized at the mid-points of their intervals. This mediocre initialization will cause a large part of the CPU time. Furthermore, the VLE formulation including the interfacing via CAPE-OPEN is not yet very efficient. In future work, we shall evaluate different formulations for the external computation to achieve a speedup. Prior work by [11, 5] promise a sizable speedup using different formulations.

The inclusion of the structural screening procedure proved to be vital for the overall set-up. Without the middle layer, the optimizer failed to solve the system at all. In our case studies, we observed behavior which we interpret in two different ways: First of all, some of the otherwise penalized nodes in the branch & bound algorithm consume a lot of time and typically the solution is aborted based on our maximum of around 1000 h. Second, solution to local infeasibility of some of the otherwise penalized subsystems leads to a degradation of the initial values of all continuous variables, which causes infeasibility in a subsequent node, which might otherwise have been successfully solved.

Of the 94 iterations, 23 iterations involved penalization of the current structure based on the screening rules. In these cases, the lower-level solvers would otherwise have probably failed to solve to any reasonable solution, so this also implies a great save in computational expenditure. Overall, we can observe a great contribution of the structural screening and pruning towards ensuring feasibility of these type of MINLP problems. Given that without screening, the optimization did not converge to a feasible solution within the given time frame, we cannot further quantify how great the speed-up due to the structural screening is.

## CONCLUSIONS & OUTLOOK

The presented, novel approach for formulation and solution of process synthesis problems through superstructure optimization can reliably solve the challenging MINLP determining the optimal structure of a separation train and find its energetic optimum. The implemented

structural screening is vital to ensure solvability depending on the problem formulation at hand.

The main advantage in using the MathML / XML-based model formulation lies in flexibility regarding setting up even larger superstructure problems and in flexibility regarding the solution environment. A switch from the GAMS / MATLAB / C++ combination to, e.g., a Python-based setting should be straight forward.

In future work, we shall attempt to further decrease the computational complexity by employing GDP-type formulations for the superstructure and export the model to other platforms, e.g., PYOMO, to speed up the solution and make use of state-of-the-art GDP solvers. Also, we will investigate reduced order models to implement the initialization of the structural instances during the optimization.

## ACKNOWLEDGEMENTS

This work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 523327609.

## REFERENCES

1. Sholl DS, Lively RP. Seven chemical separations to change the world. *Nat* 532, 435–427 (2016)
2. Chen Q, Liu Y, Seastream G, Siirola JD, Grossmann, IE. Pyosyn: A new framework for conceptual design modelling and optimization. *Comp Chem Eng* 153, 107414 (2021)
3. Papalexandri KP, Pistikopoulos EN. Generalized modular representation framework for process synthesis. *AIChE J* 42:4, 1010–1032 (1996)
4. Proios P, Pistikopoulos EN. Generalized modular framework for the representation and synthesis of complex distillation column sequences. *Ind & Eng Chem Res* 44:13, 4656–4675 (2005)
5. Ismail SR, Proios P, Pistikopoulos EN. Modular synthesis framework for combined separation / reaction systems. *AIChE J* 47:3, 629–649 (2001)
6. Demirel SE, Li J, Hasan MF. Systematic Process Intensification using building blocks. *Comp & Chem Eng* 105, 2–38 (2017)
7. Li J, Demirel SE, Hasan MF. Process Integration Using Block Superstructure. *Ind & Eng Chem Res* 57:12, 4377–4398 (2018)
8. Demirel SE, Li J, Hasan MF. Systematic Process Intensification. *Curr Opin Chem Eng* 25, 108–113 (2019)
9. Demirel SE, Li J, El-Halwagi M, Hasan MF. Sustainable Process Intensification Using Building Blocks. *ACS Sustain Chem Eng* 8:48, 17664–17679 (2020)
10. Kuhlmann H, Skiborowski M. Optimization-based

- approach to process synthesis for process intensification: General approach and application to ethanol dehydration. *Ind Eng Chem Res* 56, 13461–13481 (2017)
11. Kuhlmann H, Möller M, Skiborowski M. Analysis of tba-based etbe production by means of an optimization-based process-synthesis approach. *Chem Ing Tech* 32:3, 336–348 (2018)
  12. Kuhlmann H, Veith H, Möller M, Nguyen KP, Górák A, Skiborowski M. Optimization-based approach to process synthesis for process intensification: Synthesis of reaction-separation processes. *Ind Eng Chem Res* 57, 3639–3655 (2018)
  13. Krone D, Esche E, Asprien N, Skiborowski M, Repke JU. Enabling optimization of complex distillation configurations in GAMS with CAPE-OPEN thermodynamic models. *Comp Chem Eng* 157, 107626 (2022)
  14. Esche E, Hoffmann C, Illner M, Müller D, Fillinger S, Tolksdorf G, Bonart H, Wozny G, Repke JU. Mosaic - enabling large-scale equation-based flow sheet optimization. *Chem Ing Tech* 89, 620–635 (2017)
  15. Tolksdorf G, Esche E, Wozny G, Repke JU. Customized code generation based on user specifications for simulation and optimization. *Comp Chem Eng* 121, 670–684 (2019)
  16. Merchan VA, Esche E, Fillinger S, Tolksdorf G, Wozny G, Computer-aided process and plant development. A review of common software tools and methods and comparison against an integrated collaborative approach. *Chem Ing Tech* 88, 50–69 (2016)
  17. Wu W, Henao C, Maravelias C. A superstructure representation, generation, and modeling framework for chemical process synthesis. *AIChE J* 62:9, 3199–3214 (2016)
  18. Friedler F, Tarjan K, Huang YW, Fan LT. Combinatorial algorithms for process synthesis. *Comp Chem Eng* 16, 313–320 (1992)
  19. Krone D, Esche E, Skiborowski M, Repke, JU. Optimization-Based Process Synthesis by Phenomena-Based Building Blocks and an MINLP-Framework Featuring Structural Screening, submitted to *Comp Chem Eng* (2023)
  20. Gopal V, Biegler LT. Smoothing methods for complementarity problems in process engineering. *AIChE J* 45, 1535–1547 (1999)
  21. Lin GH, Fukushima M. A modified relaxation scheme for mathematical programs with complementarity constraints. *Ann Oper Res* 133, 63–84 (2005)
  22. Kalvelagen E. Branch-and-bound methods for an MINLP model with semi-continuous variables, <http://www.amsterdamoptimization.com/pdf/bb.pdf> (2003)
  23. Viswanathan J, Grossmann IE. A combined penalty function and outer-approximation method for MINLP optimization. *Comp Chem Eng* 14:7, 769–782 (1990)
  24. Drud AS. CONOPT - A Large-Scale GRG Code. *ORSA J Comput* 6:2, 207–216 (1994)
  25. Wächter A, Biegler LT. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math Program* 106, 25–57 (2006)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Cybersecurity, Image-Based Control, and Process Design and Instrumentation Selection

Dominic Messina<sup>a</sup>, Akkarakaran Francis Leonard<sup>a</sup>, Ryan Hightower<sup>a</sup>, Kip Nieman<sup>a</sup>, Renee O'Neill<sup>a</sup>, Paloma Beacham<sup>a</sup>, Katie Tyrrell<sup>a</sup>, Muhammad Adnan<sup>a</sup>, and Helen Durand<sup>a\*</sup>

<sup>a</sup> Wayne State University, Department of Chemical Engineering and Materials Science, Detroit, Michigan, USA

\* Corresponding Author: [helen.durand@wayne.edu](mailto:helen.durand@wayne.edu).

## ABSTRACT

Within an Industry 4.0 framework, a variety of new considerations are of increasing importance, such as securing processes against cyberattacks on the control systems or utilizing advances in image processing for image-based control. These new technologies impact relationships between process design and control. In this work, we discuss some of these potential relationships, beginning with a discussion of side channel attacks and what they suggest about ways of evaluating plant design and instrumentation selection, along with controller and security schemes, particularly as more data is collected and there is a move toward an industrial Internet of Things. Next, we highlight how the 3D computer graphics software tool set Blender can be utilized to analyze a variety of considerations related to ensuring safety of plant operation and facilitating the design of assemblies with image-based sensing.

**Keywords:** Industry 4.0, Dynamic Modelling, Nonlinear Model Predictive Control, Simulation, Cybersecurity, Instrumentation, Image-Based Control

## INTRODUCTION

Industry 4.0 is introducing new considerations in production environments, including considerations with respect to cybersecurity, imaging, and control. While these concepts are important considerations for process operation, they also have implications for next-generation design selections (and their interactions with controllers). This work considers the implications of cybersecurity concerns and the application of image-based control on the specifications and design of modern-day processes, as well as their coupling to controllers.

## CYBERSECURITY

In this section, we use a discussion of cryptography and side channel attacks to present possible future research directions related to the intersection of process design, control, and cybersecurity. Traditionally, cybersecurity has been considered to be a problem most relevant to computer scientists and information technology (IT) professionals. The details of how attacks occur can

require an understanding of details of computer hardware and software that typically go beyond traditional chemical engineering fundamentals (e.g., understanding operating system fundamentals related to bootloaders, kernels, and assembly language). However, there has been a growing interest in investigating the relevance of cybersecurity to chemical engineering decisions (e.g., process control [1,2]). Cybersecurity has also begun to be discussed from a process design perspective. For example, in [3], we discussed how different designs lead to different worst-case operating conditions under an attack (similar to an inherent safety perspective). [4] refers to using a Computer Systems/Controls Hazard and Operability Analysis, highlighting interactions between the design of the computer systems and controls safeguards and the process design. This section seeks to make additional connections between process design and control system cybersecurity, with the intent to showcase potential directions in which the process design community might be able to contribute to securing systems in a manner that seeks to promote efficiency. We focus on two areas: design concepts inspired by active detection strategies, and design concepts inspired

by side channel attacks.

## Design Lessons from Active Detection

In the first of the two cybersecurity design perspectives, we discuss learnings from our recent work in cybersecurity of control systems. Our prior work has investigated how to use control signals to disturb the process operation in a manner that would be (ideally) difficult for an attacker to predict, such that they are unlikely to evade detection because they will create process state trajectories that are not in accordance with operating expectations [5,6]. Active detection strategies such as these that attempt to probe for attacks have an advantage over passive detection strategies that they can attempt to use clever operating policies to make it difficult for an attacker to remain undetected. However, they also disrupt operation and thus may be challenging to use in practice. However, this raises the question of whether equipment could be designed that could facilitate locally disruptive behavior but global meeting of process specifications (e.g., through designs that might promote mixing in some areas and laminar flow in others to attempt to allow for active probing with certain cleverly placed actuators and sensors within the design but in a manner that would overall have a minimal impact on actual process output/performance). This analysis indicates that one potential future direction in process design (and particularly in simultaneous design and control) is to analyze whether process and equipment designs that promote an ability to probe for cyberattacks using controllers could be developed.

## Cryptography and Encryption

The second potential design direction related to cybersecurity that we discuss is inspired by side channel attacks that attempt to locate decryption keys by monitoring the power supplied to a computing system. To facilitate the discussion, we begin with a high-level discussion of encryption and then of side channel attacks, and then discuss a conceptual example indicating the potential direction in process design for cybersecurity.

### An Overview of Cryptography

Cryptography has received attention in a control context, including with respect to strategies referred to as homomorphic encryption which is a method that allows for certain mathematical operations to be performed on ciphertext such that the decrypted result of an operation on two ciphertexts is equal to the result of operating on the corresponding plaintexts. This has been considered of interest for investigating the implementation of control laws on encrypted data on the Cloud to attempt to improve the privacy of information which might be sent to the Cloud for processing (e.g., [7,8]). A fundamental aspect of cryptography entails the encoding of publicly-readable 'plaintext' into 'ciphertext' (i.e., encrypted data)

such that meaningful information can only be retrieved by intended parties. A popular kind of encryption is public key cryptography, which utilizes two separate keys to encode and decode information. This is done so that any party may encrypt a message using a widely available public key, but only the intended recipient has access to the secret key needed to decrypt. Public key cryptosystems have four main components: the public key, the secret key, and the encryption and decryption algorithms.

One way to promote privacy of information transferred throughout a control loop is to encrypt state measurements to be sent to a controller, where they must then be decrypted before computations can be performed. The resulting control action is then re-encrypted and returned to the actuators to be decrypted and actuated. This setup protects against so-called man-in-the-middle attacks, in which private information is intercepted in transmission. However, in general, side channel attacks can lead to information being obtained from a computing device that might reveal encryption keys. Thus, we discuss side channel attacks in the following section.

### Side Channel Attacks

Side channel attacks take information from the processes that generate them [9]. This information is known as a trace. Operations in circuits follow cryptographic algorithms, and the implementation of these cryptosystems can leak data about these operations. For example, timing of the messages may reveal information. Power usage is a form of information leakage as a computer would use different amounts of power based on what it is computing. There are different types of power analysis (e.g., differential power analysis (DPA) and simple power analysis (SPA)). As a side channel attack, the focus is on data leakage from an encryption-decryption process. The cryptographic operation will require power as the device computes and this is where the data leakage occurs.

[10] details how power consumption is directly related to data transmission. Data busses, metal wires within the circuitry, function as mini capacitors by charging and discharging as they transfer data between device components. This charging and discharging consumes power. In a data bus, there is a power line, or rail, that represents the states 1 or 0, and there is a ground line. Data is transmitted as states of on, 1, or off, 0. The size of the bus determines how many bits of data can be transmitted. As data is processed and transferred, electronic switches known as FETs, or field-effect transistors, open and close depending on the state.

The state can change depending on what the bit needs to be set to in the data. To change the state, charge is applied or discharged and this requires some work to be done and power consumed. On a data bus line, the transmission follows a counter and every time a bit is set to either 0 or 1. The electronic switches control

the state; if the bit is first set to 0 and then set to 1 in the next iteration, the power rail switch will close and the ground rail switch will open. When the power rail switch closes, 5 volts is now being supplied and the bit is set to 1, or “on”. The 5 volts supplied increases the power consumption, which when graphed visually could be represented by a spike in power consumption. This is important to a side channel attack because data sent through the chip will travel through these data busses, having an effect on the power consumption of the device. As there is a change in state and power is consumed, we can consider that power consumption is equivalent to the number of bits set to 1. Then, when trying to attack a system using power analysis, one can look for similarities in power consumption.

In [11], an example in which plaintext interacts with a secret key through the exclusive or (XOR) operation is performed (XOR yields a true output if only one of two conditions is true); this output is then sent to a look up table of values, which are further processed. Information is obtained after some of these operations from the circuit toward guessing keys for the encryption.

The purpose of a power analysis attack is to find patterns in the power consumption. We would expect that the number of bits set to 1 relates to power consumption, so the power consumption should be similar for outputs that share the same number of 1’s. When attacking, multiple attacks will be done to test different hypothetical keys and then an attempt will be made to determine which key is best supported by the power consumption data. Using the model of encryption and decryption with the hypothetical keys, one can obtain a hypothetical output, with its number of ones. If one of the keys chosen was right, the number of ones should relate to the power consumption, and correlations should be present in the data when evaluated for multiple rounds of the hypothetical key.

A power attack is considered a physical attack because, for example, a digital oscilloscope managed by a computer would be connected to the device under attack. While the device performs its rounds of cryptographic operations, traces of the power consumption are recorded and stored on a computer with the corresponding cryptographic data. Then, an informed guess is made for what selection function to use when partitioning the sets of traces into subsets for the averaging step used for determining if correlations exist between the partitioning strategy and the power consumption data (an example of a selection function is the predicted value of a certain bit). A piece of the hypothetical key is related to selection functions that will be used to define the subsets of traces. The averages of the subsets defined by the selection function outputs are computed for each selection function used in the previous stage. The final stage of

analysis is to analyze the test results with either data visualization or data automation, like scripts, to determine which of the hypothetical keys is best supported by the data to be the unknown key. With the key, one can work backwards to the original plaintext.

## Concepts for Process Design

Power analysis targets the consumption of power in a device to glean information from the possible operations underway; this raises the question of whether an industrial process leaks information that is intended to be encrypted or otherwise protected from attackers. To demonstrate one concept, consider an extreme case in which a process is designed where the process dynamics are fully known, and it is desired to keep track of the energy consumption of the process as a whole since that might be reflective of a sustainability objective. If the power requested of the actuators is exactly what is applied, and the only other sources of power usage are in executing known computations (except the encryption schemes), there may be a possibility that the data on the power usage could contain some information of value in a type of side channel attack based on the discussion above. This indicates that another potential direction for process design (and its intersection with control) with respect to cybersecurity is attempting to identify how process designs, combined with the measurements being taken in an era of Big Data and the industrial Internet of Things, may or may not cause hidden details of algorithms intended to promote security and privacy to be compromised.

## IMAGE-BASED CONTROL

In this section, we move away from the discussion of cybersecurity in process design toward a discussion of the role of an 3D graphics tool set called Blender in design principles, both related to how to design/research safety monitoring strategies (with the aid of simulations that allow testing of visualization components) and in the design of advanced assemblies with image-based actuation.

### Blender

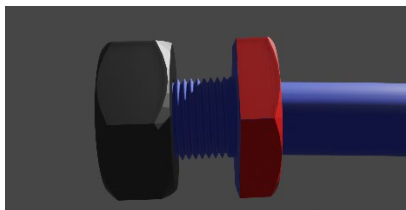
Blender is a 3D graphics tool set with capabilities for modeling, animation, and image rendering. It contains a Python interface that can be used to interface codes with the animations, providing a framework for testing ideas for design and control that require an image component. Our group has recently begun to investigate the potential of the 3D graphics software Blender toward image-based control design, with an example of a render of a rod moving stochastically and under a control policy selected based on value iteration [12]. The image-based control simulation performed using the rod utilizes the coloring of the rod at both ends to help differentiate the



rod from the background for the image processing algorithm. If the rod does not have these types of features, procedures such as edge detection may be required to help identify the boundaries of the rod, and the different sides of the rod may not be visually distinguishable. This example indicates that the selection of the process design (e.g., how the visuals will appear in a camera) can directly affect the available image-based control techniques and methodologies, again highlighting an intersection between design and control for next-generation manufacturing systems. The remaining examples of this section that utilize Blender (focused on object detection for safety purposes and materials design) highlight other potential uses of Blender in ensuring safe plant operation and in setting specifications of advanced assemblies as a step toward designing them.

### Object Detection and Safe Processes

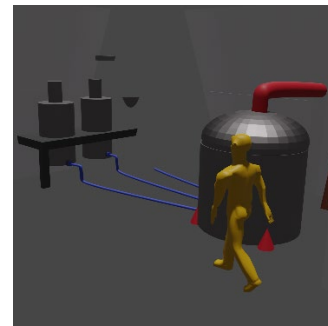
One of the key features of Blender that suggests its utility for researching the testing and designing of safety features at a plant based on images is that it has both visualization capabilities as well as an ability to import image processing Python packages that can then be used to analyze images generated from the software. To showcase this, Blender was used to represent hose and nut assembly shown in Fig. 1, where an object detection algorithm was created utilizing the Python API. The first objective of the code is to create an animation of the nut being loosened, recording a series of key frames. A second Python code was then created that performs an image detection algorithm. The position of the nut is determined by calculating the changes in red channel values in pixels between adjacent pixel rows and columns and subsequently using these to determine a central point as the location of the nut.



**Figure 1.** A render of the hose and nut assembly represented in Blender.

Another key opportunity for using Blender to aid with the design of safety enhancements at a plant is that it can be used to represent non-process components as well, such as human interactions with a process. To see this, consider the human walking in a hose room at a plant as modeled in Fig. 2. One idea for a safety algorithm would be to use images to capture the position of a human in this room and then to map those to whether the human is in a safe or an unsafe area. Without a visuali-

zation software such as Blender for testing such an algorithm, it may be harder for a process systems engineering researcher to contribute to the development of algorithms which can be used for image-based safety enhancements unless they had a physical system. Blender opens the option of being able to generate images for which tests of whether a proposed safety-enhancing algorithm is applicable can then be tested in simulation for research purposes and for better safety algorithm design. For example, using the Python programming interface in Blender, the coordinates of the human can be extracted in the image in Fig. 2 to serve as a ground truth and then compared with any coordinates obtained from an image processing algorithm to analyze how accurate the proposed image processing algorithm is. Blender thus presents a possible strategy for testing non-traditional safety monitoring components.



**Figure 2.** An image of a human walking in a hose room.

### Image-Based Control in Assemblies: Blender to Aid with Specification Determination

In this section, we describe another potential use of Blender toward design, in this case focused on how it might be used in first steps toward designing complex assemblies where images are a component. In this example that showcases its potential, we focus on a stimuli-responsive material, and how to set high-level specifications for how it should respond to visual stimuli that could then be passed downstream in the design pipeline to those who design the molecular structure to see if the material can be designed to meet those specifications that were elucidated through the Blender simulations. Stimuli-responsive material assemblies which react to external signals, including optical, audio, chemical, temperature, and physical signals by means of changing configuration, from the macroscopic assembly to molecular, may have many interesting applications in the future. In this section, we explore the process of creating an initial design of an optical stimuli responsive material assembly. An important aspect of any design phase which we expect to be incorporated in the design of stimuli responsive materials is simulation, in order to understand the dynamics of the material assembly's behavior as well as

controller performance under certain situations; as a motivating example, we develop a simulation within the 3D graphics software Blender to provide insights on what considerations should be made in the design of an optical stimuli responsive material assembly.

### Material Assembly with Image-based Control

In this example, we explore a potential use case for a stimuli-responsive material assembly which is useful when handled properly, however potentially destructive or harmful when misused or exposed to undesired situations. The concept is that we would like to design a strategy for causing this material to “sense” that it is going to be used in the harmful way, and then to break apart when it thinks it will be used in this harmful way to prevent harm. This is a complex design concept that raises many questions, both in terms of how the material should be physically designed (e.g., which molecules may even achieve such a goal), as well as a from a control and sensing point of view. However, we argue that a first step toward attempting to develop such an assembly is to attempt to design the specifications that we want it to follow, which are not obvious. For example, one could consider many ideas for how the assembly should be set up. The material might be intended to break apart immediately when it sees some type of negative signal in image data, or might be intended to do so gradually. The type of negative signals in the images should also be defined to enable testing of whether the proposed material design would work as intended or whether there would be unexpected corner/edge cases for which it breaks in an undesired manner. We suggest that the flexibility of Blender for simulating objects and their interactions physically and through image processing makes it an interesting candidate for developing potential scenarios, evaluating different “breaking” concepts in these scenarios, and then ultimately providing a test framework to evaluate whether the developed algorithms perform as intended in new scenarios (i.e., in the presence of new image-based sensor signals). In the remainder of this section, we show with a simple case study how Blender could be used toward such pathfinding studies for setting specifications for advanced assembly designs involving image sensing, which serves to suggest Blender’s utility for further use and investigation in this direction. This also is meant to motivate discussion on how new paradigms in material design and control are thought of and how they may be used to improve physical safety and provide a line of defense against physical attacks on a system.

We use Blender to create a simple macroscopic model of a non-specific material assembly composed of four blocks connected in a row down the x-axis. Attached to the first block on the long side is a camera pointed away from the assembly down the y-axis, providing optical sensing of the environment. In the camera's field of

view is a block which rotates around the z-axis, where each face is a different color, representing different stimuli. These stimuli will be used to indicate whether the block should “break apart” or not. Views of the environment with the assembly and rotating block are shown in Figures 3 and 4 where a camera can be seen attached to the assembly and facing the rotating block. The camera's initial view of the rotating cube is shown in Figure 5. We assume the assembly is equipped with a controller and actuation capable of separating the end block from the rest of the assembly. The simulation begins with the camera looking at the black face of the rotating cube; when the block rotates to reveal the blue face, the controller “breaks” the cube farthest from the camera off of the assembly. As the cube rotates, each time the center pixel changes colors, another block is broken off of the assembly until only the block with the camera remains.

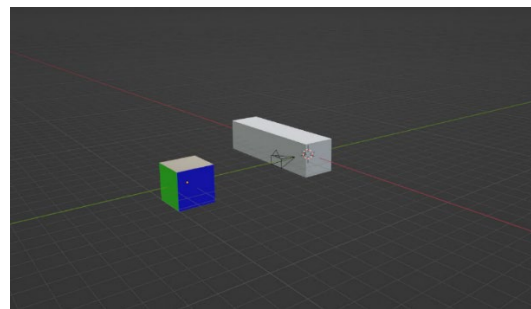


Figure 3. Initial view of assembly and rotating cube.

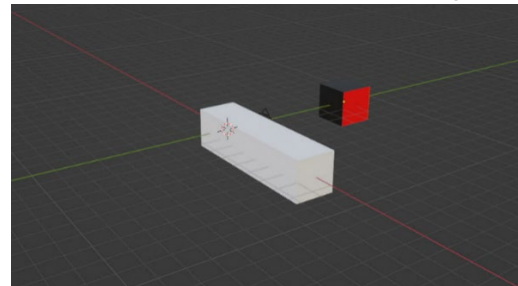


Figure 4. Alternate initial view of assembly and rotating cube.

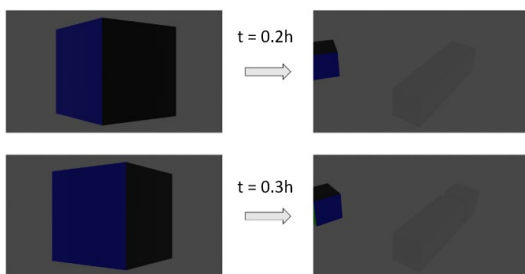


Figure 5. Initial view of rotating cube from the camera.

More specifically, Blender's Python API is used to initialize the scene described above, where four identical, non-interacting  $2 \times 2 \times 2$  unit cubes are placed in a row down the x-axis. The first block is centered at  $(0,0,0)$ , the second block at  $(2,0,0)$ , and so on, so that the blocks are placed to give the appearance of an assembly. Each block is placed with a rotation of  $(0,0,0)$ . The camera is

placed on the block centered at (0,0,0), with a position of (0,1,0) and rotation  $(\pi/2,0,0)$ . The rotating block is placed down the positive y-axis at (0,8,0) with a rotation of (0,0,0). Each of the faces on this block has a different base color, four of which will be seen as the block rotates (the "bottom" and top" faces are both colored black for simplicity). The colors are set using RGBA values (red, green, blue, alpha), where blue is (0,0,1,0), green is (0,1,0,0), red is (1,0,0,0), and black is (0,0,0,0). The block rotates around the z-axis according to the dynamic equation  $\frac{d\theta_z}{dt} = \frac{\pi}{2}$ , numerically integrated using Euler's method with a time step of 0.001 h. At every sampling period  $\Delta = 0.1$  h, a 1920x1080 pixel image of the rotating cube is rendered to a portable network graphics (PNG) file, reflecting the taking of an image by a camera sensor as a measurement of the system. This image is processed by first opening it using the Python package Pillow and loading the image's pixel map to a matrix and checking the RGB value of the center pixel (960,540) to determine which of the four colors it corresponds to. The controller has four modes of action, one associated with each of the possible colors the center pixel can hold in this simulation. Specifically, when the center pixel is black, no action is taken. When it is blue, the first cube is moved away from the assembly. When it is green, the first and second cube are moved away from the assembly, and when it is red, all three cubes are moved away from the last cube holding the camera. When signaled, each cube moves down the x-axis according to the dynamic equation  $\frac{dx}{dt} = 2$ , again numerically integrated using Euler's method and the same time step as before. The simulation is run for 2 h (equating to one full rotation of the rotating cube so that each control action is performed).

In this simulation, control actions are applied at the beginning of each sampling period based on a signal received from the image. Figure 6 demonstrates the control action applied to the first block. Specifically, at time  $t = 0.2$  h, no control action has been applied as the black face still occupies the center pixel. Between  $t = 0.2$  h and  $t = 0.3$  h, the blue face crosses over the center pixel, however the control action is not applied until  $t = 0.3$  h when the first block begins to break off from the assembly.



**Figure 6.** View the assembly and camera view of the first

control action taken at  $t = 0.3$  h (bottom left; the effect of the action at 0.3 h is shown at 0.4 h on the bottom right) compared to the system at  $t = 0.2$  h (top left; the effect of the action at 0.2 h is shown at 0.3 h on the top right).

### Remark 1

One important step in verifying the performance of the controller was determining how to use pixel data to set the behavior of the controller. In this simple case, this was "calibrated" by determining the RGB value returned by the loaded image for each color in the set of colors the controller is to see, where black corresponded to a value of (9,9,9), blue to (9,9,73), green to (9,73,9), and red to (73,9,9). When the image processing algorithm returned one of these values, a control action associated with the value is applied. This however is highly idealized; many aspects which require consideration in real processes could be analyzed using this testbed, including the effects of material properties, lighting properties, and sensor measurement noise. For example, one could imagine that instead of one discrete pixel value being used to represent a color, a range of similar values may be needed to account for variations caused by lighting, or perhaps similarly a number of pixels may need to be analyzed so that pixels affected by intense lighting variations such as glare do not negatively impact the controller.

### Remark 2

Colors were set using filter intensities between 0 and 1, however .PNG files store the values between 0 and 255. It is noted that the value of the pixels read back may not correspond to what is expected analytically; for example, setting the color blue as (0,0,1,0), one may expect a read back of a blue pixel to be (0,0,255) (where the alpha filter is not considered), however, a value of (9,9,73) was found. This sheds light on how the complexities of using a simulated environment for setting colors as well as capturing and processing images need to be carefully considered in the design of a controller which utilizes image data.

### Remark 3

The assumptions that the assembly is equipped with a controller and actuation is non-trivial. With regard to a controller, depending on the size of the assembly, it may be difficult to integrate the proper hardware with the assembly. Similar problems arise with actuation, however this is further complicated by considerations of the dynamics of the assembly which are desired under a stimulus. For example, the controlled material assembly in this simulation can be thought of to be progressively "breaking" as it is exposed to certain optical stimuli; depending on the design and material of the device, it may be difficult to not only actually provide actuation which breaks the device, but is also able to perform the action

to receive a deterministic result.

#### Remark 4

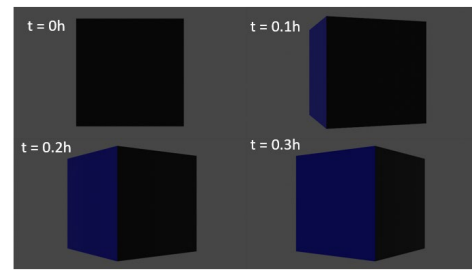
The idea of progressively breaking the material assembly reflects the desire to make material changes in a reversible or continuous manner. The reason for this is that, especially when dealing with intention, uncertainty may arise where it may be desirable from a safety perspective to begin to take actions to disable the device pre-emptively before any harm can be done, but the function of the assembly may also be lost when actions are taken to disable the device in the case harm is present (i.e. the device breaks before harm can be done). This motivates the use of a simulation test bed to fully characterize the control behavior under a wide variety of possible conditions.

#### Remark 5

The choice of having the controller only move one block at some times and one or more at other times was arbitrary and made so that the blocks do not collide after "breaking" off of the assembly. Many different assembly and actuation models of varying complexity could have been considered here depending on the intended function of the material assembly. Instead, this simulation is intended to demonstrate the potential for Blender to be used as part of a testbed framework for developing optical stimuli responsive materials, including those equipped with intent recognition.

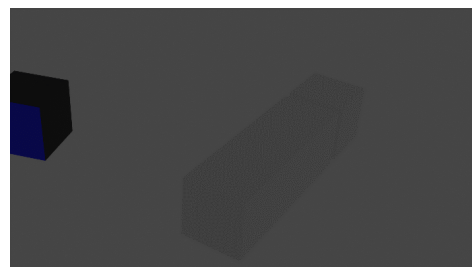
#### Material Assembly with Image Prediction-Based Control

Utilizing the same simulation, we now demonstrate the integration of image predictions in the control law. Specifically, at each sampling time, the controller receives an image of the current system (i.e. an image of the rotating cube from the camera attached to the material assembly). By integrating the dynamics of the rotating cube forward in time, predictions of future images captured by the assembly camera can be generated by Blender and used to preemptively signal control actions to be applied. In this simulation, the dynamics of the system are integrated forward in time to produce image predictions at 0.1, 0.2, and 0.3 h into the future at every sampling time. Using the same algorithm to determine control actions (i.e. analyzing the color of the center pixel of each image), each predicted image is processed. In the current algorithm, the control action signaled by the prediction from 0.3 h in the future is applied at the beginning of the sampling period (this is demonstrated in Figure 7 where the image prediction for  $t = 0.3$  h signals the first block to begin to break off at the beginning of the first sampling period at  $t = 0$  h, where the effect is shown at time  $t = 0.1$  h in Figure 8).



**Figure 7.** At the beginning of the first sampling period at  $t = 0$  h, the controller receives the top left image. The controller integrates the dynamics of the rotating cube forward in time to produce predictions of the image it will see at  $t = 0.1$  h,  $t = 0.2$  h, and  $t = 0.3$  h in the future.

The ability of the controller in this simulation to pre-emptively act based on predicted images demonstrates how the behavior of image responsive systems (in this case, an image responsive material assembly) can be compared with and without image predictions to design a desired response. Further testing may be conducted to tune the desired controller response; for example, consider that the response of the predictive controller in the example above is considered too aggressive in the sense that control actions are being applied based on predictions of images which are too far in the future (i.e. irreversible control actions are being applied based on predictions which we are less confident in). It may be desired to change how the control signals are used, for example using the prediction from 0.1 h in the future (instead of 0.3 h) to signal the controller, or changing how the controller responds to signals (such as breaking each cube off of the assembly at a slower rate) to achieve a desired response. This indicates that Blender may be used as a simulation testbed to design controlled material assemblies and tune aspects of their behavior, including control design and the integration of image predictions, as well as image processing algorithms. These simulations more broadly showcase the ability of Blender to be used to develop simulations which facilitate the testing of integrated image-based and image prediction-based control strategies and image processing algorithms.



**Figure 8.** View of the assembly at  $t = 0.1$  h after the predictive controller signaled a control action at  $t = 0$  h to



break the first cube off the assembly.

## Conclusions

This work provided a perspective on two areas in which process design (and its integration with control) could be impacted, and new avenues opened, by Industry 4.0 considerations related to control system cybersecurity and the use of image-based control and safety systems. We began with a discussion of two ideas of potential avenues for cybersecurity in process design, one which was inspired by the impacts of active attack detection mechanisms on process objectives (i.e., that designs be developed which can facilitate probing for attacks during operation but without impacting overall profits) and one which was inspired by the ability of power supply attacks to use physical measurements on a computing device to backtrack encryption information (i.e., that designs and instrumentation/information availability strategies be analyzed for whether they have any ability to leak information that could reduce security/privacy). We then discussed the potential utility of Blender for testing various considerations related to images in next-generation manufacturing systems design, including in the design of safety monitoring schemes, as well as for advanced assemblies (e.g., stimuli-responsive materials that should respond to certain image data).

## ACKNOWLEDGEMENTS

Financial support from the National Science Foundation CBET-1839675, CNS-1932026, and CBET-2143469, Air Force Office of Scientific Research FA9550-19-1-0059, the National Aeronautics and Space Administration (NASA) under award number 90NSSC30M0124, Michigan Space Grant Consortium (MSGC) and Wayne State University is gratefully acknowledged. We would like to thank Nicholas Deeb for his work on investigating operating system fundamentals and communicating those principles to us.

## REFERENCES

1. Parker S, Wu Z, and Christofides PD. Cybersecurity in process control, operations, and supply chain. *Comput. Chem. Eng.* 108169 (2023).
2. Narasimhan S, El-Farra NH, Ellis MJ. A control-switching approach for cyberattack detection in process systems with minimal false alarms. *AIChE J.* 68:e17875 (2022).
3. Durand H, Wegener M. Mitigating safety concerns and profit/production losses for chemical process control systems under cyberattacks via design/control methods. *Mathematics* 8:499 (2020).
4. Cormier A, Ng C. Integrating cybersecurity in hazard and risk analyses. *Journal of Loss Prevention in the Process Industries* 6: 104044 (2020).
5. Oyama H, Durand H. Integrated cyberattack detection and resilient control strategies using Lyapunov-based economic model predictive control. *AIChE J.* 66:e17084 (2020).
6. Oyama H, Messina D, Rangan KK, Leonard AF, Nieman K, Durand H, Tyrrell K, Hinzman K, Williamson M. Development of directed randomization for discussing a minimal security architecture. *Digital Chemical Engineering* 6:100065 (2023).
7. F. Farokhi, I. Shames, and N. Batterham, "Secure and private control using semi-homomorphic encryption," *Control Engineering Practice*, vol. 67, pp. 13–20, Oct. 2017.
8. A. B. Alexandru, M. Morari, and G. J. Pappas, "Cloud-Based MPC with Encrypted Data," in 2018 IEEE Conference on Decision and Control (CDC). Miami Beach, FL: IEEE, 2018, pp. 5014–5019.
9. MIT OpenCourseWare. 16. Side-Channel Attacks. <https://youtu.be/3v5Von-oNUg?si=b5XuqBQ8dzSUTAjn>
10. Flynn C. Introduction to Side-Channel Power Analysis (SCA, DPA). <https://youtu.be/OIX-p4AGhWs?si=yleaesMyXWP7j4Cb>
11. Kocher, Paul, et al. "Introduction to differential power analysis." *Journal of Cryptographic Engineering* 1 (2011): 5–27.
12. Akkarakaran Francis Leonard, Gjonaj G, Rahman M, and Durand H. Virtual Test Beds for Image-Based Control Simulations Using Blender. *Processes* 12 (2024): 279.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Design Space Identification of the Rotary Tablet Press

Mohammad Shahab<sup>a\*</sup>, Sunidhi Bachawala<sup>b</sup>, Marcial Gonzalez<sup>b</sup>, Gintaras Reklaitis<sup>a</sup>, and Zoltan Nagy<sup>a</sup>

<sup>a</sup> Davidson School of Chemical Engineering, Purdue University, West Lafayette, IN-47907, USA

<sup>b</sup> School of Mechanical Engineering, Purdue University, West Lafayette, IN 47907, USA

\* Corresponding Author: [moshahab@purdue.edu](mailto:moshahab@purdue.edu).

## ABSTRACT

The determination of the design space (DS) in a pharmaceutical process is a crucial aspect of the quality-by-design (QbD) initiative which promotes quality built into the desired product. This is achieved through a deep understanding of how the critical quality attributes (CQAs) and process parameters (CPPs) interact that have been demonstrated to provide quality assurance. For computational inexpensive models, the original process model can be directly deployed to identify the design space. One such crucial process is the Tablet Press (TP), which directly compresses the powder blend into individual units of the final product or adds dry or wet granulation to meet specific formulation needs. In this work, we identify the design space of input variables in a TP such that there is a (probabilistic) guarantee that the tablets meet the quality constraints under a set of operating conditions. A reduced-order model of TP is assigned for this purpose where the effects of lubricants and glidants are used to characterize the design space to achieve the desired tablet CQAs. The probabilistic design space, which takes into account interactions between crucial process parameters and important quality characteristics including model uncertainty, is also approximated because of the high cost associated with the comprehensive experiments.

**Keywords:** design space, tablet press, direct compression, pharmaceutical process, optimization

## INTRODUCTION

The “quality-by-design” (QbD) paradigm put forward by the ICH Q8 guideline on pharmaceutical development states that quality should be built into the products instead of tested into them [1]. The QbD concept allows the practitioner to embrace a thorough and comprehensive approach towards pharmaceutical processes and product development. This leads to a more systematic understanding of the intricate relationships between material attributes, process parameters (CPPs), and product quality (CQAs) for the manufacture of new drugs. The manufacturer may get an advantage from this understanding and receive regulatory clearance to manufacture at any condition and within a broad operating regime if there is sufficient scientific proof that the process will produce a product of acceptable quality. Such an operational regime is called the design space and is also introduced in the ICH guideline. The regulatory approval provides the boundaries within which the material attributes and process parameters can be changed without further approval. However, changes beyond the design space

values mandate a regulatory post-approval change process.

The most widely available literature for design space identification uses an empirical approach and generally follows the steps of: identifying the knowledge space; design-of-experiments (DoE) measuring the relation between the CPPs and CQAs within the knowledge space; using empirical regression methods to define boundaries of the design space; and validation experiments to confirm the design space [3]. The empirical approach is highly favorable when the model is complicated (e.g., multiple unit operations, integrated flowsheet model), and the design space analysis can be challenging due to the computational cost associated with the simulation of the process model.

However, in the presence of a relatively inexpensive computational model, the original model based on the mechanistic equations of the process can be directly used to characterize the design space. Instead of using experimental data to generate empirical relationships, the equations in these models are constructed from a series of presumptions about the physical system and

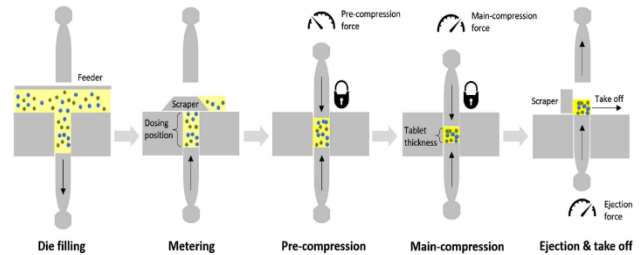
conservation principles of physics and chemistry. The degree to which these assumptions are true determines whether these models hold up when applied to novel situations. There are instances of mechanistic models being applied to the identification of design spaces; however, fewer examples exist for pharmaceutical processes [4-5].

The design space concept has been based on empirical relationships since its beginning; however, the current practice is to evolve by using models that are more mechanistic. The discussion remains about how much experimental evidence is necessary to establish the design space, which changes across regions and application scenarios. There are three broad approaches to identifying design spaces that result in desirable quality attributes: optimization methods, Bayesian inference, and knowledge space sampling. Optimization techniques help identify operational parameters, which ensures that the process acquiesces to a constraint set by performing the flexibility analysis [6]. Bayesian approaches use limited process data and variability and can also incorporate prior knowledge about the process to define a design space that includes uncertainty [7]. Another alternative approach to identifying the design space is to generate a mesh of sample points in the process parameter space and perform simulations at each of those points to determine if the predicted product quality disobeys any constraint [8].

The current work contributes to this fast-evolving domain, where we investigate a steady-state mechanistic tablet press (TP) model [2]. For the purpose of using QbD methods in a direct compression tableting process—where the dry blended component materials are compressed into tablets—the TP model may be used to predict the CPPs and CQAs of tablets. The reduced-order TP model has been used previously to describe the effects of glidants and lubricants on tablet CQAs [2] and implement moving horizon estimation-based non-linear model predictive control (MHE-NMPC) for the tablet press at Purdue’s pharmaceutical continuous manufacturing pilot plant [9]. In the current study, the mechanistic model is deployed for identifying the design space using explicit knowledge space sampling, and the importance of various process parameters is discussed based on the optimal areas of operation that follow a set of constraints. The study investigates the effects of glidants and lubricants and other process parameters on making desired quality tablets via two separate experimental campaigns and their probabilistic design spaces are characterized.

The paper is organized as follows: the next section explains the reduced-order model for the TP, where the role of glidants and lubricants on tablet quality is explained. Subsequently, the design of experiments and the parameter estimation of the mechanistic model are described. We then present the design space results and

recommend the optimal areas of process operation to manufacture tablets of the desired quality. The robust design space is also determined when the process model includes uncertainty by performing Monte Carlo simulations for different probability values as acceptable minimum.



**Figure 1.** Steps in a rotary tablet press process (taken from [10])

## TABLET PRESS MODEL

The tablet press is a multi-stage process that involves the following primary actions at each station: die filling, metering, pre-compression, main-compression, tablet ejection and take-off from lower punch. The metering stage adjusts the dosing position to change the amount of powder within the die after die filling the feed frame. The die is then locked between the upper and lower punches throughout the pre-compression and main-compression stage until tablet ejection and take-off takes place. While the main compression works to compact and solidify the powder into tablets, the pre-compression helps to release trapped air in the die and reorganize the particle packing.

In order to lower frictional losses and improve powder flow during die filling and the mechanical compression-formed solid tablet formation process, lubricants and glidants are essential components. Consequently, the porosity and tensile strength of tablets will be monitored and controlled using models for lubricant/glidant effects in die filling and compression operations. The effects of glidant and lubricant concentrations and mixing conditions are specifically captured by these mechanistic models.

The tablet weight,  $W$ , formed using NATOLI D-type tooling is affected by the process parameters such as dosing position ( $h_{fill}$ ), turret speed ( $n_T$ ), and diameter of the tablet ( $D$ ) and is computed as

$$W = \rho_b V_{fill} \left( -\varphi_1 \frac{n_F}{n_T} + \varphi_2 \frac{h_{fill}}{D} + \varphi_3 \left( \frac{h_{fill}}{D} \right)^2 \right) \quad (1)$$

Where  $\varphi_1$ ,  $\varphi_2$ ,  $\varphi_3$  refer to the model parameters (to be estimated using the experimental data) and  $V_{fill}$ ,  $\rho_b$ , and  $n_F$  is the die cavity volume, powder bulk density, and feed frame speed, respectively. The powder bulk density  $\rho_b$  is dependent on the glidant/lubricant concentration ( $c$ )

and mixing time (or shear imparted during mixing) ( $\gamma$ ) which follows the asymptotic relationship as

$$\rho_b = \rho_{b,\infty} - \frac{\rho_{b,\infty} - \rho_{b,0}}{1 + C_p} \text{ with } C_p = \frac{c_l^{r_1}(\gamma + \gamma_0)^{r_2}}{r_3} \quad (2)$$

where  $\rho_{b,\infty}$ , and  $\rho_{b,0}$  represent the bulk densities when the shear imparted during mixing is infinite and zero respectively,  $\gamma_0$  is the initial shear imparted during pre-mixing, and  $r_1, r_2, r_3$  are model parameters estimated from the experimental data. The bulk density in (2) increases with increase in concentration or mixing time of glidant/lubricant and reaches an asymptotic value. The die cavity volume is calculated as follows

$$V_{fill} = \frac{\pi D^2 h_{fill}}{4} + \frac{\pi h(\frac{3D^2}{4} + h^2)}{6} \quad (3)$$

where  $h$  is the cup depth. The main compression force ( $F_{main}$ ) which is an important process variable can be estimated using the Kawakita equation [11] for the effect of silica (independent of the glidant conditions) by

$$F_{main} = \frac{\pi D^2(\rho_{in-die} - \rho_c)}{4b(\rho_{in-die}(a-1) + \rho_c)} \quad (4)$$

where  $a$  and  $b$  are the Kawakita parameters,  $\rho_c$  is the critical density of the powder and the in-die relative density,  $\rho_{in-die}$  is given by

$$\rho_{in-die} = \frac{W}{\rho_t V_{in-die}} \quad (5)$$

where  $\rho_t$  is the true density of the powder and  $V_{in-die}$  is the die-cavity volume with main compression thickness  $h_{in-die}$  given by

$$V_{in-die} = \frac{\pi D^2 h_{in-die}}{4} + \frac{\pi h(\frac{3D^2}{4} + h^2)}{3} \quad (6)$$

The lubrication conditions are found to affect the compression force for the MgSt blends and this is incorporated by modifying  $a$  as

$$a = a_\infty - \frac{a_0 - a_\infty}{1 + C_c} \text{ with } C_c = \frac{c_l^{p_1}(\gamma + \gamma_0)^{p_2}}{p_3} \quad (7)$$

where  $a_0, a_\infty, p_1, p_2,$  and  $p_3$  are model parameters. Here, the compaction force increases with increasing lubrication. The elastic recovery ( $\varepsilon_p$ ) model which is part of the tablet ejection stage is insensitive to the glidant mixing conditions and can be calculated by

$$\varepsilon_p = \varepsilon_0 \frac{\rho_{in-die} - \rho_{c,\varepsilon}}{1 - \rho_{c,\varepsilon}} \quad (8)$$

However, the lubricant conditions affect the elastic recovery and an increase in the former increases the latter. This behavior is captured by modeling  $\varepsilon_0$  as

$$\varepsilon_0 = \varepsilon_\infty - \frac{\varepsilon_\varphi - \varepsilon_\infty}{1 + C_\varepsilon} \text{ with } C_\varepsilon = \frac{c_l^{q_1}(\gamma + \gamma_0)^{q_2}}{q_3} \quad (9)$$

where  $\varepsilon_\varphi, \varepsilon_\infty, q_1, q_2,$  and  $q_3$  are model parameters. The tablet density  $\rho_{tab}$  can then be calculated using

$$\rho_{tab} = \rho_{in-die} (1 - \varepsilon_p) \quad (10)$$

The tensile strength ( $\sigma_t$ ) of a tablet is a crucial CQA which affects the tablet dissolution along with the tablet weight  $W$ . Both lubricant/glidant concentration and mixing time affect the tensile strength and the relationship is governed by Kuentz and Luenberger, 2000 [12]

$$\sigma_t = \sigma_0 \left[ 1 - e^{(\rho_{tab} - \rho_{c,\sigma})} \left( \frac{1 - \rho_{tab}}{1 - \rho_{c,\sigma}} \right) \right] \text{ with } \sigma_0 = \frac{\sigma_{0,\varphi}}{1 + C_\sigma}$$

$$\rho_{c,\sigma} = \rho_{c,\sigma,\infty} - \frac{\rho_{c,\sigma,\infty} - \rho_{c,\sigma,\varphi}}{1 + C_\sigma} \text{ and } C_\sigma = \frac{c_l^{b_1}(\gamma + \gamma_0)^{b_2}}{b_3} \quad (11)$$

Where  $\rho_{c,\sigma,\varphi}, \rho_{c,\sigma,\infty}, b_1, b_2, b_3$  are the model parameters [13] and  $\sigma_{0,\varphi}$  and  $\rho_{c,\sigma,\varphi}$  represents the tensile strength and critical relative density when there is no lubrication,  $C_\sigma = 0$ . As a result of the tensile strength model, soft tablets with lower tensile strength are formed as the concentration or mixing time of glidant or lubricant in the formulation increases. However, the decrease in tablet tensile strength manufactured using silica blends is solely because of the variations in the blended material properties, but the lower tensile strength of tablets formed using Magnesium Stearate (MgSt) blends would in addition be due to the increased elastic recovery of lubricated tablets.

## METHODOLOGY

Tablets are the most common oral solid dosage form that can be produced by direct compression or enhanced by either wet or dry granulation to meet specific formulation needs. The direct compression line in the pharmaceutical continuous manufacturing pilot plant at Purdue University was used for the studies in this work. The materials used in the current study include 10% w/w acetaminophen (APAP) as the API, microcrystalline cellulose Avicel PH200 (MCC) as the excipient, and glidant colloidal silica and lubricant Magnesium Stearate (MgSt) at different concentrations. Colloidal silica is an excipient which is useful for improving powder flowability and MgSt helps reduce internal friction during compaction and tablet-tooling friction during ejection. These excipients have a substantial impact on the powder's surface characteristics, such as the strength of the solid bridges created during compaction, as well as bulk properties, including bulk density, even when added in very small amounts [14]. This shift in the blend's characteristics naturally affects the tableting procedure and the final tablet's CQAs, including its dissolving profile [13, 15], which in turn affects the active pharmaceutical product's (API) bioavailability.

For the experimental campaigns, APAP and MCC were mixed in a tote blender with 0-0.2% w/w silica in the first experimental campaign and with 0-2% w/w MgSt in the second campaign for 10-30 minutes. The dosing

position and the in-die (main compression) thickness values were maintained between 7–11 mm (9–13 for lubricant MgSt) and at 3.1 mm, respectively, to manufacture tablets that have a wide range of relative densities. The design-of-experiments DoE was conducted separately for the two experimental campaigns to study the effects of silica and MgSt on tablet quality. The MATLAB function lhsdesign was used to build a Latin hypercube sampling (Viana, 2013) of turret speed, dosing position, concentration, and mixing time to generate 30 experiments for the silica blends and 20 experiments for the MgSt blends. Next, the powder blends were compressed into tablets using a NATOLI-NP400 tablet press using D-type tooling with shallow cup of depth 0.33 mm, which features a total of 22 punch-die stations with die-size 8 mm. A SOTAX AT4 tablet tester was utilized to measure the weight, hardness, diameter, and thickness of 50 tablets under steady-state manufacturing conditions for each run in the DoE. At the beginning of each experimental run, 0.5 kg of the mix was added to the tablet press hopper.

The TP model parameters are estimated by minimizing the least squares which is typically solved as an optimization problem where the objective is to minimize the sum of squared errors (*SSE*) between the model predicted and observed values.

$$SSE = \sum_{i=1}^n (y_{mod,i} - y_{obs,i})^2 \quad (12)$$

where  $y_{mod,i}$  and  $y_{obs,i}$  are the  $i$ th estimated value using the TP model and  $i$ th experimental data of the tablet CQAs (tablet weight, main compression force, and tablet tensile strength), respectively, and  $n$  is the number of experimental samples. The sum of squared errors is minimized, and the corresponding optimal values of model parameters and their uncertainty are stored for design space characterization.

There are 4 CPPs in the DoE namely turret speed, dosing position, concentration, and mixing time. In the tablet press, the feed frame speed gets automatically adjusted to be ~15 rpm greater than the turret speed. While high feeder to turret speed ratios can lead to uneven die-filling and tablet weight variability, the current DoE maintains low turret and feeder speeds, minimizing their impact on tablet weight. The dosing position ( $h_{fill}$ ) determines the tablet weight, while the main compression thickness ( $h_{in-die}$ ) affects the tablet density and hence the tensile strength of the tablet. Therefore, the crucial CPPs for the design space are the dosing position, main compression thickness, concentration, and mixing time and their realizable bounds are shown in Table 1. The main compression thickness is flexible enough to vary within operable limits and can be used to widen the useful design space regions. The tablet CQAs considered in this study include main compression force, tablet weight, and tensile strength, and the desired specifications are mentioned in Table 2. The compression force is essentially

not included in the tablet CQAs but it helps in regulating the lower bounds of dosing position and main compression thickness that results in useful design space regions. Therefore, it is included as a quality constraint. The objective of this study is to identify the design space of input variables ( $h_{fill}$ ,  $h_{in-die}$ ,  $c_L$ ,  $\gamma$ ) such that the tablets are guaranteed to meet the quality constraints ( $W$ ,  $F_{main}$ ,  $\sigma_t$ ) under the operating limits of the tablet press.

**Table 1:** The critical process parameters (CPPs) and their bounds

CPP	Low limit	High limit
Dosing position ( $h_{fill}$ )	6 mm	11 mm
Comp. thickness ( $h_{in-die}$ )	1.5 mm	3 mm
Silica conc. ( $c_i$ )	0.1 %	0.2 %
MgSt conc. ( $c_i$ )	0.1 %	2 %
Mixing time ( $\gamma$ )	10 mins.	30 mins.

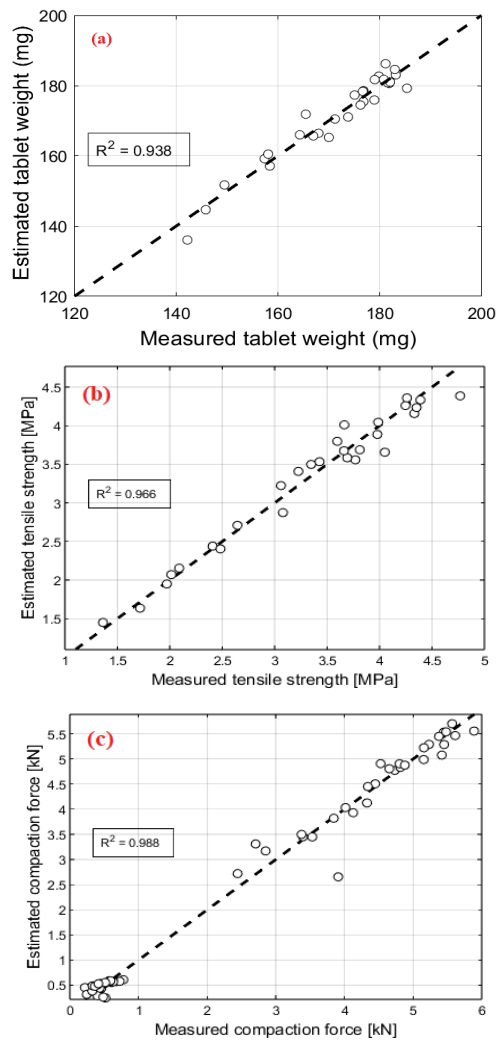
**Table 2:** The critical quality attributes (CQAs) and their specifications

CPP	Low	High
Tablet weight ( $W$ )	125 mg	175 mg
Compression Force ( $F_{main}$ )	2 kN	50 kN
Tensile strength ( $\sigma_t$ )	2 MPa	15 MPa

As previously stated, an explicit sampling of the knowledge space is used to identify the design space. For the deterministic design space, the process parameters are discretized by creating a fine mesh of sample points within their predefined bounds. To determine if the projected CQAs jointly observe the constraints for each discretized sample, simulations are performed for each of these discretized CPPs. A deterministic design space is the outcome of this mapping.

For the probabilistic design space, additionally, at each of the discretized CPPs, Monte Carlo simulations are executed on the full model  $N$  times using the model parameter values in the space of their uncertainty bounds. The uncertainty information is incorporated in the design space using the computed standard deviation values of model parameters from the least squares approach. The fitted model predictions against experimental data for different CQAs can be seen in Figure 2. Then, for each of the discretized process parameter samples, the probability of meeting the constraints jointly is estimated. The probability is computed based on the fraction of times the Monte Carlo simulations resulted in CQA values that complied with all the constraints. For each discretized sample, the probability is set equal to this fraction. This allows for the propagation of model uncertainty to the predicted CQAs, and the Monte Carlo simulation serves well in showing the possible values of the model prediction as explicit probability maps which

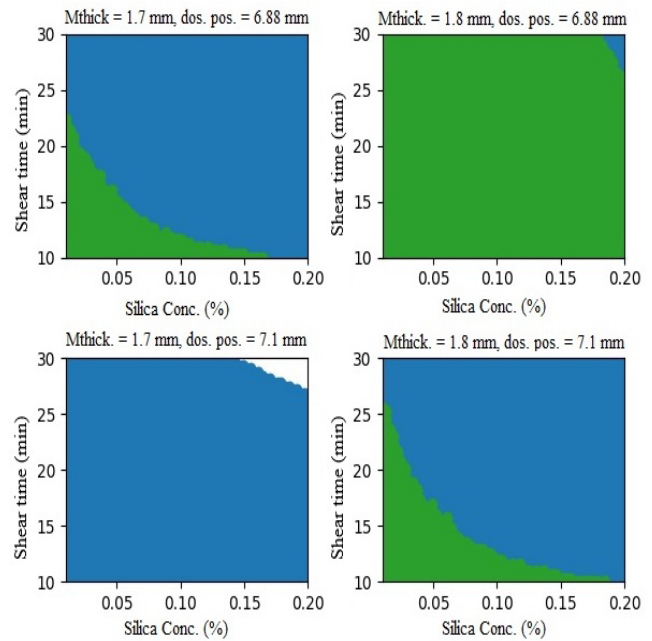
accounts for model variability and common cause variability.



**Figure 2.** Fitted TP model predictions versus experimental data for (a) tablet weight, (b) tensile strength, (c) compaction force.

The explicit sampling method is effective in identifying the probabilistic design space, and the process parameter grid can be made finer through discretization, and there is no restriction on the sampling size of the uncertain parameters. It is worth mentioning that a major limitation of this approach lies in the large number of simulations that need to be performed, and the design of effective strategies for knowledge space sampling is an active research area but that is beyond the scope of this work. Despite the computational challenges, the explicit sampling method is the most straightforward approach to estimate even complex design spaces. However, the goal here is to recommend optimal areas of operation for the tablet press within the process parameter space that results in quality tablets utilizing the straightforward and

effective explicit sampling approach.



**Figure 3.** Design space plots for silica blends following CQA constraints mentioned in Table 2. Blue region is the deterministic design space (without accounting for model uncertainty) and green region is the probabilistic design space with 85% confidence.

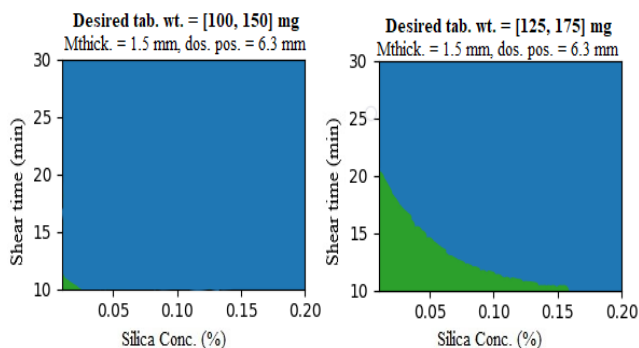
## RESULTS AND DISCUSSIONS

From the tablet press model described in the second section, it can be seen that the CPPs dosing position  $h_{fill}$ , lubricant or glidant concentration  $c_L$ , and mixing time  $\gamma$ , affect the CQA tablet weight  $W$  (from equations 1, 2, and 3). The main compression force  $F_{main}$ , which is another CQA, is additionally influenced by the main compression thickness (equations 4, 5, 6, and 7). The tensile strength  $\sigma_t$  is also dependent on all the CPPs (equations 8, 9, 10, and 11). Therefore, the tablet press model can be lumped into three equations: one for tablet weight, a second for main compression force, and a third for tensile strength. The equations have not been mentioned here due to their comprehensiveness. However, the equations can be used to solve for the design space variables within the bounds shown in Table 1, and the equations themselves would be constrained due to the limits specified in Table 2.

Figure (2) depict the multidimensional design space in simplified graphical form for the silica blends. Several simulations were performed for various combinations of process parameters and sampling model parameters within their uncertainty bounds in order to identify the multidimensional design space. For demonstration, 1000 Monte Carlo simulation realizations were used in this case, with an acceptable minimum probability of 85%.



The blue region is where the three quality constraints are met without accounting for the model uncertainty and the green region is where the probability of meeting all the quality constraints is acceptable. It is important to emphasize here that the lower feasible bounds of the dosing position and the compression thickness is regulated by the main compression force and their upper feasible bounds is controlled by the tensile strength constraint. From the design space plots of silica blends, it is observed that the closer you operate near the nominal value of compression thickness (~ 2.1 mm) and dosing position (~6.9 mm), wider would be the probabilistic design space. Although the comprehensive probability maps have not been shown here, but the variation in design space can be explained as follows: for lower dosing positions values (closer to lower bound in Table 1), lower thickness values are favorable and for high dosing positions values (closer to upper bound), higher thickness values are required for tablet quality to obey the constraints. This is only valid within the feasible limit of process parameters. The design space plots for the lubricated blends were found to very narrow and small even



for the deterministic case and are therefore not reported.

**Figure 4.** Design space plots for different tablet weight CQA limits. Blue region is the deterministic design space (without accounting for model uncertainty) and green region is the probabilistic design space with 85% confidence.

We also demonstrate the effectiveness of the TP model in capturing the design space boundaries by showing the design space plots with varying tablet weight limits. The other two CQA constraints remain unchanged as mentioned in Table 2. To illustrate the case, we have considered a new limit for tablet weight  $W = [100, 150]$  mg and compare it with the original case but the operating conditions on dosing position and compression thickness is kept fixed at (6.3 mm, 1.5 mm) and other constraints also remain unchanged. The silica blends are used for this case study. The change in feasible region boundary for the probabilistic design space is shown in Figure 4. Here the green region shrinks towards bottom left for smaller tablet weight CQA.

## CONCLUSIONS

The study investigated the determination of design space for the tablet press operation to achieve desired tablet quality. The design space plots showed the effects of glidant or lubricant on the tablet quality. Tablets made from silica blends showed much promise where a design space point in a sufficient broad feasible region can be used to perform validation experiments. The efficacy of the TP model is also depicted in the case where the variation in probabilistic design space boundary is captured when the required tablet weight changes. An immediate future work includes performing the validation experiments based on the identified design space plots.

The future work includes various improvisation to the current study. First, the TP model needs to be tested against variations in measured tablet CQAs. Second, to develop design space plots for much smaller or larger tablets and to study the effect of die size on such tablets and how the TP model performs with such variations. Finally, an empirical approach such as Bayesian modeling based on the DoE can be used to explore the design space of the TP and compare its efficacy against the model-based methodology.

## ACKNOWLEDGEMENTS

The authors acknowledge support by NSF (CMMI-EPSRC: RIFTMaP, 2140452).

## REFERENCES

1. ICH. Quality Guideline Q8(R2) Pharmaceutical Development. (2009)
2. Bachawala S., Gonzalez M. Development of mechanistic reduced order models (ROMs) for glidant and lubricant effects in continuous manufacturing of pharmaceutical solid-dosage forms. *Comput. Aided Chem. Eng.* 51:1129-1134 (2022)
3. Huang, J., Kaul, G., Cai, C., Chatlapalli, R., Hernandez-Abad, P., Ghosh, K. and Nagi, A., Quality by design case study: an integrated multivariate approach to drug product and process development. *Int. journal of pharm.*, 382(1-2), pp.23-32. (2009)
4. Brueggemeier, S.B., Reiff, E.A., Lyngberg, O.K., Hobson, L.A. and Tabora, J.E., Modeling-based approach towards quality by design for the ibipinabant API step. *Org. process res. & dev.*, 16(4), pp.567-576. (2012)
5. Prpich, A., am Ende, M.T., Katzschner, T., Lubczyk, V., Weyhers, H. and Bernhard, G., Drug product modeling predictions for scale-up of tablet film coating—a quality by design approach. *Comput &*

- Chem. Eng., 34(7), pp.1092-1097. (2010)
6. Grossmann, I.E. and Morari, M., Operability, resiliency, and flexibility: Process design objectives for a changing world. (1983)
  7. Chatzizacharia, K.A. and Hatzivramidis, D.T., Design space approach for pharmaceutical tablet development. *Ind. & Eng. Chem. Res.*, 53(30), pp.12003-12009. (2014)
  8. Garcia-Munoz, S., Luciani, C.V., Vaidyaraman, S. and Seibert, K.D., Definition of design spaces using mechanistic models and geometric projections of probability maps. *Org. process Res. & Dev.*, 19(8), pp.1012-1023. (2015)
  9. Huang, Y.S., Sheriff, M.Z., Bachawala, S., Gonzalez, M., Nagy, Z.K. and Reklaitis, G.V., Evaluation of a combined mhe-nmpc approach to handle plant-model mismatch in a rotary tablet press. *Processes*, 9(9), p.1612. (2021)
  10. Su, Q., Bommireddy, Y., Shah, Y., Ganesh, S., Moreno, M., Liu, J., Gonzalez, M., Yazdanpanah, N., O'Connor, T., Reklaitis, G.V. and Nagy, Z.K., Data reconciliation in the Quality-by-Design (QbD) implementation of pharmaceutical continuous tablet manufacturing. *Int. journal of pharm.*, 563, pp.259-272. (2019)
  11. Kawakita, K. and Lüdde, K.H., Some considerations on powder compression equations. *Powder technology*, 4(2), pp.61-68. (1971)
  12. Kuentz, M. and Leuenberger, H., A new theoretical approach to tablet strength of a binary mixture consisting of a well and a poorly compactable substance. *Eur. journal of pharm. and biopharma.*, 49(2), pp.151-159. (2000)
  13. Razavi, S.M., Gonzalez, M. and Cuitiño, A.M., Quantification of lubrication and particle size distribution effects on tensile strength and stiffness of tablets. *Powder Technology*, 336, pp.360-374. (2018)
  14. Mehrotra, A., Llusa, M., Faqih, A., Levin, M. and Muzzio, F.J., Influence of shear intensity and total shear on properties of blends and tablets of lactose and cellulose lubricated with magnesium stearate. *Int. journal of pharm.*, 336(2), pp.284-291. (2007)
  15. Van Veen, B., Bolhuis, G.K., Wu, Y.S., Zuurman, K. and Frijlink, H.W., 2005. Compaction mechanism and tablet strength of unlubricated and lubricated (silicified) microcrystalline cellulose. *Eur. journal of Pharm. and Biopharm.*, 59(1), pp.133-138. (2005)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Recent Advances of PyROS: A Pyomo Solver for Nonconvex Two-Stage Robust Optimization in Process Systems Engineering

Jason A. F. Sherman<sup>a</sup>, Natalie M. Isenberg<sup>a, b</sup>, John D. Sirola<sup>c</sup> and Chrysanthos E. Gounaris<sup>a\*</sup>

<sup>a</sup> Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA

<sup>b</sup> Brookhaven National Laboratory, Upton, NY

<sup>c</sup> Sandia National Laboratories, Albuquerque, NM

\* Corresponding Author: [gounaris@cmu.edu](mailto:gounaris@cmu.edu).

---

## ABSTRACT

In this work, we present recent algorithmic and implementation advances of the nonconvex two-stage robust optimization solver PyROS. Our advances include extensions of the scope of PyROS to models with uncertain variable bounds, improvements to the formulations and/or initializations of the various subproblems used by the underlying cutting set algorithm, and extensions to the pre-implemented uncertainty set interfaces. The effectiveness of PyROS is demonstrated through the results of an original benchmarking study on a library of over 8,500 small-scale instances, with variations in the nonlinearities, degree-of-freedom partitioning, uncertainty sets, and polynomial decision rule approximations. To demonstrate the utility of PyROS for large-scale process models, we present the results of a carbon capture case study. Overall, our results highlight the effectiveness of PyROS for obtaining robust solutions to optimization problems with uncertain equality constraints.

---

**Keywords:** Design Under Uncertainty, Pyomo, Algorithms, Optimization, Process Design

## INTRODUCTION

Two-stage robust optimization (RO) is a useful framework for obtaining robust designs of process and energy systems. Optimization models for process and energy systems are often subject to parametric uncertainty, arising from the use of empirical correlations for property models, temporal variations in feedstock quality, or economic stochasticity [1]. Since any changes in the prevailing values of these uncertain parameters may significantly affect the model outputs, deterministic optimization may yield system designs which are suboptimal or infeasible under off-nominal scenarios. Moreover, the design of a plant should account for any built-in controllability that will allow the operator to adjust plant response during operation.

The recently developed two-stage RO solver PyROS [2, 3] is designed to obtain robust solutions to process models. A typical process design model features nonconvexities, design and recourse degrees of freedom, and a prevalence of equality constraints constituting the

“simulation part” of the model. The equality constraints are often highly complex, to the extent that they cannot, in general, be reformulated out of the model. Until recently, there were no existing RO tools capable of solving nonconvex two-stage RO problems with general nonlinear inequality constraints. Based on a generalization [1] of the robust cutting set algorithm of [4], PyROS is the first two-stage RO solver that systematically handles the equality constraints without reformulation requirements [3].

In this work, we present an overview of the cutting set methodology upon which the PyROS solver is based, recent algorithmic and implementation advances of the PyROS solver, a high-throughput benchmarking study to demonstrate the effectiveness of PyROS for a wide range of small-scale nonconvex two-stage RO instances, and a post combustion carbon capture case study to show the utility of PyROS for obtaining risk-averse designs for critical systems with large-scale nonconvex optimization models.

## METHODOLOGY OVERVIEW

### Formulations

Let  $x \in R^{n_x}, z \in R^{n_z}, y \in R^{n_y}, q \in R^{n_q}$  denote the first-stage variables, second-stage variables, state variables, and uncertain parameters, respectively, for a model of interest. The first-stage variables may be assigned a value from a known domain  $\mathcal{X} \subseteq R^{n_x}$ , described by nonlinear constraints involving  $x$  only. We assume the uncertain parameters  $q$  are restricted to an *a priori* known, compact, exogenous uncertainty set  $Q \subset R^{n_q}$ , containing a nominal value  $q^{\text{nom}}$ . For fixed  $x$  and  $q$ , the second-stage variables  $z$  can be adjusted freely, and the state variables  $y$  are then implicitly determined from the model constraints.

In a deterministic optimization setting, the model of interest is a nonlinear program (NLP) of the form

$$\begin{aligned} \min_{\substack{x \in \mathcal{X} \\ z \in R^{n_z} \\ y \in R^{n_y}}} \quad & f_1(x) + f_2(x, z, y; q^{\text{nom}}) \\ \text{s.t.} \quad & g_i(x, z, y; q^{\text{nom}}) \leq 0 \quad \forall i \in \mathcal{J} \\ & h_j(x, z, y; q^{\text{nom}}) = 0 \quad \forall j \in \mathcal{J} \end{aligned} \quad (1)$$

in which  $f_1: \mathcal{X} \rightarrow R$  denotes the first-stage objective function,  $f_2: \mathcal{X} \times R^{n_z} \times R^{n_y} \times Q$  denotes the second-stage objective function,  $g_i: \mathcal{X} \times R^{n_z} \times R^{n_y} \times Q, i \in \mathcal{J}$  denotes the inequality constraint functions, and  $h_j: \mathcal{X} \times R^{n_z} \times R^{n_y} \times Q, j \in \mathcal{J}$  denotes the equality constraint functions. The two-stage RO counterpart of the model (1) is

$$\begin{aligned} \min_{x \in \mathcal{X}} \quad & \max_{q \in Q} \quad \min_{\substack{z \in R^{n_z} \\ y \in R^{n_y}}} f_1(x) + f_2(x, z, y, q) \\ \text{s.t.} \quad & g_i(x, z, y, q) \leq 0 \quad \forall i \in \mathcal{J} \\ & h_j(x, z, y, q) = 0 \quad \forall j \in \mathcal{J} \end{aligned} \quad (2)$$

### Solution Approach: Generalized Robust Cutting Set Algorithm

The two-stage RO problem (2) can be solved using the generalized robust cutting set algorithm presented in [1, 3]. We now present a brief overview. First, the adjustability of the second-stage variables with respect to the uncertain parameters is approximated with decision rules. For each  $l = 1, 2, \dots, n_z$ , let  $d_l \in R^{n_{d,l}}$  be a vector of *decision rule coefficients*, each of which is considered an additional first-stage variable. We model the second-stage variable  $z_l$  as a scalar function  $v_l: R^{n_{d,l}} \times Q \rightarrow R$  of the coefficients  $d_l$  and the uncertain parameters  $q$ . In this case, the RO problem (2) becomes

$$\begin{aligned} \min_{\substack{x \in \mathcal{X} \\ d \in R^{n_d}}} \quad & \max_{\substack{q \in Q \\ z \in R^{n_z} \\ y \in R^{n_y}}} \quad \min_{\zeta \in R} \zeta \\ \text{s.t.} \quad & \zeta \geq f_1(x) + f_2(x, z, y, q) \\ & g_i(x, z, y, q) \leq 0 \quad \forall i \in \mathcal{J} \\ \text{s.t.} \quad & h_j(x, z, y, q) = 0 \quad \forall j \in \mathcal{J} \\ & z_l = v_l(d_l, q) \quad \forall l \in \mathcal{L} \end{aligned} \quad (3)$$

in which  $d = (d_1, d_2, \dots, d_{n_z})$  and  $\mathcal{L} = \{1, 2, \dots, n_z\}$ . We

assume that for fixed  $x \in \mathcal{X}, z \in R^{n_z}, q \in Q$ , the state variables  $y \in R^{n_y}$  are *uniquely* determined by the equality constraints  $h_j(x, z, y, q) = 0 \forall j \in \mathcal{J}$ .

Observe that (3) constitutes a restriction of (2) through the addition of the decision rule equations. Given (3), we may then obtain a feasible first-stage solution  $(x, d)$  with the following iterative approach. For each iteration  $k = 0, 1, \dots$ :

- A sampled relaxation of (3), termed the *master problem*, is solved to obtain a candidate solution  $(x^*, d^*)$ .
- For each  $i \in \mathcal{J}$ , a *separation problem* is solved to determine, if it exists, a parameter realization  $q \in Q$  for which there is no feasible adjustment of  $z, y$  satisfying the inequality constraint  $g_i(x^*, z, y, q) \leq 0$ , subject to the model's equality constraints and decision rule policy (with  $d_l = d_l^*$  fixed for every  $l \in \mathcal{L}$ ). A similar separation problem is solved for the epigraph constraint.
- If a realization for which infeasibility is detected is found in the previous step, then the constraints subject to this realization are added to the master problem, and we proceed with iteration  $k + 1$ . Otherwise,  $x^*$  is said to be *robust feasible*, and the algorithm terminates. If  $x^*$  is robust feasible, and the master problem has been solved to global optimality, then  $x^*$  is said to be *robust optimal*.

To reduce the risk of numerical issues, we are careful to initialize the variables of the master problem of each iteration  $k = 1, 2, \dots$  to an optimal solution of a closely related slack variable-based feasibility problem.

In [1], a procedure for empirically evaluating the quality of the robust feasible or robust optimal solution returned by the cutting-set algorithm is provided.

## PYROS SOLVER INTERFACE

### Implementation

The Pyomo Robust Optimization Solver (PyROS) is a Python-based implementation of the cutting set algorithm presented in the previous section; the official documentation is contained in [2]. Built on top of the open-source algebraic modeling language Pyomo, PyROS has been designed for ease of use, such that a user can quickly extend a deterministic optimization workflow to a two-stage RO workflow. Successful invocation of the PyROS two-stage RO solver requires only the following arguments:

- The deterministic model
- The first-stage degree-of-freedom variables
- The second-stage degree-of-freedom variables

- The uncertain model parameters
- The uncertainty set
- Subordinate NLP optimizers. At least one optimizer must be capable of solving nonconvex NLPs to global optimality.

For a Pyomo user with a modest level of experience, the uncertainty set is the only argument for which instantiation may immediately pose a challenge. Therefore, to facilitate the construction of uncertainty set objects, PyROS provides an abstract uncertainty set class, and a suite of pre-implemented concrete subclasses, based on uncertainty quantifications often used in the RO literature. The pre-implemented classes include hyperrectangular, polyhedral, ellipsoidal, and discrete uncertainty set types; full descriptions of all available set types are provided in [2, 3]. If none of the pre-implemented classes meet the user's modeling needs, then the user may implement a custom-written subclass of the abstract uncertainty set class.

## Recent Updates

Below, we synopsise a number of implementation features that were contributed to the PyROS codebase in recent months.

### Improved Subproblem Initializations

Given that the master and separation subproblems of the underlying cutting set algorithm may, in general, be nonconvex NLPs, the success of solving these subproblems may be heavily contingent on the initial point provided to the subordinate NLP optimizers. To this end, we have opted to initialize the master problems with the solution to a slack-based feasibility problem formulation, such as to obtain a feasible, or near-feasible, initial point. To reduce the level of infeasibility of the initial point for the feasibility problem, the variables of the feasibility problem of iteration  $k$  are initialized based on the solution to the master and separation problems of the previous iteration  $k - 1$ . Similarly, each of the separation subproblems of iteration  $k$  is initialized based on the solution to the master problem of iteration  $k$ .

### Acceleration of the Separation Step

The separation problem subroutines have been refactored to enhance the performance of those subroutines, and by extension, the PyROS solver, and facilitate improvements to the initializations of subproblems in subsequent iterations. Moreover, based on the assumption that the equality constraints of (3) uniquely determine the state variables  $y$  for fixed  $x \in \mathcal{X}, z \in R^{n_z}, q \in \mathcal{Q}$ , we have implemented an efficiency to significantly reduce the computational expense of the separation step for problems with discrete uncertainty sets.

### Extensions to the Uncertainty Set Interfaces

The budget and factor model set classes of the original PyROS release have been generalized.

### Updated Solver Log Output

PyROS now features a solver output logging system which provides a breakdown of the model component statistics, a summary of the progress of the underlying cutting set algorithm at the end of each iteration, and a summary of the solver termination statistics.

## BENCHMARKING STUDY

### Test Set and Computational Environment

To evaluate the performance of the PyROS solver, we sought to establish a comprehensive library of small-scale benchmark instances upon which a high-throughput computational study could be performed. Ten deterministic NLP models of varying size and nonlinearity were collected from the open deterministic NLP modeling libraries [5, 6]. From this collection of deterministic NLPs, a library of 8,591 two-stage RO problem instances was assembled by varying the degree-of-freedom partitioning, subset of model parameters considered uncertain, uncertainty set type and parameterization, and decision rule order specification.

All two-stage model instances were instantiated and solved in a Python 3.9 virtual environment with Pyomo 6.6.1dev0/PyROS 1.2.7. Instances were solved on a single core of an Intel® Xeon® CPU E5-2687W v3 @ 3.10 GHz, running Ubuntu Linux 20.04 with 64 GB RAM. For the subordinate local and global NLP optimizers, we used IPOPT 3.14.6/MA27 [7] and BARON v23.6.23/CPLEX 22.1.0 [8], respectively. Instances were solved to a robust feasibility tolerance of  $1 \times 10^{-3}$  and wall time limit of 400 s. For the `hydrothermal` and `optmass` benchmark instances, the master problems were solved to local optimality; otherwise, all subproblems were solved to global optimality. Additional details about the instances and computational environment used in this study are presented in [3].

### Results

Table 1 summarizes the results of our computational study. Of the 8,591 instances, 88.9% were solved to the desired robust optimality or feasibility target. The average wall time and iteration requirements for successfully solved instances were found to be low. The relatively high wall time and iteration requirements for `himmel6`, `haverly`, `hydro`, and `hydrothermal` instances were found to be attributable to the size and/or nonlinearities present in the original deterministic models, which rendered the master and/or separation subproblems more difficult to solve. For instances in the model library that were not successfully solved, the outcome was most often found to be attributable to a pathological separation



subproblem which could not be solved successfully by the subordinate global optimizer due to the wall time requirements exceeding the specified limit of 400s. Regardless, the success rate of 88.9% and low iteration and wall time requirements suggest that PyROS is able to obtain robust feasible or robust optimal solutions to non-convex two-stage robust NLPs with a high level of reliability.

**Table 1:** Computational performance statistics for high-throughput computational study of the PyROS solver. Averages are reported for the successfully solved instances only.

Base Model	No. of Instances		Average	
	Total	Successful	Wall $t$ (s)	Iterations
haverly	923	520	9.3	3.6
himmelp6	497	446	9.1	2.4
hydro	923	661	14.9	4.2
hydro-thermal	923	846	24.0	3.1
lewispol	923	923	1.2	1.0
optcntrl	923	870	3.7	2.2
optmass	923	862	4.3	2.7
s353	710	703	1.1	1.7
s381	923	922	1.8	2.6
s382	923	882	1.9	2.4

## CASE STUDY: AMINE-BASED POST COMBUSTION CARBON CAPTURE

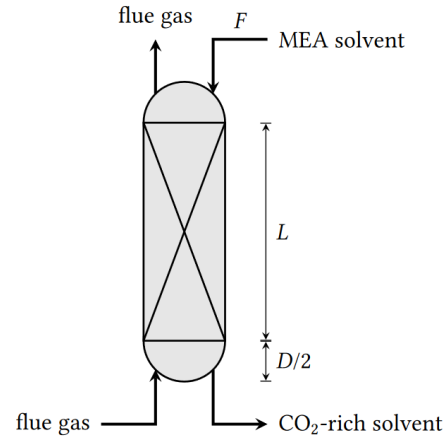
### Introduction

Carbon capture and storage technologies are considered vital to minimal cost pathways for reducing the carbon footprints of process systems to meet global climate targets [9, 10]. Along with other organizations worldwide, the United States Department of Energy has spearheaded initiatives for developing computational tools to accelerate the design and deployment of carbon capture and storage systems [11, 12]. One issue with computational model-based tools is the susceptibility of the underlying mathematical optimization models to parametric uncertainty. Towards developing a computational framework for risk-averse carbon capture process designs, models for amine-based CO<sub>2</sub> absorption processes have been identified as benchmarks, given the relative maturity of amine-based absorption technologies [11, 12]. In this section, we present a case-study based on a model of a CO<sub>2</sub> absorption column using monoethanolamine (MEA) solvent.

### Modeling Framework

Our case study involves a high-fidelity, one-dimensional, steady state, rate-based model for an MEA-based CO<sub>2</sub> absorption column. Figure 1 shows a schematic

diagram. The column length and diameter are denoted by  $L$  and  $D$ , respectively, and we assume the column heads are hemispherical. Flue gas, rich in CO<sub>2</sub>, enters the base of the column from an upstream process. A stream of CO<sub>2</sub>-lean MEA solvent enters the top of the column at a molar flow rate  $F$  to remove CO<sub>2</sub> from the gaseous mixture prior to emission.



**Figure 1:** Schematic for the MEA-based CO<sub>2</sub> absorption column model. The column length and diameter are denoted by  $L$  and  $D$ , respectively. The lean solvent flow rate is denoted by  $F$ .

The axial dimension of the column is discretized into a grid of  $N = 40$  finite elements. For each  $i = 1, 2, \dots, N$ , we denote the flooding fraction in the  $i$ th finite element by  $u_i$ . We are free to adjust the column dimensions  $L, D$  during design of the system (i.e. as first-stage variables) and the solvent flow rate  $F$  during operation (i.e. as a second-stage variable). The length-to-diameter ratio of the column must fall between 1.2 and 30. We require that the CO<sub>2</sub> capture rate of the column meet a pre-specified target, and that the flooding fractions  $u_1, u_2, \dots, u_N$  fall within pre-specified bounds.

Using the notation of the previous paragraphs, the deterministic optimization model can be written as

$$\begin{aligned}
 \min_{L, D, F, u, y} \quad & \frac{1}{4}\pi D^2 L + \frac{1}{3}\pi D^3 + \frac{F}{2} \\
 \text{s.t.} \quad & h(L, D, F, u, y; q) = 0 \\
 & 1.2D \leq L \leq 30D \\
 & r^L - f^{\text{CO}_2}(F, y) \leq 0 \\
 & 0.5e \leq u \leq 0.8e
 \end{aligned} \tag{4}$$

in which  $u = (u_1, u_2, \dots, u_N)$  is a vector whose components are the finite element flooding fractions;  $y$  is a vector of approximately 5,000 state variables;  $h$  denotes a nonlinear vector of approximately 5000 +  $N$  equality constraint functions, governed by material and energy balances and transport and thermodynamic property models;  $q$  is a vector of six uncertain model parameters, discussed later in the text;  $r^L$  is a constant denoting the CO<sub>2</sub> capture

**Table 2:** Deterministic optimization results for MEA-based CO<sub>2</sub> absorption column model.

Capture Target (%)	Capture Rate (%)	$L$ (m)	$D$ (m)	Volume ( $10^3$ m <sup>3</sup> )	$L/D$ Ratio	Flow rate $F$ (kmol/s)	Optimal Cost
85.0	85.0	18.19	15.19	6.93	1.2	13.24	13.55
87.5	87.5	18.24	15.20	6.99	1.2	13.63	13.81
90.0	90.0	18.30	15.25	7.05	1.2	14.02	14.07
92.5	92.5	18.35	15.29	7.12	1.2	14.41	14.32
95.0	95.0	18.40	15.34	7.18	1.2	14.81	14.58
96.0	96.0	18.42	15.35	7.20	1.2	14.96	14.68
96.5	96.5	18.43	15.36	7.21	1.2	15.04	14.73
97.0	97.0	18.44	15.37	7.22	1.2	15.12	14.79
97.5	97.5	18.45	15.38	7.24	1.2	15.21	14.84

**Table 3:** PyROS-based robust optimization results for MEA-based CO<sub>2</sub> absorption column model.

Minimum Capture Rate (%)	Robust Column Proxy Cost and DOF ( $L, D, F$ ) Values [m, m, kmol/s] (Expected $\pm$ Standard Deviation)			
	0% confidence (deterministic)	90% confidence	95% confidence	99% confidence
90.0	14.02 $\pm$ 0.08 (18.30, 15.25, 13.93 $\pm$ 0.17)	14.28 $\pm$ 0.11 (18.46, 15.38, 14.07 $\pm$ 0.22)	14.29 $\pm$ 0.11 (18.47, 15.39, 14.02 $\pm$ 0.22)	14.31 $\pm$ 0.12 (18.49, 15.41, 14.08 $\pm$ 0.25)
92.5	14.28 $\pm$ 0.09 (18.35, 15.29, 14.32 $\pm$ 0.17)	14.50 $\pm$ 0.10 (18.49, 15.41, 14.41 $\pm$ 0.20)	14.51 $\pm$ 0.10 (18.50, 15.42, 14.41 $\pm$ 0.20)	14.55 $\pm$ 0.13 (18.52, 15.43, 14.41 $\pm$ 0.26)
95.0	14.53 $\pm$ 0.09 (18.40, 15.34, 14.81 $\pm$ 0.18)	14.74 $\pm$ 0.11 (18.52, 15.44, 14.81 $\pm$ 0.23)	14.76 $\pm$ 0.11 (18.54, 15.45, 14.85 $\pm$ 0.23)	robust infeasible

target, and  $f^{\text{CO}_2}$  is a scalar function which evaluates the CO<sub>2</sub> capture rate. The nonlinearity of the component functions of  $h$  renders the deterministic model (4) non-convex.

We reiterate that  $L, D, F$  are degree-of-freedom variables, such that the flooding fractions  $u$  and other state variables  $y$  are implicitly determined by the equality constraints of (4). The objective of (4) is a proxy for the total capital and operating cost of the column. In particular, the capital cost is represented by the column volume  $(1/4)\pi D^2 L + (1/3)\pi D^3$ , while the operating cost is represented by the solvent flow rate, scaled by a factor of  $1/2$ , so that both costs share the same order of magnitude.

The uncertain parameters  $q$  consist of two liquid phase chemical reaction equilibrium constants and four vapor-liquid equilibrium constants, downselected according to the results of an uncertainty propagation study

of the deterministic optimization model and associated thermodynamic and physical property submodels. The resulting six-dimensional uncertainty set takes the form of a confidence ellipsoid centered on the nominal uncertain parameter realization, with a pre-specified confidence level and covariance matrix.

### Deterministic Optimization Results

Subject to the nominal realization of the uncertain parameters  $q$ , the deterministic absorption column model was solved to local optimality with GAMS 45.0/CONOPT3 [13] for capture targets  $r^L$  ranging from 85.0% to 97.5%. Table 2 shows the results. In all cases, the capture target constraint of (4) was found to be active, and the  $L/D$  ratio was found to achieve the lower bound of 1.2. Observe that the optimal column volume, flow rate, and objective were all found to increase as the capture target was

increased. We also note that, based on the model thermodynamics, a theoretical upper bound for the CO<sub>2</sub> capture rate is 98.2%, and we were unable to obtain nominally feasible designs for capture targets above 98.1%.

## Robust Optimization Results

Through rigorous sensitivity analysis studies, we were able to establish that the nominally optimal column designs are robust infeasible subject to quantifications of uncertainty in the VLE and reaction equilibrium parameters. To this end, the two-stage robust counterpart of the deterministic model (4) was solved with PyROS, subject to the degree of freedom partitioning and confidence ellipsoidal uncertainty quantification discussed previously, and a static decision rule approximation. We have elected to optimize for each of three different CO<sub>2</sub> capture targets and three levels of ellipsoidal confidence.

Table 3 summarizes the RO results for the MEA-based absorption column model. For all capture targets, the robust feasible column designs were found to be significantly more expensive than their deterministic counterparts. Nevertheless, the robust optimal designs for a given capture target were found to be comparable in cost to nominally feasible designs for higher capture targets. Observe, for example, that the robust feasible cost of  $14.31 \pm 0.12$  for 90% capture under a 99% confidence ellipsoidal uncertainty, is significantly lower than the cost  $14.53 \pm 0.09$  of the deterministic design for a 95% capture target. For the case of a 95% capture target and 99% confidence level, the robustness requirements were too stringent for the resulting problem to admit a risk-averse design. We expect that incorporating additional recourse, through the usage of a full CO<sub>2</sub> capture flowsheet model in lieu of the standalone absorption column model used in the present study, will afford robust designs for capture targets and confidence level requirements as high as 95% and 99%, respectively. It is worth noting that the difference between the robust feasible cost and the deterministically optimal cost provides an upper bound for the value worth investing in uncertainty reduction through additional data acquisition. Overall, our results highlight the ability of the PyROS solver to obtain economical, risk-averse designs of an amine-based post combustion carbon capture system with a high-fidelity NLP model.

## CONCLUSIONS

We have presented recent advances in the non-convex two-stage RO solver PyROS. The interface is designed to facilitate extensions of deterministic optimization models to two-stage RO workflows. Our computational benchmarking study demonstrates the effectiveness of PyROS for identifying robust feasible or robust optimal solutions to general nonlinear two-stage RO problems for a broad range of uncertainty

quantifications. Finally, our post combustion carbon capture case study shows the applicability of PyROS to the robust design of process systems. In future work, we intend to perform additional improvements to the initializations of the various subproblems, perform RO studies of the MEA-based absorption column model subject to non-static decision rule approximations, and extend the post combustion carbon capture case study to a full flowsheet optimization model.

## ACKNOWLEDGEMENTS

This work was conducted as part of the U.S. Department of Energy's Institute for the Design of Advanced Energy Systems (IDAES) and Carbon Capture Simulation for Industry Impact (CCSI2) initiative supported by the Office of Fossil Energy and Carbon Management, through the Simulation-based Engineering/Crosscutting Research Program and Carbon Capture Program. Natalie M. Isenberg also acknowledges support from a U.S. Department of Energy, Office of Science Graduate Student Research award.

The authors would like to acknowledge the contributions of Anca G. Ostace, Douglas A. Allan, Andrew Lee and Miguel A. Zamarripa of the National Energy Technology Laboratory to modeling the amine-based carbon capture system serving as the case study in this paper.

## DISCLAIMER

This project was funded by the Department of Energy, National Energy Technology Laboratory, an agency of the United States Government, through a support contract. Neither the United States Government nor any agency thereof, nor any of its employees, nor the support contractor, nor any of their employees, makes any warranty, expressor implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

This article has been authored by an employee of National Technology & Engineering Solutions of Sandia, LLC under Contract No. DE-NA0003525 with the U.S. Department of Energy (DOE). The employee owns all right, title and interest in and to the article and is solely responsible for its contents. The United States Government retains and the publisher, by accepting the article

for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this article or allow others to do so, for United States Government purposes. The DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan <https://www.energy.gov/downloads/doe-public-access-plan>.

## REFERENCES

1. Isenberg, N. M., Akula, P., Eslick, J. C., Bhattacharyya, D., Miller, D. C., & Gounaris, C. E. (2021). A generalized cutting-set approach for nonlinear robust optimization in process systems engineering. *AIChE Journal*, *67*(5), e17175.
2. Isenberg, NM, Sherman, JA, Sirola, JD, Gounaris, CE. *PyROS Solver*. Pyomo Official Documentation. Retrieved November 21, 2023, from [https://pyomo.readthedocs.io/en/stable/contributed\\_packages/pyros.html](https://pyomo.readthedocs.io/en/stable/contributed_packages/pyros.html).
3. Isenberg, NM, Sherman, JA, Sirola, JD, & Gounaris, CE. *PyROS: The Pyomo Robust Optimization Solver*. Forthcoming. 2024.
4. Mutapcic, A., & Boyd, S. (2009). Cutting-set methods for robust convex optimization with pessimizing oracles. *Optimization Methods & Software*, *24*(3), 381-406.
5. GAMS Development Corporation. *gamsworld: GAMS World model library*. Retrieved November 21, 2023, from <https://github.com/GAMS-dev/gamsworld>.
6. GAMS Corporation. *The GAMS Model Library*. Retrieved November 21, 2023, from [https://www.gams.com/latest/gamslib\\_ml/libhtml/index.html](https://www.gams.com/latest/gamslib_ml/libhtml/index.html).
7. Wächter, A., & Biegler, L. T. (2006). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical programming*, *106*, 25-57.
8. Tawarmalani, M., & Sahinidis, N. V. (2005). A polyhedral branch-and-cut approach to global optimization. *Mathematical programming*, *103*(2), 225-249.
9. Intergovernmental Panel on Climate Change. *IPCC Special Report on Renewable Energy Sources and Climate Change Mitigation*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press; 2012.
10. Bui M, Adjiman CS, Bardow A, Anthony EJ, Boston A, Brown S, Fennell PS, Fuss S, Galindo A, Hackett LA, Hallett JP. Carbon capture and storage (CCS): the way forward. *Energy & Environmental Science*. 2018;11(5):1062-176.
11. Morgan JC, Bhattacharyya D, Tong C, Miller DC. Uncertainty quantification of property models: Methodology and its application to CO<sub>2</sub>-loaded aqueous MEA solutions. *AIChE Journal*. 2015 Jun;61(6):1822-39.
12. Akula P, Lee A, Eslick J, Bhattacharyya D, Miller DC. A modified electrolyte non-random two-liquid model with analytical expression for excess enthalpy: Application to the MEA-H<sub>2</sub>O-CO<sub>2</sub> system. *AIChE Journal*. 2023 Jan 1:e17935.
13. Drud, A. S. (1994). CONOPT—a large-scale GRG code. *ORSA Journal on computing*, *6*(2), 207-216.

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Optimal Design Approaches for Cost-Effective Manufacturing and Deployment of Chemical Process Families with Economies of Numbers

Georgia Stinchfield<sup>a</sup>, Sherzoy Jan<sup>a</sup>, Joshua C. Morgan<sup>b,c</sup>, Miguel Zamarripa<sup>b,c</sup>, and Carl D. Laird<sup>a\*</sup>

<sup>a</sup> Carnegie Mellon University, Department of Chemical Engineering, Pittsburgh, PA 15213, USA

<sup>b</sup> National Energy Technology Laboratory (NETL), Pittsburgh, PA 15236, USA

<sup>c</sup> NETL Support Contractor, Pittsburgh, PA 15236, USA

\* Corresponding Author: [claird@andrew.cmu.edu](mailto:claird@andrew.cmu.edu)

## ABSTRACT

Developing methods for rapid, large-scale deployment of carbon capture systems is critical for meeting climate change goals. Optimization-based decisions can be employed at the design and manufacturing phases to minimize the costs of deployment and operation. Manufacturing standardization results in significant cost savings due to economies of numbers. Building on previous work, we present a process family design approach to design a set of carbon capture systems while explicitly including economies of numbers savings within the formulation. Our formulation optimizes both the number and characteristics of the common components in the platform and simultaneously designs the resulting set of carbon capture systems. Savings from economies of numbers are explicitly included in the formulation to determine the number of components in the platform. We show and discuss the savings we gain from economies of numbers.

**Keywords:** Optimization, Process Design, Energy Systems, Carbon Capture

## INTRODUCTION

Effectively combatting climate change relies on large-scale deployment of critical chemical process systems, such as carbon capture or water desalination. In traditional process system design approaches, engineers uniquely design each installation focused on economies of scale. However, this approach is expensive and leads to long deployment timelines. Modular design approaches derive savings from economies of numbers by offering a catalog of small, stackable designs. However, a pure modular approach neglects the benefits of economies of scale.

In this paper, we develop a rigorous, optimization-based design method, inspired by product family design literature [1], that designs a family of process variants while simultaneously optimizing a platform of unit module designs that can be shared across this set of processes. This approach seeks to exploit both economies of scale and economies of numbers to minimize costs while achieving manufacturing standardization and reduced deployment timelines. While well-developed

technologies can benefit from this approach, we apply it to novel process systems like carbon capture due to the importance of making this technology widely available within a short time period.

This work builds on our optimization formulation for process family design [2] and extends it to explicitly include the benefits of economies of numbers. Economies of numbers (sometimes referred to as economies of learning) is a well-documented cost-saving phenomenon [3,4]. It characterizes the manufacturing cost savings due to standardization; in particular, it captures the correlation between cost reduction and the number of times a particular product has been manufactured. Following an approach like that described by Gazzaneo et al. (2022), we develop a costing expression that captures total manufacturing costs as a function of the number of unit modules produced.

If the platform has a small number of unit module designs, there are fewer options to share across all the process variants. Therefore, we will be manufacturing more of each design and gaining increased benefits from economies of numbers. However, increasing the number



of unit module designs in the platform gives each process variant more options leading to a more “optimized” design (at the cost of reducing savings due to economies of numbers). With the optimization formulation in Stinchfield et al. (2023), the number of unit module designs included in the platform must be pre-specified. In this work, by including the economies of numbers explicitly, we allow the mathematical programming formulation to determine the optimal number of unit module designs to include in the platform. We demonstrate this approach on multiple case studies, including an MEA-based carbon capture system and a water desalination process.

## LITERATURE REVIEW

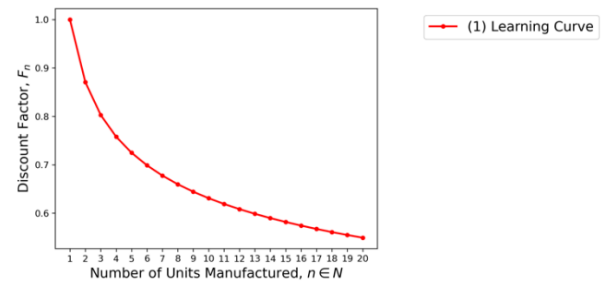
Modular and product family design approaches have been studied in many different industrial applications. Gonzales-Zugasti (2000) describes how a product family approach can be applied to NASA’s exploratory space missions beginning with a two-stage optimization approach [7]. Simpson et. al. (2004) used a genetic algorithm for product family design optimization for the selection of parts of aircraft [14]. Pirmoradi et al. (2015) devised a two-phase platform configuration approach utilizing sensitivity analysis, metamodeling, and a black-box optimization strategy to craft a variety of universal electric motor designs [9]. In general, research focused on product family design focuses on heuristic and stochastic optimization techniques, rather than employing rigorous deterministic mathematical programming algorithms.

Product family design derives significant savings from the standardization of elements within a product. Manufacturing cost savings associated with standardization are due to *economies of numbers* (sometimes referred to as *economies of learning*). Increasing the number of products to be produced results in lower per unit costs, as documented by Wright et al. (1936). Specifically, the authors reported how the per unit cost of manufacturing an airplane *decreased* with respect to the number of airplanes manufactured [10]. This correlation has since been documented across many industries with different levels of cost reduction.

The effect of economies of numbers can be captured mathematically as a function of the number of units manufactured,  $n$ , and their resulting discount factor,  $F_n$ . The learning rate  $\alpha$  captures the impact of production levels on cost, as shown in (1).

$$F_n = n^{-\alpha} \quad (1)$$

Equation (1) is also referred to as the learning curve and is shown graphically in Figure 1, with an  $\alpha = 0.2$ .



**Figure 1.** Example Learning Curve<sup>[13]</sup>  
( $\alpha = 0.2$ )

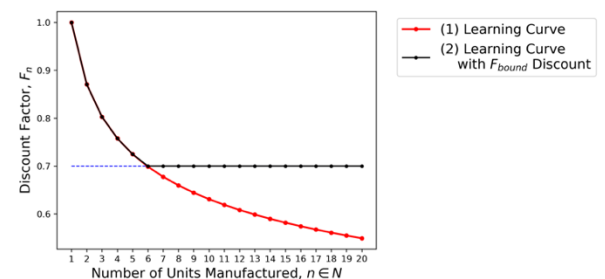
Learning curves are product specific. They exhibit the general behavior shown in Figure 1 but vary significantly depending on the specific industry and unit manufactured. In practice,  $\alpha$  is selected based on experience or data. Argot and Epple (1990) explored factors affecting this parameter, including organizational forgetting, turnover, and transfer of productivity gains, and how they can potentially reduce costs [11].

Economies of numbers has been mentioned occasionally, but favorably, in literature related to chemical and process systems engineering. Liebermann (1984) investigated this concept and used a many-shot approach to try and determine which factors affected the learning rates in a plant or chemical manufacturing context [3]. Weber et al. (2019) described how economies of numbers apply to the chemical process industries [12]. Gazzaneo et al. (2022) proposed a novel techno-economic framework for costing intensified modular systems in the process engineering industry [4].

This cost benefit from manufacturing standardization applies only to part of the overall process cost (e.g., labor but not materials). Therefore, we expect these learning curves to decay toward an asymptote. Gazzaneo et al. (2022) define a piecewise function that includes a lower limit on the discount factor, as shown in (2).

$$F_n = \begin{cases} n^{-\alpha}, & \text{if } n^{-\alpha} \geq F_{bound} \\ F_{bound}, & \text{if } n^{-\alpha} < F_{bound} \end{cases} \quad (2)$$

This modified relationship is shown in Figure 2 with  $\alpha = 0.2$  and  $F_{bound} = 0.7$ .



**Figure 2.** Example Learning Curve with Maximum Discount<sup>[4]</sup>  
( $\alpha = 0.2$  and  $F_{bound} = 0.70$ )

In this work, we use a similar economies of numbers correlation that also includes a maximum discount,  $F_{bound}$ , with a smooth transition. Our work then incorporates this correlation into a mixed-integer linear programming formulation for the simultaneous design of process families.

## PROBLEM APPROACH

In previous work, we have presented a discretization-based Mixed Integer Linear Program (MILP) to solve the process family design problem [2]. In this section, we present modifications to this formulation that include cost trade-offs associated with economies of numbers and avoid explicitly specifying the size of the platform. To find a comprehensive description of the problem, please see Stinchfield et al. (2023).

### Problem Description

Given a process architecture, we wish to specify the designs for a set of *process variants*. A process variant  $v$  requires customization of the process architecture to meet a set of specific requirements; the specifications are parameterized in the vector  $\mathbf{r}_v$ . For example, take an industrial refrigeration system that a grocery store chain wishes to deploy at 10 different stores. While the general refrigeration system architecture will be the same at each location (i.e., the units required to build the system), the design details for each store (e.g., the size and maximum cooling capacity) can be different.

The *process system architecture* refers to the flow-sheet of the process system; this defines all the *unit module types* necessary to build an instance of the process. In the case of the refrigeration system example, unit module types could include the compressor, condenser, valve, and evaporator. We store this set of unit module types in the set  $M$ . Each variant  $v$  will have exactly one of each unit module type  $m \in M$ . The variable vector  $\mathbf{d}_{v,m}$  is the *unit module type design* for each unit module type  $m$  and each variant  $v$ .

Our goal is to design and deploy multiple variants  $v \in V$  in a cost-optimal manner by determining the unit module type designs  $\mathbf{d}_{v,m}$  and operating variables  $\mathbf{o}_v$ . We wish to save on engineering and manufacturing costs by optimizing a platform of common unit module designs to share across the process variants. This means there are fewer unique units to design and reduced manufacturing costs for the shared designs due to economies of numbers. We separate the unit module types  $M$  into those that are designed *commonly* (stored in the set  $C$  and included in the platform) and the remaining ones that are designed *uniquely* for each variant (stored in the set  $U$ ). The sets  $C$  and  $U$  are disjoint ( $C \cap U = \emptyset$ ) while their union recovers the set  $M$  ( $C \cup U = M$ ). The design specifications for the shared platform unit module types are captured by the corresponding variable vectors  $\mathbf{d}_{c,l}$ . We index the

common designs by a label  $l \in L_c$  to differentiate between options (e.g., different sizes of a compressor). The collection of common designs  $\mathbf{d}_{c,l}$  for all common unit module types  $c \in C$  determines the platform  $\mathcal{P}$ . All unique unit module types  $u \in U$  are designed specifically for each variant; they do not require any standardized elements.

### Problem Formulation

To determine an optimal process system design for a single variant with a set of design requirements, a traditional approach would start by building the set of equations that defines the system (i.e., physics, costing, etc.). An optimization could then be performed, parameterized with the requirements, where the designs for all unit module types  $m \in M$  and operational decisions are decision variables. In most applications, this system of equations will be a nonlinear program (NLP). In our approach, we must design *multiple* systems, one for each variant  $v \in V$ . Additionally, the unit module designs included in the platform are optimized simultaneously. And, for each variant, the common unit module design must be selected from those in the platform. This introduces discrete decisions and leads to a Mixed-Integer Nonlinear Program (MINLP).

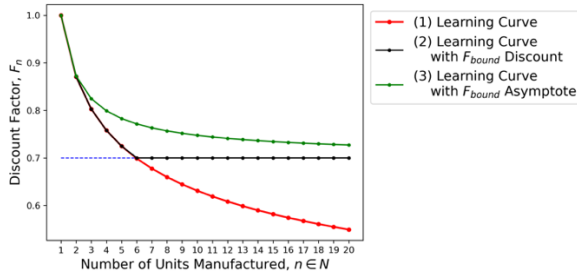
Successful optimization of MINLPs in process systems engineering can be a challenge; oftentimes, it is an active area of research. In the process family design setting, it can quickly become impractical to directly solve the MINLP with the entire set of design, costing, and performance equations in the overall formulation. Our approach to this challenge was to develop a Mixed-Integer Linear Programming (MILP) formulation that relies on a discretized set of *candidate* designs for the common unit module types  $c \in C$ . For each possible combination of candidate common unit module designs and each process variant  $v \in V$ , we optimize the process system for the unique unit module designs, operating variables, and cost. We call this a design *alternative* for variant  $v$ . From here, we build a set of feasible and infeasible design combinations and their associated costs. These form the input data for the MILP formulation.

We select a combination of common unit module designs to assign to a particular variant using the binary decision variable  $x_{v,a}$ . Here,  $a$  refers to a particular design *alternative*, and all feasible alternatives for a variant  $v$  are stored in the set  $A_v$ . The binary decision variables  $y_{c,l,n}$  determine the designs of common unit module types to include in the platform and the number that are manufactured. The parameter  $M_c$  sets the upper limit on the number of designs of unit module type  $c$  to include in the platform. Of course, the problem is constrained to ensure we do not select alternative  $a$  for a variant  $v$  unless we have also selected the corresponding unit module designs in the platform.

As described in the previous section, we aim to capture cost savings associated with economies of numbers within our formulation. Like Gazzaneo et al. (2022), we use a function that approaches  $F_{bound}$  as  $N \rightarrow \infty$ , and the smooth formulation is shown in (3).

$$F_n = F_{bound} + (1 - F_{bound}) \times n^{-\beta} \quad (3)$$

This function captures the cost reduction due to economies of numbers while accounting for fixed costs (e.g., materials) that do not decrease with increasing  $n$ . The parameter  $\beta$  also represents a learning rate. However, since this is only applied to a portion of the overall manufacturing costs, we use different values from  $\alpha$  to capture similar behavior to that in Gazzaneo et al. (2022), as shown in Figure 3.



**Figure 3.** Learning Curve with Asymptotic Approach towards Max. Discount ( $\alpha = 0.2$ ,  $F_{bound} = 0.70$ ,  $\beta = 0.8$ )

We pre-compute each possible discounted unit module cost based on the number of times the unit module type could be manufactured. The base cost of each unit module type  $c$  design  $l$  is stored in  $p_{c,l}$ . We store the discounted costs in  $\widehat{p}_{c,l}^n$ , where each entry represents the cost of unit module type  $c$  for common design  $l$  if it is manufactured  $n$  times. We introduce the binary variable  $y_{c,l,n} \in \{0,1\}$  to indicate the number of manufactured units of unit module type  $c$  and common design  $l$ .

Formulation (4) describes the MILP we used to design a process family from a set of candidate common designs with discounts from economies of numbers.

$$\min_{x,y,\rho} \sum_{v \in V} w_v \sum_{a \in A_v} x_{v,a} c_{v,a} - \rho \quad (4a)$$

$$\text{s. t.} \quad \sum_{a \in A_v} x_{v,a} = 1 \quad \forall v \in V \quad (4b)$$

$$\sum_{l \in L_c} (1 - y_{c,l,0}) \leq M_c \quad \forall c \in C \quad (4c)$$

$$x_{v,a} \leq 1 - y_{c,l,0} \quad \forall c \in C, l \in L_c, (c, l) \in Q_a$$

$$\sum_{n=0}^N y_{c,l,n} = 1 \quad \forall c \in C, l \in L_c$$

$$\sum_{n=0}^N n \times y_{c,l,n} = \sum_{v \in V} \sum_{a \in A_{v,c,l}} w_v x_{v,a}$$

$$\rho = \sum_{c \in C} \sum_{l \in L_c} \sum_{n=0}^N n \times y_{c,l,n} \times (p_{c,l} - \widehat{p}_{c,l}^n)$$

$$0 \leq x_{v,a} \leq 1$$

$$y_{c,l,n} \in \{0,1\}$$

$$\rho \in \mathbb{R}^1$$

This formulation generally performs two main functions, (1) selecting which candidate design should be included within the process platform and (2) which common designs should be used at each variant, all while considering the effects of economies of numbers. The objective (4a) minimizes the total weighted cost of all variants, including the total savings from economies of numbers, contained in the variable  $\rho$ . Constraint (4b) ensures only one alternative (i.e., combination of common unit module designs) is selected for each variant. Constraint (4c) sets an upper limit of  $M_c$  on the number of common designs selected to be in the platform. Constraint (4d) ensures an alternative can only be selected if we also choose to select the required unit module designs for the platform. (4e) and (4f) constrain the binary indicator  $y_{c,l,n}$  to be 1 if design  $l$  for unit module type  $c$  has been selected to be manufactured  $n$ -times and 0 otherwise. (4g) calculates the total cost savings attributed to unit module manufacturing standardization for the given process family and stores the entire discount in the variable  $\rho$ . (4h) – (4j) defines the domain for the three optimization variables. Notably, (4g) defines  $x_{v,a}$  to be a continuous variable between the bounds of 0 and 1. For all case studies,  $x_{v,a}$  converges to binary decision, most likely due to similarities in the formulation to the  $P$ -Median optimization problem [13].

## CASE STUDY

In this section, we describe the case study used to demonstrate our optimization approach. We chose a monoethanolamine solvent-based carbon capture system simulated in Aspen Plus® as a part of the CCSi2 initiative [6]. The CCSi2 initiative focused on developing computational tools and models for accelerating the commercialization of carbon capture technologies. The flowsheet described by Morgan et al. (2022) is shown in Figure 4.

Solvent-based processes, particularly amines, are a mature class of technology for CO<sub>2</sub> capture from point sources of power generation and industrial processes. The solvent-based CO<sub>2</sub> capture process is shown schematically in Figure 4.

(4d)  
(4e)

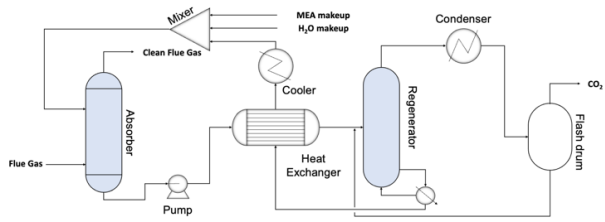


Figure 4. MEA Carbon Capture Flowsheet

As shown in Figure 4, the point source flue gas containing CO<sub>2</sub> enters the bottom of the absorber and is contacted countercurrently with solvent flowing down the column. The CO<sub>2</sub>-lean solvent stream enters at the top of the column. The transfer of CO<sub>2</sub> from the gas to the liquid occurs through reactive absorption, and the mass transfer area is generally provided by structured packing. Since the reaction of CO<sub>2</sub> with amine is exothermic, solvent intercooling is often included in the process design to expel the heat of absorption associated with the reaction. At one or more locations in the column, a portion of the solvent is extracted and cooled by cooling water before being returned to the column. This generally results in lowering the temperature profile in the absorber and thus increasing the driving force for CO<sub>2</sub> uptake in the liquid. The flue gas with reduced CO<sub>2</sub> content exits the top of the absorber, and the CO<sub>2</sub>-rich solvent exits the bottom. The rich solvent is pressurized to avoid flashing at higher temperatures required for solvent regeneration and is heated in the lean/rich heat exchanger by the lean solvent exiting the bottom of the stripper. In the stripper the CO<sub>2</sub> is separated from the solvent with the energy requirement for the endothermic reaction provided by the steam input to the reboiler. The stream exiting the top of the stripper primarily contains CO<sub>2</sub> and H<sub>2</sub>O, the latter of which is condensed and returned to the stripper as reflux. This results in a high purity stream of CO<sub>2</sub>, which is compressed and sent for sequestration or utilization. The lean solvent exits the bottom of the stripper and is cooled by the rich solvent in the lean/rich heat exchanger. The trim cooler, which uses cooling water, provides the residual duty required to cool the solvent prior to its return to the top of the absorber column.

This model has been adapted in this work for a process family design problem that accommodates a variety of flue gas feed conditions including a range of flowrates and CO<sub>2</sub> concentrations representing different industrial flue gas sources. We define the set of variants  $v \in V$  to be different combinations of flue gas flow rates and flue gas CO<sub>2</sub> concentrations. This was motivated by the fact that these quantities vary significantly by carbon capture application, and the design decisions depend heavily on these two quantities but differ significantly at each potential carbon capture location, depending on the

capacities of the plants and the source of CO<sub>2</sub>. For this case study, we consider seven flue gas flow rates and nine CO<sub>2</sub> concentrations.

The set of unit module types for this system contains all of those types defined in the flowsheet, which is to say  $M = [\text{absorber, pump, heat exchanger, regenerator, condenser, flash drum, cooler, mixer}]$ . For the commonly designed absorber and regenerator (identified by the blue units shown graphically in Figure 4), we design for a specific *volume* of each column. To run simulations in Aspen Plus®, absorber and regenerator designs were specified using reported parameters from NCCC [6] as a baseline. We tested six absorber diameters (0.5m, 0.6m, ..., 1.0m) and eight regenerator diameters (0.3m, 0.4m, ..., 1.0m). Furthermore, to optimize the lean loading, we considered five different CO<sub>2</sub> lean loading concentrations (0.16, 0.17, 0.18, 0.19, 0.20) and selected the best lean loading for each combination of design for the absorber, regenerator, and process variant based on lowest total annualized cost. Given the 63 different process variants we wish to design, six absorber designs, eight regenerator designs, and five CO<sub>2</sub> lean loadings considered in this case study required 15,120 simulations. After running each simulation, the results were used to calculate the total annualized cost for each design alternative.

## RESULTS & DISCUSSIONS

In this section, we present the results from the case study described in the previous section by employing the methodology described in the Problem Approach section. We discuss key findings from the results.

We designed the process family using the optimization formulation described in (4). We do not include (4c), instead allowing the formulation to select however many of the candidate unit module designs to include in the platform. We built this optimization formulation in Python using the open-source algebraic modeling language, Pyomo [5]. The optimization resulted in the selection of four common designs for the absorber (out of six possible) and four common designs for the regenerator (out of eight possible), as shown in Figure 5.

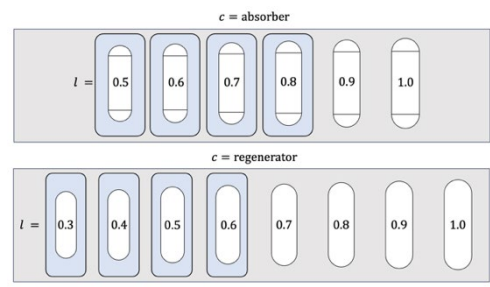
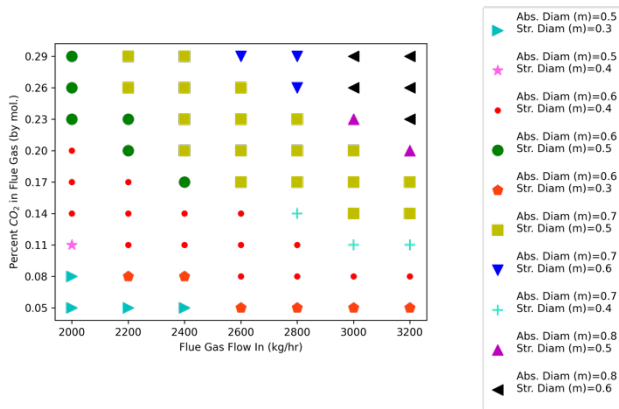


Figure 5. Carbon Capture Platform,  $\mathcal{P}$



The common designs selected for the platform are identified by the blue shaded boxes. To differentiate between the designs,  $l$  corresponds to the diameter (in  $m$ .) of the design.

The design of the platform  $\mathcal{P}$  is determined simultaneously with the design of the process family  $\mathcal{F}$  in the optimization formulation. From the process platform constructed and shown in Figure 5, the corresponding optimal design of the process family is shown in Figure 6.



**Figure 6.** Carbon Capture Process Family,  $\mathcal{F}$

Figure 6 describes which combinations of absorber and regenerator designs offered in the platform  $\mathcal{P}$  (shown in Figure 5) are assigned to each variant  $v \in V$ . The x-axis corresponds to the flue gas flow rate that describes a particular variant  $v$ ; the y-axis corresponds to the percentage of  $\text{CO}_2$  in the flue gas at variant  $v$ .

Using Gurobi<sup>®</sup>, it took under a second to solve this problem. The formulation presented in (4) for this case study resulted in 1,376 constraints, 646 continuous variables, and 896 binary variables. The objective of this optimization resulted in a total annualized cost, discounted by savings due to economies of numbers, of \$72.5M. We used a learning rate of  $\beta = 0.8$  and a maximum discount factor of  $F_{bound} = 0.7$ . Savings associated with economies of numbers, captured in the value for the variable  $\rho$ , came to approximately \$2.38M annually. The percentage of savings associated with the overall cost is approximately 3.3%. If we only consider the capital costs for this system, the percent savings is 26.8%.

To compare this approach to a more traditional method, we optimized each carbon capture variant independently. To do this for the case study presented, we selected the combination of candidate absorber and regenerator designs that minimized the cost of the variant, with no discounts due to economies of numbers or limitations on the number of common designs that could be used. In this way, we design each variant *individually* rather than as a family. The overall objective cost of this individual optimization came to \$74.86M, which is more than \$2M more expensive than taking the

process family design approach. This demonstrates the importance of considering economies of numbers in the design of chemical process systems due to the benefits it can provide.

## CONCLUSIONS & FUTURE WORK

Process family design lends insight into how standardization within the process systems engineering manufacturing industry can potentially save money. In particular, the optimization approach described in this paper aims to capture and fully exploit cost savings by quantifying the impact of economies of numbers on the design of a process family. We demonstrated this approach in a carbon capture case study motivated by the need to deploy these systems rapidly and cost-effectively to combat the effects of climate change. The results of this case study showed a decrease in total annualized cost compared to taking the discretized approach described in Stinchfield et al. (2023) and a traditional engineering approach. The optimization formulation presented in this work selected the size of the platform, rather than having to pre-define the size, which further allowed the formulation to determine the optimal trade-off between standardization and customization of the system. We are investigating decomposition strategies for the MINLP formulations from Stinchfield et al. (2023).

## ACKNOWLEDGEMENTS & DISCLAIMER

This effort was partially funded by the U.S. Department of Energy's Institute for the Design of Advanced Energy Systems (IDAES) supported by the Office of Fossil Energy and Carbon Management's Simulation-Based Engineering/Crosscutting Research Program.

This project was funded by the Department of Energy, National Energy Technology Laboratory an agency of the United States Government, through a support contract. Neither the United States Government nor any agency thereof, nor any of their employees, nor the support contractor, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## REFERENCES



1. Jianxin J, Simpson TW, Siddique Z. Product family design and platform-based product development: a state-of-the-art review. *Journal of Intelligent Manufacturing* 18: 5-29 (2007).
2. Stinchfield G, Biegler LT, Eslick JC, Jacobson C, Miller DC, Siirola J, Zhang C, Zhang Q, Laird CD. Optimization-Based Approaches for Design of Chemical Process Families using ReLU Surrogates. *In proceedings, FOCOPO/CPC* (2023).
3. Lieberman M. The Learning Curve and Pricing in the Chemical Processing Industries. *RAND Journal of Economics* 15: 198-213 (1984).
4. Gazzaneo V, Watson M, Ramsayer CB, Kilwein ZA, Alves V, Lima FV. A techno-economic analysis framework for intensified modular systems. *Journal of Advanced Manufacturing and Processing* 4.3 (2022)
5. Bynum ML, Hackebeil GA, Hart WE, Laird CD, Nicholson BL, Siirola JD, Watson JP, Woodruff DL. Pyomo-optimization modeling in python. 67. Berlin/Heidelberg, Germany: Springer (2021).
6. Morgan JC, Chinen AS, Omell B, Bhattacharyya D, A. Deshpande. *Carbon capture simulation for industry impact (CCSI2)*, 3.2.1. mea steady state model (2022)
7. Gonzalez-Zugasti JP. Models for Platform-Based Product Family Design. *Massachusetts Institute of Technology*, (2000)
8. Zhou K, Du G, Jiao RJ. Personalized service product family design optimization considering crowdsourced service operations. *Computers & Industrial Engineering*, 166:107973 (2022)
9. Pirmoradi Z, Hajikolaie KH, Wang GG. Designing scalable product families by the radial basis function-high-dimensional model representation metamodelling technique. *Engineering Optimization* 47:1423-1439 (2015)
10. Wright TP. Factors Affecting the Cost of Airplanes. *Journal of the Aeronautical Sciences*, 3 (1936)
11. Argote L, Epple D. Learning curves in manufacturing. *Science, New Series*, 247:920-924, (1990)
12. Weber RS, Snowden-Swan LJ. The economics of numbering up a chemical process enterprise. *Journal of Advanced Manufacturing Process* 1 (1-2), e10011 (2019)
13. ReVelle CS, Ralph WS. Central facilities location. *Geographical analysis* 2.1: 30-42 (1970)
14. Simpson TWD'Souza BS. Assessing variable levels of platform commonality within a product family using a multiobjective genetic algorithm. *Concurrent Engineering* 12:119-129, (2004)

and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator

# A Study on Accelerated Convergence of Cyclic Steady State in Adsorption Process Simulations

Sai Gokul Subraveti<sup>a\*</sup>, Kian Karimi<sup>b</sup>, Matteo Gazzani<sup>b</sup>, and Rahul Anantharaman<sup>a</sup>

<sup>a</sup> SINTEF Energy Research, Trondheim, Norway

<sup>b</sup> Utrecht University, Copernicus Institute of Sustainable Development, Princetonlaan 8a, 3584 CB Utrecht, the Netherlands

\* Corresponding Author: [sai.gokul.subraveti@sintef.no](mailto:sai.gokul.subraveti@sintef.no).

---

## ABSTRACT

Cyclic adsorption processes attain a cyclic-steady state (CSS) condition by undergoing repeated cycles in time, owing to their transient and modular nature. Mathematically, solving a set of underlying nonlinear partial differential equations iteratively for different steps in a cycle until the CSS condition is attained presents a computational challenge, making the simulation and optimization of cyclic adsorption processes time-consuming. This paper focuses on expediting the CSS convergence in adsorption process simulations by implementing two vector-based acceleration methods that offer quadratic convergence akin to Newton's methods. These methods are straightforward to implement, requiring no prior knowledge of the first derivatives (or Jacobian). The study demonstrates the efficacy of accelerated convergence by considering two adsorption processes that exhibit complex dynamics, namely, a four-step vacuum swing adsorption and a six-step temperature swing adsorption cycles for post-combustion CO<sub>2</sub> capture. The case studies showcase the potential for improved computational efficiency in adsorption process simulations.

---

**Keywords:** cyclic adsorption processes, modeling, optimization, process design, acceleration methods

## INTRODUCTION

Cyclic adsorption processes are widely employed for industrial gas separations [1, 2]. The underlying separation mechanism involves the solid adsorbent to selectively adsorb one or more gases from a gaseous mixture. Depending on the choice of regeneration, i.e., by either varying temperature or pressure or both, the adsorbed gases are recovered or removed. Accordingly, different process operational modes such as pressure swing adsorption (PSA), vacuum swing adsorption (VSA), temperature swing adsorption (TSA), etc. are implemented [3].

Unlike distillation and absorption, which typically operate at steady-state conditions, cyclic adsorption processes are transient in nature and are operated cyclically in a sequence of constituent steps, thus attaining a cyclic steady-state condition (CSS), instead of a true steady state [2]. From a process design perspective, these processes are dictated by the set of nonlinear partial differential equations (PDEs) resulting from mass, momentum, and energy balances. Further, the modular nature allows these processes to undergo several

constituent steps in each cycle. The choice and the sequence of steps, as well as the direction of flows, can give rise to several process configurations [4], which need to be optimized [5]. Another critical design parameter is the choice of the solid adsorbent and the contactor structure. Therefore, designing cyclic adsorption processes for a given adsorbent requires solving the nonlinear PDEs for each step repeatedly until CSS and then calculating the key process performance indicators based on the transient profiles of state variables obtained at the CSS. The CSS ensures that these profiles no longer vary as the numerical iterations are continued.

Despite such flexibility in adsorption process design, the main bottleneck for designing optimal process configurations lies in very high computational costs associated with repeated solving of nonlinear PDEs in process simulations to attain the CSS condition from an arbitrary initial condition. Although simulating one process configuration for a given adsorbent takes a few minutes of computational time [5], process models are often coupled with optimizers where thousands of operating conditions are probed, which makes the process design and

optimization computationally demanding. Given the flexibility of synthesizing different process configurations and the recent advent of a multitude of adsorbents for gas separations [6], the problem of finding the optimal design for a given separation further exacerbates as evaluating different process configurations and adsorbents will be a computationally daunting task.

To speed up adsorption process simulations, there have been some previous efforts to accelerate the determination of CSS [7-13]. For instance, Croft and LeVan in their first paper [7] incorporated Newton's method to speed up the iterations to attain CSS. Newton's method provided quadratic convergence near the CSS solution and converged faster than the cycle iterative procedure. However, the extra computation time for the calculation of the Jacobian matrix in each Newton iteration impaired the convergence acceleration [7, 8]. In the third paper of the series [9], Ding and LeVan developed several acceleration algorithms for the convergence of adsorption process simulations, which include a hybrid Newton-Broyden method, an iterative secant method, a sensitivity interpolation method, and a dynamic error tolerance method. These methods achieved better accelerations

compared to the Newton's method [9]. Despite this, some of the methods rely on the prior knowledge of the Jacobian matrix. On the other hand, Nilchan and Pantelides [11] completely discretized the PDEs into algebraic equations and imposed the CSS condition as a constraint. Perturbation techniques have also been explored for accelerated convergence of CSS [12]. Moreover, Pai et al. [13] recently developed machine learning models to predict approximate CSS condition which is then fed as an initial condition to the PDE-based process model for rapid determination of CSS condition. The machine learning model for initial condition prediction was developed for a four-step VSA cycle based on the CSS profiles from hundreds of different operating conditions. In the context of other related separations, methods such as the single shooting method were considered for the convergence of semicontinuous distillation process simulations [14].

This study focuses on implementing two acceleration methods for the convergence of the CSS condition in adsorption process simulations, with an overall goal of significantly reducing the computational times of current adsorption process design and optimization tools. These acceleration methods require no prior knowledge of the

**Table 1:** Governing partial differential equations of the adsorption process model.

**Gas-phase component mass balance**

$$\frac{\partial y_i}{\partial t} + \frac{y_i}{P} \frac{\partial P}{\partial t} - \frac{y_i}{P} \frac{\partial T}{\partial t} = \frac{T}{P} D_v \frac{\partial}{\partial z} \left( \frac{P}{T} \frac{\partial y_i}{\partial z} \right) - \frac{T}{P} \frac{\partial}{\partial z} \left( \frac{y_i P}{T} v \right) - \frac{RT}{P} \frac{1 - \varepsilon}{\varepsilon} \frac{\partial q_i}{\partial t}$$

**Gas-phase total mass balance**

$$\frac{1}{P} \frac{\partial P}{\partial t} - \frac{1}{T} \frac{\partial T}{\partial t} = - \frac{T}{P} \frac{\partial}{\partial z} \left( \frac{P}{T} v \right) - \frac{RT}{P} \frac{1 - \varepsilon}{\varepsilon} \sum_i^m \frac{\partial q_i}{\partial t}$$

**Column energy balance**

$$\left[ \frac{1 - \varepsilon}{\varepsilon} \left( \rho_s C_{p,s} + C_{p,a} \sum_{i=1}^m q_i \right) \right] \frac{\partial T}{\partial t} = \frac{K_z}{\varepsilon} \frac{\partial^2 T}{\partial z^2} - \frac{C_{p,g}}{R} \frac{\partial(vP)}{\partial z} - \frac{C_{p,g}}{R} \frac{\partial P}{\partial t} - \frac{1 - \varepsilon}{\varepsilon} C_{p,a} T \sum_{i=1}^m \frac{\partial q_i}{\partial t} + \frac{1 - \varepsilon}{\varepsilon} \sum_{i=1}^m (-\Delta H_i) \frac{\partial q_i}{\partial t} - \frac{2h_{in}(T - T_w)}{\varepsilon r_{in}}$$

**Wall energy balance**

$$\rho_w C_{p,w} \frac{\partial T_w}{\partial t} = K_w \frac{\partial^2 T_w}{\partial z^2} + \frac{2r_{in} h_{in}(T - T_w)}{r_{out}^2 - r_{in}^2} - \frac{2r_{out} h_{out}(T_w - T_a)}{r_{out}^2 - r_{in}^2}$$

**Darcy's pressure drop**

$$v = \frac{4}{150\mu} \left( \frac{\varepsilon}{1 - \varepsilon} \right)^2 r_p^2 \left( - \frac{\partial P}{\partial z} \right)$$

**Linear driving force model**

$$\frac{\partial q_i}{\partial t} = k_i (q_i^* - q_i)$$

Jacobian matrix and are easier to implement for any cyclic process. Simulation of two different adsorption processes is chosen to test the two acceleration methods. As a first case, a four-step VSA cycle designed for post-combustion CO<sub>2</sub> capture [5] was used for the process simulation. The second case explores the simulation of a six-step TSA cycle designed again for post-combustion CO<sub>2</sub> capture [15]. Despite both processes exhibiting complex dynamics and taking several cycles to achieve CSS, they differ in their process dynamics and how they converge to CSS. This makes them an ideal case study for testing the vector-based acceleration methods.

## METHODOLOGY

### Adsorption Process Model

The mathematical model for simulating adsorption column dynamics consists of a system of nonlinear PDEs obtained by solving mass, momentum, and energy balances [5]. The model assumptions include (1) ideal gas, (2) axially dispersed plug flow representing the bulk flow, (3) linear driving force model describes the bulk-to-surface mass transfer, (4) there exist no radial gradients for state variables, i.e., composition, pressure, and temperature, (5) Darcy law used for pressure drop calculations, (6) thermal equilibrium between the gas and the solid phases, and (7) uniform column properties. The resulting governing equations based on these assumptions are provided in **Table 1**.

The model is numerically solved by discretizing the spatial terms in PDEs into 30 finite volumes with a total variation diminishing (TVD) scheme involving the van-Leer flux limiter [5]. The resulting ordinary differential equations (ODEs) are then integrated based on standard ODE solvers in Python [16]. The process simulations are carried out using a uni-bed approach where a single column undergoes all constituent steps sequentially. The column is initialized with feed composition at low pressure and the cycle is simulated until the CSS condition. At CSS, state variables such as composition, pressure, and temperature profiles are obtained.

### Cyclic-steady state (CSS) condition

Several mathematical criteria can be employed for the attainment of CSS from a pre-defined initial condition in adsorption process simulations [5, 17-19]. For instance, few studies assumed that the process reaches CSS after a very large number of cycle iterations [16]. Most of the other studies, however, assume that the CSS criterion is met when absolute integral differences in axial profiles of state variables [5] or absolute overall mass balance errors [5, 18] fall below a tolerance limit. Effendy et al. recently proposed a rigorous CSS criterion, reducing the differences between the current and CSS states of axial profiles to a set tolerance limit without needing prior

knowledge of the CSS [19].

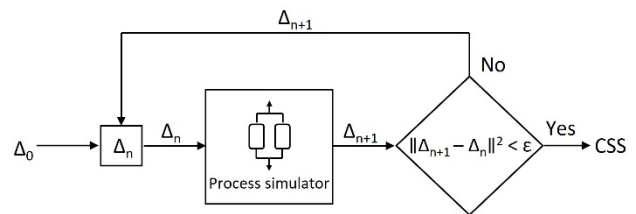
This study uses the same definition for the CSS condition as that of the case study [5], which assumes that the CSS criterion is met when the overall mass balance error for the process is less than 0.5%.

### Methods for Accelerated Convergence of CSS condition in process simulations

An important basis for implementing the acceleration methods in the present study is that the process simulation of cyclic adsorption processes is a fixed point (or Picart) iteration problem [19]. Schematically shown in **Figure 1**, a number of cycles are simulated (cycle iterations) from an arbitrary initial condition to a fixed-point CSS condition. If  $f$  is a function representing the adsorption process simulator, then the fixed-point iteration problem can be written as

$$\Delta_{n+1} = f(\Delta_n) \quad (1)$$

In the above equation, the adsorption process simulator,  $f$ , takes the previous state of the adsorption column ( $\Delta_n$ ) as an input and returns the subsequent state of the adsorption column ( $\Delta_{n+1}$ ). Note that  $\Delta$  represents a vector of state variables such as gas-phase composition, solid-phase concentrations, pressure, and column and wall temperatures across the adsorption columns. After



**Figure 1.** Schematic of a cyclic adsorption process simulator. Adapted from [19].

several iterations, the CSS condition is achieved. Although fixed point iterations are the simplest way to obtain nonlinear solutions without any prior knowledge of  $f$  and its derivative [20], linear convergence to CSS makes it computationally slow [19].

Alternatively, Eq. 1 can be reformulated as a root-finding problem:

$$\varphi(\Delta) = f(\Delta) - \Delta = 0 \quad (2)$$

A common approach for solving Eq. 2 is Newton's method,

$$\Delta_{n+1} = \Delta_n - \frac{\varphi(\Delta_n)}{\varphi'(\Delta_n)} \quad (3)$$

which provides a quadratic and faster convergence to solutions, compared to fixed-point iterations. The main drawback of Newton's method as a convergence

accelerator is the prior knowledge of first derivatives or Jacobian, i.e.,  $\varphi'(\Delta)$ , which is not straightforward to calculate for many realistic systems.

Without affecting the order of convergence, Newton's method can be modified into "derivative-free" algorithms such as Steffensen's methods or its equivalents [19, 20]. Steffensen's convergence acceleration for a scalar nonlinear fixed-point equation,  $x = g(x)$ , can be represented as:

$$x_{n+3} = x_n - \frac{(x_{n+1} - x_n)^2}{(x_{n+2} - 2x_{n+1} + x_n)} \quad (4)$$

Practical problems such as adsorption process simulations often deal with many unknown variables and systems of nonlinear PDEs. Obtaining the solutions typically requires the spatial discretization of PDEs which introduces a vector of state variables ( $\Delta$ ) across the column. Several vector-based acceleration methods are available in the literature to deal with vector nonlinear fixed-point problems [20].

In this study, two vector-based acceleration methods proposed in the literature are considered for the accelerated convergence of the CSS condition in adsorption process simulations. Particularly, the acceleration algorithms proposed by Irons and Tuck [22] and Graves-Morris [23] are evaluated for CSS convergence acceleration. It is worth noting that both methods are the vector extensions of Steffensen's scalar method (Eq. 4).

### Irons and Tuck Acceleration Method

The vector acceleration method proposed by Iron and Tuck [22] can be expressed as:

$$\Delta_{n+3} = \Delta_{n+2} - \frac{(\Delta_{n+2} - \Delta_{n+1}) \cdot (\Delta_{n+2} - 2\Delta_{n+1} + \Delta_n)}{\|\Delta_{n+2} - 2\Delta_{n+1} + \Delta_n\|^2} (\Delta_{n+2} - \Delta_{n+1}) \quad (5)$$

Note that this acceleration method is invoked alternately with a basic fixed-point cycle iteration.

### Graves-Morris Acceleration Method

The approach of Graves-Morris [23] is given by:

$$\Delta_{n+3} = \Delta_{n+1} - \frac{\|\Delta_{n+2} - 2\Delta_{n+1} + \Delta_n\|^2}{(\Delta_{n+1} - \Delta_n) \cdot (\Delta_{n+2} - 2\Delta_{n+1} + \Delta_n)} (\Delta_{n+2} - \Delta_{n+1}) \quad (6)$$

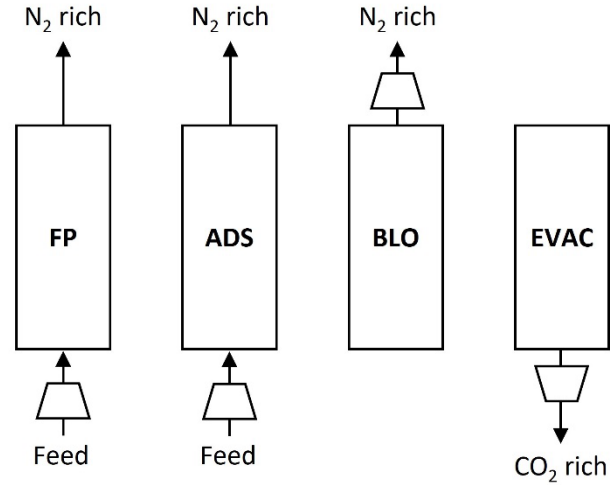
This accelerator is also applied alternately with a basic fixed-point iteration in adsorption process simulations.

## RESULTS AND DISCUSSION

Two case studies involving different adsorption processes are considered to test the convergence accelerator methods. The first examines a VSA cycle, where the process dynamics are heavily influenced by the pressure swing between the adsorption and desorption steps. The second explores a TSA cycle, where the dynamics are driven by the temperature swing during the process. Thus, the two cases differ in terms of their process dynamics and their convergence to CSS.

### Case 1: Four-step VSA cycle

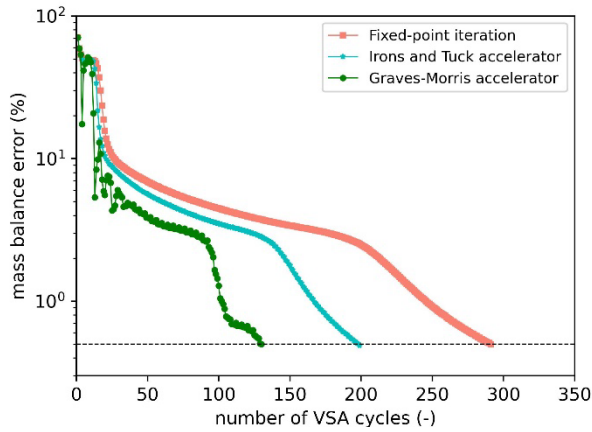
Here, the vector acceleration methods are incorporated into the process simulation of a four-step VSA cycle separating the binary mixture of CO<sub>2</sub>/N<sub>2</sub> for postcombustion CO<sub>2</sub> capture [5]. The VSA process was designed to recover CO<sub>2</sub> from a binary mixture of 15% CO<sub>2</sub> and 85% N<sub>2</sub> at ambient pressure, which represents dry flue gas from coal-fired power plants.



**Figure 2.** Four-step vacuum swing adsorption cycle designed for post-combustion CO<sub>2</sub> capture [5].

**Figure 2** illustrates the schematic of the four-step VSA cycle. The cycle consists of feed pressurization (FP), adsorption (ADS), co-current blowdown (BLO), and counter-current evacuation (EVAC) steps. Each step can be implemented using appropriate boundary conditions and isotherm parameters provided by Haghpanah et al. [5]. Commercial zeolite 13X was used as the adsorbent. The example considered represents the complex dynamics of adsorption processes, and several cycles are needed to reach CSS. Similar to reference work, CO<sub>2</sub> and N<sub>2</sub> isotherms on zeolite 13X were described using the competitive dual-site Langmuir isotherm model. Finally, the cycle operating conditions used for the simulation are as follows:  $P_H = 1$  bar,  $P_I = 0.2$  bar,  $P_L = 0.1$  bar,  $v_0 = 1.0$  m s<sup>-1</sup>,  $t_{FP} = 15$  s,  $t_{ADS} = 15$  s,  $t_{BLO} = 30$  s,  $t_{EVAC} = 40$  s. Note that  $P_H$ ,  $P_I$ , and  $P_L$  are the feed, intermediate, and low pressures in adsorption, blowdown, and evacuation steps, respectively. The feed is introduced into the column in the adsorption step with an interstitial velocity  $v_0$ .  $t_{FP}$ ,  $t_{ADS}$ ,  $t_{BLO}$ , and  $t_{EVAC}$  are the step durations of feed pressurization, adsorption, blowdown, and evacuation steps, respectively. The column was first initialized with feed composition, and the simulation was conducted until CSS.





**Figure 3.** Convergence of overall mass balance error for the four-step VSA process simulation using the fixed-point iteration, Irons and Tuck [22] algorithm, and Graves-Morris method [23].

**Figure 3** shows the convergence of the overall mass balance error for the simulation. Without any accelerator, the fixed-point iteration took 291 cycles to reach the CSS. The four-step VSA process simulations are repeated for the same operating conditions with the two acceleration methods. The convergence of the overall mass balance for the simulation with the accelerators can also be seen in **Figure 3**. Using the acceleration algorithm by Irons and Tuck, the CSS convergence was achieved after 199 cycles. The computations were reduced by approximately one-third. On the other hand, Graves-Morris's acceleration method almost reduced the number of iterations to CSS by less than half, i.e., 130 cycles, outperforming the Irons and Tuck algorithm. The CO<sub>2</sub> purity and recovery obtained at CSS using all the three methods are reported in **Table 2**. The acceleration methods offered faster convergence to CSS while achieving the same performance as that of the conventional fixed-point iteration with rather straightforward implementation.

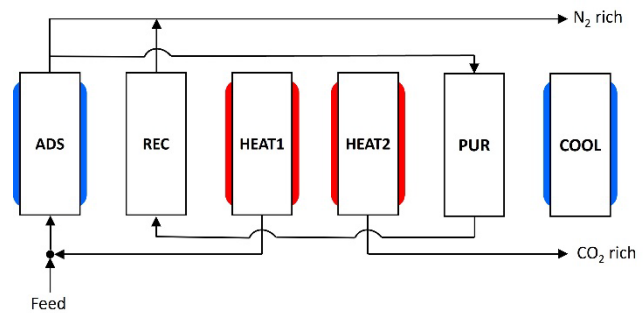
**Table 2:** Performance of the VSA process using three different approaches for the convergence of CSS.

Method	CO <sub>2</sub> purity (%)	CO <sub>2</sub> recovery (%)
Fixed-point iteration	79.64	37.36
Irons & Tuck [21]	79.65	37.36
Graves-Moris [22]	79.64	37.36

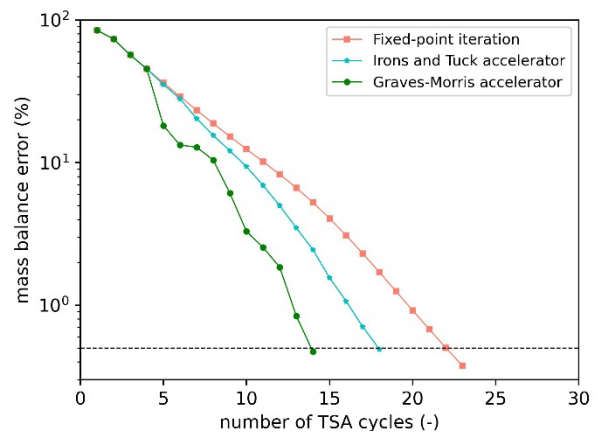
### Case 2: Six-step TSA cycle

In this case study, the vector acceleration methods are employed in the process simulation of a six-step TSA cycle proposed by Joss et al. [15] to separate the binary

feed mixture of 12% CO<sub>2</sub> and 88% N<sub>2</sub> using commercial zeolite 13X. **Figure 4** shows the six-step TSA cycle consisting of the adsorption (ADS), recovery (REC), two heating (HEAT), purge (PUR), and cooling (COOL) steps. For simulating this process, the cycle operating conditions and the simulation parameters have been retrieved from Joss et al. [15]. The adsorption step occurs under a constant pressure of 1.3 bar and at an ambient temperature of 300 K. In the heating steps, the column is externally heated to 420 K to desorb CO<sub>2</sub>. The cycle is simulated for the following operating conditions: T<sub>COOL</sub> = 300 K, T<sub>HEAT</sub> = 420 K, v<sub>0</sub> = 1.4 m s<sup>-1</sup>, t<sub>ADS</sub> = 150 s, t<sub>REC</sub> = t<sub>PUR</sub> = 25 s, t<sub>HEAT1</sub> = 150 s, t<sub>HEAT2</sub> = 600 s, t<sub>COOL</sub> = 600 s. It is worth noting that the durations of TSA cycles are much longer compared to VSA cycles due to heat transfer limitations. Here, a column that is completely saturated with N<sub>2</sub> is used as an initial condition to simulate the cycle until CSS.



**Figure 4.** Six-step temperature swing adsorption cycle proposed by Joss et al. [15] for post-combustion CO<sub>2</sub> capture.



**Figure 5.** Convergence of overall mass balance error for the six-step TSA process simulation using the fixed-point iteration, Irons and Tuck [22] algorithm, and Graves-Morris method [23].

**Figure 5** and **Table 3** present the comparative performances of the vector acceleration methods for this

simulation. As can be seen from **Figure 5**, the fixed-point iteration needed 23 cycles to reach CSS, where the overall mass balance convergence is less than 0.5%. When the accelerators are incorporated into the process simulations, the cycle converges to CSS in 18 and 14 cycles using Irons and Tuck and Graves-Morris's methods, respectively. While the TSA cycle reached CSS in fewer cycles compared to the VSA cycle, each TSA cycle requires a significantly longer computational time. Hence, utilizing accelerator methods facilitated almost 1.25 – 1.67 times faster CSS convergence compared to the fixed-point iteration. The true impact of accelerators on computational speeds up will further be realized in process optimizations. It is worth reiterating that the implementation of the considered acceleration methods is straightforward. Table 3 reports the comparable CO<sub>2</sub> purity and recovery obtained at CSS using all the methods.

**Table 3:** Performance of the TSA process using three different approaches for the convergence of CSS.

Method	CO <sub>2</sub> purity (%)	CO <sub>2</sub> recovery (%)
Fixed-point iteration	93.08	99.02
Irons & Tuck [21]	93.06	98.92
Graves-Morris [22]	93.06	98.94

## CONCLUSIONS

The computational complexity in cyclic adsorption process simulations mainly arises from solving a set of nonlinear partial differential equations iteratively for every cycle until the cyclic-steady state (CSS) is achieved making the simulation and optimization routines of these processes exceedingly time-consuming. Despite prior efforts to expedite CSS convergence, the computational burden of CSS still persists in adsorption process simulations.

This study contributes to overcoming the computational challenges of CSS determination in adsorption process simulations by implementing two vector-based acceleration methods proposed previously in the literature. The acceleration methods of Irons and Tuck and Graves-Morris can offer quadratic convergence near the CSS solutions, like Newton's method without requiring the prior information of the first derivatives or Jacobian. These methods are straightforward to implement in adsorption process simulations. The acceleration capabilities of these methods are demonstrated by considering two different adsorption processes, namely, a four-step VSA cycle and a six-step TSA cycle, which undergo complex dynamics and take multiple cycles to reach CSS. The results showed that the Graves-Morris accelerator provided expedited convergence by speeding the convergence of VSA simulations more than double. For the same case, Irons and Tuck's method reduced the cycle

iterations to CSS by one-third. When the accelerator methods are incorporated into the process simulations of a six-step TSA cycle, the Graves-Morris accelerator marginally provides better convergence than Irons and Tuck's method by reducing the number of iterations to 14 from the original 23 fixed-point cycle iterations. In the context of process optimization, where thousands of operating conditions must be probed, these acceleration methods can significantly reduce the overall computational times with no extra effort. This work is a first step in addressing the computational challenges associated with adsorption process simulations. In the future, these methods will be extended to different adsorption processes and efforts will be directed to further improve the convergence of CSS. For instance, incorporating machine learning principles into the acceleration framework can further boost CSS convergence speeds.

## ACKNOWLEDGEMENTS

This work was carried out as a part of the project "Removing non-CO<sub>2</sub> greenhouse gas emissions to support ambitious climate transitions (REPAIR)" (Project number: 101069905) funded by the European Commission via the European Climate, Infrastructure and Environment Executive Agency (CINEA) within the Horizon Europe framework.

## REFERENCES

1. Sircar S. Pressure Swing Adsorption. *Ind. Eng. Chem. Res.* 41: 1389–1392 (2002).
2. Ruthven DM, Farooq S, Knaebel KS. *Pressure Swing Adsorption*. VCH Publishers: New York (1994).
3. Rajendran A, Subraveti SG, Pai KN, Prasad V, Li Z. How Can (or Why Should) Process Engineering Aid the Screening and Discovery of Solid Sorbents for CO<sub>2</sub> Capture? *Acc. Chem. Res.* 56 (17): 2354–2365 (2023).
4. Haghpanah R, Nilam R, Rajendran A, Farooq S, Karimi IA. Cycle Synthesis and Optimization of a VSA Process for Postcombustion CO<sub>2</sub> Capture. *AIChE J.* 59: 4735–4748 (2013).
5. Haghpanah R, Majumder A, Nilam R, Rajendran A, Farooq S, Karimi IA, Amanullah M. Multiobjective Optimization of a Four-Step Adsorption Process for Postcombustion CO<sub>2</sub> Capture Via Finite Volume Simulation. *Ind. Eng. Chem. Res.* 52: 4249–4265 (2013).
6. Farmahini AH, Krishnamurthy S, Friedrich D, Brandani S, Sarkisov L. Performance-Based Screening of Porous Materials for Carbon Capture. *Chem. Rev.* 121: 10666–10741 (2021).
7. Croft DT, LeVan MD. Periodic States of Adsorption Cycles—I. Direct Determination and Stability. *Chem.*

- Eng. Sci.* 49: 1821–1829 (1994).
8. Croft DT, LeVan MD. Periodic States of Adsorption Cycles—II. Solution Spaces and Multiplicity. *Chem. Eng. Sci.* 49: 1831–1841 (1994).
  9. Ding Y, LeVan MD. Periodic States of Adsorption Cycles III. Convergence Acceleration for Direct Determination. *Chem. Eng. Sci.* 56: 5217–5230 (2001).
  10. Ding Y, Croft DT, LeVan MD. Periodic States of Adsorption Cycles IV. Direct Optimization. *Chem. Eng. Sci.* 57: 4521–4531 (2002).
  11. Nilchan S, Pantelides C. On the Optimisation of Periodic Adsorption Processes. *Adsorption* 4: 113–147 (1998).
  12. Wilson SJ, Webley PA. Perturbation Techniques for Accelerated Convergence of Cyclic Steady State (CSS) in Oxygen VSA Simulations. *Chem. Eng. Sci.* 57: 4145–4159 (2002).
  13. Pai KN, Prasad V, Rajendran A. Experimentally Validated Machine Learning Frameworks for Accelerated Prediction of Cyclic Steady State and Optimization of Pressure Swing Adsorption Processes. *Sep. Purif. Technol.* 241: 116651 (2020).
  14. Madabhushi PB, Adams TA. On the Application of Shooting Method for Determining Semicontinuous Distillation Limit Cycles. *Chem. Eng. Res. Des.* 160: 370–382 (2020).
  15. Joss L, Gazzani M, Mazzotti M. Rational design of temperature swing adsorption cycles for post-combustion CO<sub>2</sub> capture. *Chem. Eng. Sci.* 158: 381–394 (2017).
  16. Shampine LF, Reichelt MW. The MATLAB ODE Suite. *SIAM J. Sci. Comput.*, 18: 1–22 (1997).
  17. Bhadra S, Farooq S. Separation of Methane-Nitrogen Mixture by Pressure Swing Adsorption for Natural Gas Upgrading. *Ind. Eng. Chem. Res.* 50: 14030–14045 (2011).
  18. Jayaraman A, Hernandez-Maldonado AJ, Yang RT, Chinn D, Munson CL, Mohr DH. Clinoptilolites for Nitrogen/Methane Separation. *Chem. Eng. Sci.* 59: 2407–2417 (2004).
  19. Effendy S, Farooq S, Ruthven DM. A Rigorous Criterion for Approach to Cyclic Steady-State in PSA Simulations. *Chem. Eng. Sci.* 160: 313–320 (2017).
  20. Ramière I, Helfer T. Iterative Residual-Based Vector Methods to Accelerate Fixed Point Iterations. *Comput. Math. with Appl.* 70 (9): 2210–2226 (2015).
  21. Cordero A, Hueso JL, Martínez E, Torregrosa JR. Steffensen Type Methods for Solving Nonlinear Equations. *J. Comput. Appl. Math.* 236 (12): 3058–3064 (2012).
  22. Irons BM, Tuck R. A Version of the Aitken Accelerator for Computer Iteration. *Int. J. Numer. Meth. Eng.* 1:275–277 (1969).
  23. Graves-Morris PR. Extrapolation Method for Vector

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Optimal Design of Intensified Towers for CO<sub>2</sub> Capture with Internal, Printed Heat Exchangers

Stephen Summits<sup>a</sup>, Paul Akula<sup>a</sup>, Debangsu Bhattacharyya<sup>a,d\*</sup>, Grigorios Panagakos<sup>b,d</sup>, Benjamin Omell<sup>c</sup>, and Michael Matuszewski<sup>c</sup>

<sup>a</sup> Department of Chemical and Biomedical Engineering, West Virginia University, Morgantown, West Virginia 26506, USA

<sup>b</sup> Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA

<sup>c</sup> National Energy Technology Laboratory (NETL), Pittsburgh, Pennsylvania 15236, USA

<sup>d</sup> NETL Support Contractor, Pittsburgh, PA 15236, USA

\* Corresponding Author: debangsu.bhattacharyya@mail.wvu.edu.

## ABSTRACT

Solvent-based carbon capture processes typically suffer from the temperature rise of the solvent due to the heat of absorption of CO<sub>2</sub>. This increased temperature is not thermodynamically favorable and results in a significant reduction in performance in the absorber column. As opposed to interstage coolers, which only remove, cool, and return the solvent at discrete locations in the column, internal coolers that are integrated with the packing can cool the process inline, which can result in improved efficiency. This work presents the modeling of these internal coolers within an existing generic, equation-oriented absorber column model that can cool the process while allowing for simultaneous mass transfer. Optimization of this model is also performed, which is capable of optimally choosing the best locations to place these devices, such that heat removal and mass transfer area are balanced. Results of the optimization have shown that optimally placed cooling elements result in a significant increase in the capture efficiency of the process, compared to a similar column with no internal cooling, with a common trend being the cooling of the column in the temperature bulge region. It is observed that by optimally placing an internal cooler, the solvent flow rate can be decreased, and the CO<sub>2</sub> lean loading can be increased while still maintaining the same efficiency. These process changes can lead to a substantial reduction in costs due to lower reboiler duty.

**Keywords:** CO<sub>2</sub> capture, optimization, monoethanolamine, process intensification

## INTRODUCTION

Post-combustion CO<sub>2</sub> capture is a critical approach for achieving net-zero emissions. Among many potential technologies for post-combustion CO<sub>2</sub> capture, solvent-based capture technologies are at the forefront due to their maturity, ease of operation, availability of efficient contactors for solvent-based capture systems, and many other advantages.

A common family of aqueous solvents utilized in this process is one that contains amine functional groups. The solvent absorption process excels compared to other capture methods, such as solid sorbents, due to its ability to react with carbon dioxide even at low partial pressure, as well as having a high capacity of absorption [1]. There are, however, two key disadvantages associated with

this process. The first disadvantage is that the high concentration of water in these solvents requires a significant amount of steam for the use of the reboiler in the regeneration process [2]. This steam utilization results in the energy intensity of the process being up to 4 MJ per ton of CO<sub>2</sub> captured, which accounts for the majority of the operating cost. The second disadvantage is the high exothermic heat of absorption for CO<sub>2</sub>. In the case of a 30 wt% MEA solvent, the heat of absorption can range from 84 to 100 kJ/mol CO<sub>2</sub> [3, 4]. This heat can cause temperatures within the absorber tower to rise substantially, thus increasing the equilibrium pressure of CO<sub>2</sub> by several orders of magnitude, resulting in a reduced mass transfer rate [3].

To keep the solvent capture process operating as efficiently as possible, heat needs to be removed from

the absorber column to operate the tower under more favorable conditions. Therefore, it is common to use inter-stage coolers [5, 6], which withdraw a portion of the solvent or the entire solvent from discrete locations within the column, and cool and return it in the next stage below. Karimi et al. showed that optimal placement of an inter-cooler within the absorber column can result in energy savings as high as 7.27% by reducing the amount of solvent flow required for operation [7]. However, while cooling is achieved at discrete locations, intercoolers fail to achieve continuous heat removal along the height of the tower to reach an optimal temperature profile to maximize performance and/or economics. Therefore, the use of a cooling method that is integrated with the packing of the tower is needed.

Additive manufacturing has become a rapidly growing method of developing innovative technologies, with advancements in 3D printing leading to products that would be impossible to create using traditional manufacturing methods. A recent novel application of 3D printing is structured metal packing that can be utilized for a packed column. The intensified packing device created at Oak Ridge National Laboratory is a 3D printed structured packing element with a double-walled design, which creates two disjoint flow channels in which the process fluid can be kept in a flow channel separate from the cooling fluid in the other [8]. Tests utilizing this device have shown that carbon capture can be increased between 3 to 15% compared to columns with no cooling [9, 10].

A disadvantage of such a device is that the cooling water channels reduce the total amount of available process volume, which reduces the mass transfer area. In addition, there are several other aspects that need to be evaluated for the optimal configuration of absorbers in the presence of such intensified towers. This work seeks to answer the following questions. Where are the optimal locations at which to implement internal cooling such that performance is maximized? What is the best configuration of cooling water flow within the devices? How does variance in operating set points affect the optimal design and performance of the internal coolers?

## MODELING

### Column Model

The absorber column model employed in this work was developed by Akula et al. [11]. This model is a rate-based solvent absorber model that can be applied to many processes by using the built-in physical and chemical property models. For convenience, this work uses 30 wt% MEA as the solvent.

For the reactive absorption process, rate-based models are significantly more accurate than equilibrium-based models, but are more computationally expensive

due to the consideration of transport through the films, especially the liquid film. To reduce this computation expense through removing the need to model film interactions, an enhancement factor method is used that accounts for the increase in mass transfer due to chemical reactions by using a set of algebraic equations [12]. Equations 1 and 2 are used to describe the enhancement factor,  $E$ , in terms of dimensionless concentration of MEA,  $Y_{MEA}^i$ .

$$E = 1 + (E_{\infty}^* - 1) \frac{(1 - Y_{MEA}^i)}{(1 - Y_{CO_2}^b)} \quad (1)$$

$$E = Ha \sqrt{Y_{MEA}^i} \frac{(1 - Y_{CO_2}^*)}{(1 - Y_{CO_2}^b)} \quad (2)$$

The phase energy balance of the model considers three heat transfer mechanisms/sources/sinks: convective heat transfer between phases, heat of absorption of  $CO_2$ , and heat of vaporization of water. The interphase heat transfer is calculated by using a corrected heat transfer coefficient, which is required for high mass fluxes (Equation 3). Both heat of absorption and vaporization are accounted for in the liquid phase energy balance (Equation 4). The heat of absorption is fixed at -85 kJ/mol  $CO_2$  for simplicity [13], while the heat of vaporization being used was derived from [14].

$$Q_V = h'_V a_e (T_L - T_V) \quad (3)$$

$$Q_L = Q_V + N_{CO_2,V} \Delta H_{abs} - N_{H_2O,V} \Delta H_{vap} \quad (4)$$

### Internal Heat Exchanger Model

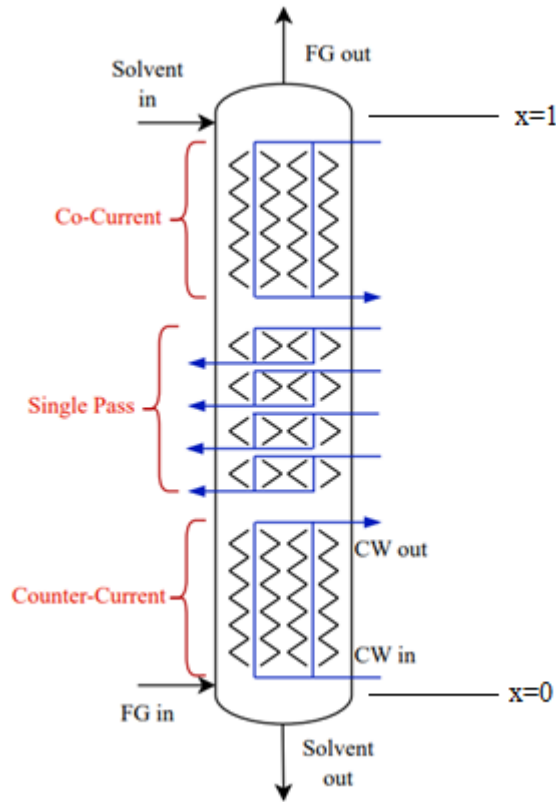
The modeling for the internal heat exchanger packing was modified from Moore et al., which considered a single bed absorber column. This model was modified to account for individual discretized elements of the column in which an internal heat exchanger can be independently placed from other elements. In this model, it is assumed that all heat removal from the process to the cooling water occurs through the liquid phase due to the higher wetted area and conductivity in the liquid phase leading to a negligible amount of heat being directly transferred to the gaseous phase. It was also assumed to be smooth transition between standard and intensified packing, if flooding velocity is not surpassed. The two decision variables are a binary variable,  $y$ , for each column element to indicate placement of an internal heat exchanger and a voidage term,  $\epsilon^{cw}$ , that accounts for the volume occupied by the intensified cooler. Equation 5 calculates this heat transfer in each column element where the overall heat transfer coefficient, assumed to be constant, is from [9], which is a conservative estimate using the packing area, rather than wetted area, and is based on experimental data. In this equation,  $i$  is length index of the column,  $U$  is the overall heat transfer coefficient,  $a$  is the specific geometric area of the packing,  $T^{cw}$  and  $T^L$  are the temperatures of the cooling water and solvent phase,



respectively, and  $Q^{cw}$  is the heat transfer rate through the boundary. Equation 6 calculates the updated voidage by accounting for the volume occupied by the intensified coolers where  $\varepsilon^\circ$  is the standard voidage of the packing being used, and  $\varepsilon$  is the resulting voidage of the process.

$$Q_i^{cw} = y_i U a_i (T_i^{cw} - T_i^L) \quad (5)$$

$$\varepsilon_i = \varepsilon^\circ_i - \varepsilon_i^{cw} \quad (6)$$



**Figure 1.** Configuration of cooling water flow through absorber tower

Modeling the different possible flow configurations of the cooling water through the intensified packing is necessary since different flow configurations can result in a significant difference in heat removal efficiency. Three different flow directions were included in this model: co- and counter-current flow, relative to liquid phase flow, and single pass flow. In this case, the single pass flow is a limiting case for the model since the cooling water enters and exits the column in the same finite element of the model resulting in the best case for heat removal rate. Figure 1 shows how each of these configurations moves cooling water through the column. These flow directions are selected through an integer variable,  $d^{cw}$ , which can take the values of -1, 0, or 1 corresponding to co-current, single pass, and counter-current configurations, respectively. The energy balance of the cooling water is performed by creating lower bound inequality

constraints, which will be active depending on the value of  $d^{cw}$ .

To determine the best placement locations for the internal heat exchanger, an objective function is required. Three objective functions are utilized in this work, the first of these is minimization of CO<sub>2</sub> emissions,  $F_{CO_2}^{V,out}$  (Equation 7). The next two are minimization of column height,  $H$  (Equation 8) and minimization of liquid to gas ratio (Equation 9) in which  $F^{L,in}$  and  $F^{V,in}$  is the molar inlet flow rate of the solvent and gaseous phases, respectively. An additional performance constraint is included with the last two objectives, which sets a lower bound on the allowed capture efficiency. These last two objectives were chosen as each can be used to estimate potential reductions in costs of the system since the absorber column is a large factor in the total capital cost and solvent flow rate is directly proportional to reboiler duty, which is greatest factor in operational costs.

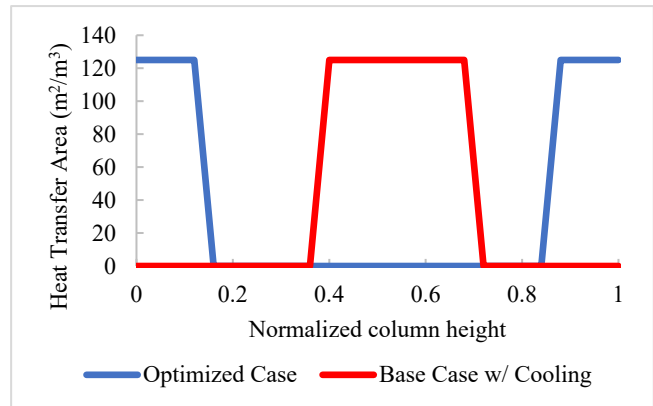
$$\min_{\varepsilon^{cw}, y, d} F_{CO_2}^{V,out} \quad (7)$$

$$\min_{\varepsilon^{cw}, y, d} H \quad (8)$$

$$\min_{\varepsilon^{cw}, y, d} \frac{F^{L,in}}{F^{V,in}} \quad (9)$$

The absorber and internal cooler models were implemented using the IDAES platform, which is built on top of the Pyomo optimization suite [15].

## RESULTS

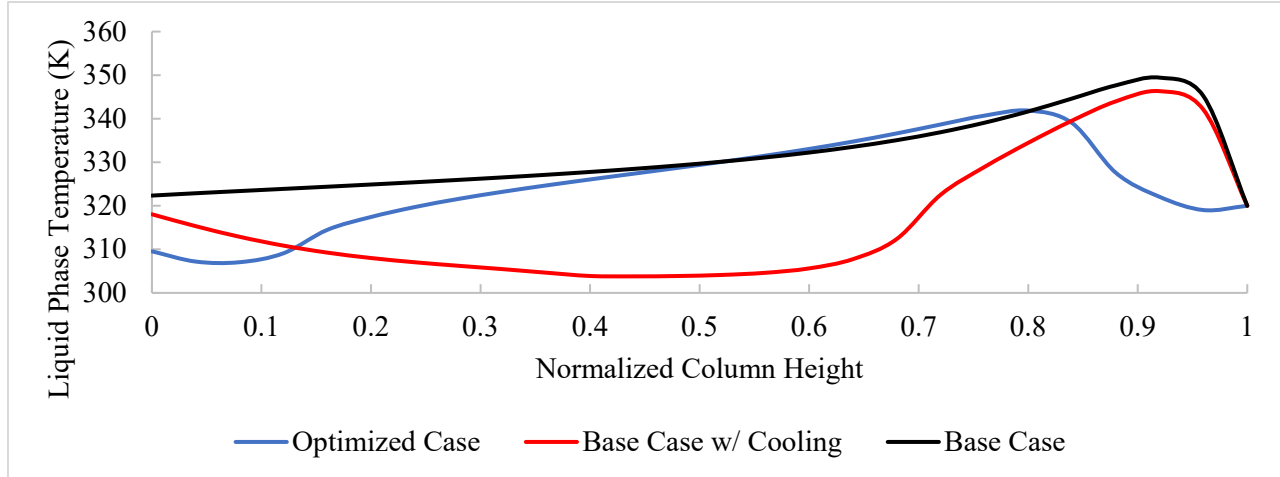


**Figure 2.** Optimal and suboptimal internal heat exchanger placement for minimizing CO<sub>2</sub> emissions.

The base configurations for the column design and process specifications are shown in Table 1. This first set of results utilizes the pilot scale configuration. Using this configuration, the placement of the internal heat exchangers was optimized by minimizing CO<sub>2</sub> emissions. Figure 2 shows the placement and area of the heat exchangers in the optimal solution, which were in the top and bottom 15% of the column. A separate case was also

**Table 1.** Column and process configurations for pilot-scale and process-scale models.

	Height (m)	Diameter (m)	Gas Inlet (mol/s)	CO <sub>2</sub> Gas Conc.	Liquid-Gas Ratio	CO <sub>2</sub> Lean Loading
Case 1 (Pilot Scale)	15	0.65	22	0.12	1.77	0.15
Case 2 (Process Scale)	20	12	12,000	0.042	1.83	0.22

**Figure 3.** Solvent temperature profile for base case, base case with suboptimal cooling, and optimized case.

simulated in which the internal heat exchangers were placed in the center of the column with a similar amount of total available heat transfer area.

The liquid phase temperature profile for the base case without cooling, base case with suboptimal cooling, and optimized cooling are shown in Figure 3, with the capture efficiencies for each case shown in Table 2. The profile base case with out cooling exhibits a much higher average temperature across the length of the column. This higher temperature results in larger thermodynamic limitations to the mass transfer of CO<sub>2</sub>, especially in the upper end of the column where a significant portion of mass transfer is occurring (indicated by the peak in temperature). The base case with suboptimal cooling appears to have a greater extent of heat removal when compared to the optimal case but results in lower capture efficiency. This is due to the given operating conditions. As previously mentioned, the majority of CO<sub>2</sub> mass occurs at the top 10–20% of the column, which is where the peak of the temperature bulge caused by the absorption energy is located in the base case. This causes this section of the column to be heavily thermodynamically limited, as opposed to physical limitations, such as mass transfer area, which explain the optimality of placing an internal cooling element in this region. The temperature profile in the bottom 15% of the column is comparatively level indicating very little mass transfer of CO<sub>2</sub>. In this region, mass transfer is being solely limited by the reduced driving force due to higher loading in the solvent, which results in placement of the intensified packing as a means to

further increase the total amount of heat removed from the system.

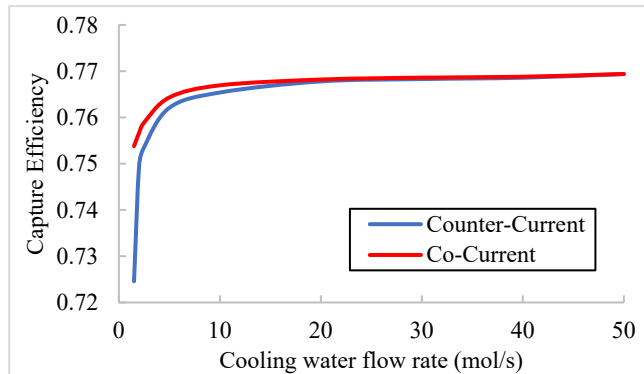
**Table 2.** Comparison of capture efficiency for cases with and without internal cooling

Case	Capture Efficiency
Base Case	72.46%
Base Case w/ Cooling	75.63%
Optimized Case	76.94%

As opposed to the top region of the column, the middle region is mass-transfer limited; therefore, maximizing mass transfer area to obtain high performance is desired. Finally, the bottom section of the column becomes thermodynamically limited; therefore, the optimal configuration is to include the cooling section there.

Using these optimized placements, a study was then conducted that investigated the variance of capture performance in different cooling water flow directions and flow rates. Both counter-current and co-current configurations were simulated using cooling water flow rates between 1.5 and 50 mol/s. The capture efficiencies for each of these configurations are shown in Figure 4. Below 20 mol/s of cooling water, the co-current configuration shows a slight advantage, which is due to how the solvent temperature changes in the upper internal cooler section. Analyzing these temperature profiles shows that the solvent enters the upper heat exchanger at the inlet temperature of 320 K and exits it at 338 K at a column height of 0.85. So, if cooling water is flowing in the

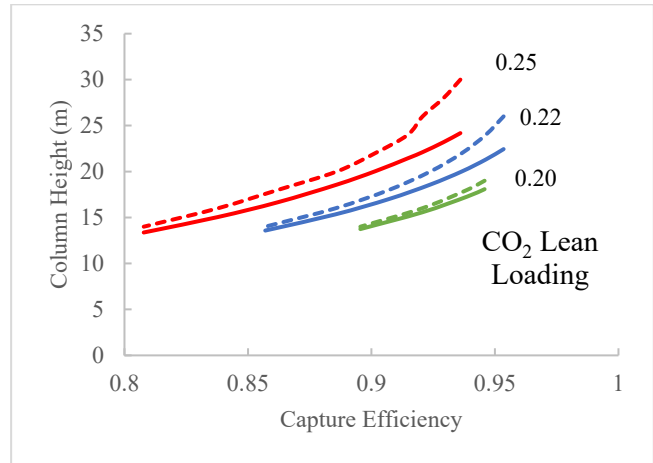
counter-current direction, there is a pinch at the top of the column limiting the heat removal rate. However, once flow rates increase above 20 mol/s, the difference in performance between the two configurations is nearly the same.



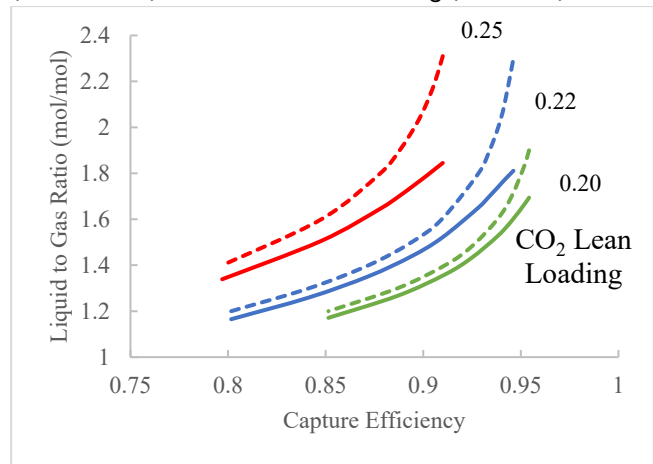
**Figure 4.** Capture efficiency at varying cooling water flow rates and flow direction.

The next results shown are from using the objective functions for minimizing the height and liquid-gas ratios (Equations 8 and 9). These objective functions are useful since each of these values correlates with major costs within the process, in which the column height accounts for an investment cost and the solvent flow correlates with the steam requirement in the reboiler. For these studies, the process scale model configuration was utilized, and each aspect was optimized at different CO<sub>2</sub> lean loadings and at different minimum capture efficiencies. The minimization of the column height in Figure 5 shows a maximum reduction of 6 meters when operating at a lean loading of 0.25 with a capture efficiency of 94%. The minimization of the liquid gas ratio in Figure 6 shows that the solvent rate can be reduced by upwards of 20% at the same lean loading of 0.25 operating at a capture efficiency of 91%.

Results for the column height minimization are shown in Figure 5. This data can be interpreted in two ways. When picking any point along the dotted line representing a column without internal cooling, moving down from this point to the corresponding solid line shows how much the column height can be reduced at a fixed capture efficiency. Alternatively, the horizontal shift shows how the capture efficiency can be increased with optimally placed internal coolers at a fixed column height. The improvement of reduced column height when using internal cooling units increases at higher capture rates and higher lean loading of the solvent. This same trend can be seen in the results for minimizing the liquid-gas ratio (Figure 6).



**Figure 5.** Variation is column height subject to capture efficiency, and lean loading without internal cooling (dashed line) and with internal cooling (solid line).



**Figure 6.** Variation is column height subject to capture efficiency, and lean loading without internal cooling (dashed line) and with internal cooling (solid line).

The reduction in both column height and solvent flow rate with the use of intensified packing has substantial implications on the cost saving for the process. The ability to reduce the height absorber bed, by up to 6 meters, while still retaining equal capture performance suggests that capital costs can be significantly reduced. Similarly, being able to reduce solvent flow rate by up to 20% corresponds with a similar reduction in the operating costs of the steam reboiler in the solvent capture process, which is a major factor in the determination of the levelized cost of capture of CO<sub>2</sub>. Due to the infancy of this technology, an accurate cost model for the intensified packing is still required to allow for a robust economic optimization of the process to determine the trade off of these potential cost savings for the cost of the packing.

There are still other factors of this technology yet that need to be further investigated. The first of which is overall heat transfer coefficient of the packing. In a

dynamic process, this coefficient is likely to vary based on the variance of the loading of flue gas to the column which can influence the heat removal rate. In terms of performance rating, due to the significant reductions found in process conditions through placement optimization and the use of a conservative value for heat transfer coefficient, the intensified packing is likely to retain an advantage over conventional structured packing even with a wide margin of uncertainty in heat transfer performance. Another factor to further consider is the comparison in performance to and absorbent system utilizing intercoolers, which again raises the need for an accurate cost for the intensified packing to determine if the cost of the complexity of the design outweighs the benefit of providing in-line, continuous cooling to the process.

## CONCLUSION

An existing model for solvent-based carbon capture is modified to implement a model for an internal cooling element that is integrated within the structured packing of the column. The internal cooling model was designed so that the placement of these elements can be used as a decision variable in an MINLP problem. The flow direction of the cooling water is also taken into consideration with the possibility of co- and counter-current flows, along with a single pass option.

The results of this work have shown the feasibility of utilizing internal heat exchangers to boost the capture performance of a solvent capture system. Depending on the configuration, the capture efficiency is shown to increase as much as 5%, which can account for a significant reduction in CO<sub>2</sub> emissions being released into the atmosphere. As shown, the optimization for the placement of internal heat exchangers is necessary since a deviation from optimum placement leads to a reduction in capture efficiency. The results also show that when comparing co- and counter-current flows of cooling water through these elements, it is optimal to use the co-current configuration. In the counter-current configuration, a pinch point is created that significantly affects the effectiveness of cooling.

The case studies on optimizing the height and solvent flow rates have shown that significant reductions of up to 20% in each can be made. This reduction is even greater when operating at higher capture efficiencies, which can make this technology ideal for that area of implementation. Although a reduction in the absorber column height can result in significant cost reduction, the real benefit of implementing internal coolers is the reduction in steam costs from reboiler operation. This cost is heavily affected by the solvent flow rate and CO<sub>2</sub> loading of the lean solvent. Thus, the placement of an intensified absorber not only improves the performance of the absorber but the economics of the overall process due to

the resulting effect on the stripper operation.

## ACKNOWLEDGMENTS

The authors graciously acknowledge funding from the U.S. Department of Energy, Office of Fossil Energy and Carbon Management, through the Carbon Capture Program.

## DISCLAIMER

This project was funded by the U.S. Department of Energy, National Energy Technology Laboratory an agency of the United States Government, through a support contract. Neither the United States Government nor any agency thereof, nor any of its employees, nor the support contractor, nor any of their employees, makes any warranty, expressor implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## REFERENCES

1. Meng F, Meng Y, Ju T. Research progress of aqueous amine solution for CO<sub>2</sub> capture: A review. *Renew Sust Energ Rev* (2022)
2. Park J, Yoon S, Oh SY. Improving energy efficiency for a low-temperature CO<sub>2</sub> separation process in natural gas processing. *Energy* (2021)
3. Kim I, Svendsen HF. Heat of Absorption of Carbon Dioxide (CO<sub>2</sub>) in Monoethanolamine (MEA) and 2-(Aminoethyl)ethanolamine (AEEA) Solutions. *Ind Eng Chem Res* 46:5803–5809 (2007)
4. Akula P, Lee A, Eslick J. A modified electrolyte non-random two-liquid model with analytical expression for excess enthalpy: Application to the MEA-H<sub>2</sub>O-CO<sub>2</sub> system. *AIChE J* (2023)
5. Chang H, Shih CM. Simulation and Optimization for Power Plant Flue Gas CO<sub>2</sub> Absorption-Stripping Systems. *Sep Sci Technol* (2007)
6. Plaza JM, Wagener DV, Rochelle GT. Modeling CO<sub>2</sub> capture with aqueous monoethanolamine. *Enrgy Proced* 1:1171–1178 (2009)
7. Karimi M, Hillestad M, Svendsen HF. Investigation of intercooling effect in CO<sub>2</sub> capture energy consumption. *Enrgy Proced* 4:1601–1607 (2011)
8. Bolton S, Kasturi A, Palko S. 3D printed structures for optimized carbon capture technology in packed

bed columns. *Sep Sci Technol* 54:2047–2058 (2019)

9. Miramontes E, Jiang EA, Love LJ. Process intensification of CO<sub>2</sub> absorption using a 3D printed intensified packing device. *AIChE J* (2020)
10. Thompson JA, Tsouris C. Rate-Based Absorption Modeling for Post combustion CO<sub>2</sub> Capture with Additively Manufactured Structured Packing. *Ind Eng Chem Res* 60:14845–14855 (2021)
11. Akula P, Eslick J, Bhattacharyya D, Miller DC. Model Development, Validation, and Optimization of an MEA-Based Post-Combustion CO<sub>2</sub> Capture Process under Part-Load and Variable Capture Operations. *Ind Eng Chem Res* 60:5176–5193 (2021)
12. Gaspar J, Fosbøl PL. A general enhancement factor model for absorption and desorption systems: A CO<sub>2</sub> capture case-study. *Chem Eng Sci* 138:203–215 (2015)
13. Kohl AL, Nielsen RB. Chapter 2 - Alkanolamines for Hydrogen Sulfide and Carbon Dioxide Removal. In: Gas Purification Ed: 5. Gulf Professional Publishing (1997)
14. Que H, Chen C-C. Thermodynamic Modeling of the NH<sub>3</sub>–CO<sub>2</sub>–H<sub>2</sub>O System with Electrolyte NRTL Model. *Ind Eng Chem Res* 50:11406–11421 (2011)
15. Lee A, Ghouse J, Eslick J, Laird C, Siirola J, Zamarripa M, Gunter D, Shinn J, Dowling A, Bhattacharyya D, Biegler L, Burgard A, Miller D. The IDAES Process Modeling Framework and Model Library – Flexibility for Process Simulation, Optimization and Control. *J Adv Manuf Process* (2021)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Design of Plastic Waste Chemical Recycling Process Considering Uncertainty

Zhifei Yuliu<sup>a</sup>, Yuqing Luo<sup>a</sup>, and Marianthi Ierapetritou<sup>a\*</sup>

<sup>a</sup> University of Delaware, Department of Chemical and Biomolecular Engineering, Newark, DE, USA

\* Corresponding Author: [mgi@udel.edu](mailto:mgi@udel.edu).

---

## ABSTRACT

Chemical recycling of plastics is a promising technology to reduce carbon footprint and ease the pressure of waste treatment. Specifically, highly efficient conversion technologies for polyolefins will be the most effective solution to address the plastic waste crisis, given that polyolefins are the primary contributors to global plastic production. Significant challenges encountered by plastic waste valorization facilities include the uncertainty in the composition of the waste feedstock, process yield, and product price. These variabilities can lead to compromised performance or even render operations infeasible. To address these challenges, this work applied the robust optimization-based framework to design an integrated polyolefin chemical recycling plant. Data-driven surrogate model was built to capture the separation units' behavior and reduce the computational complexity of the optimization problem. It was found that when process yield and price uncertainties were considered, wax products became more favorable, and pyrolysis became the preferred reaction technology.

**Keywords:** Process Design, Design Under Uncertainty, Optimization, Polymers, Technoeconomic Analysis, Plastic Waste

## INTRODUCTION

Global plastic waste has been on the rise, making efficient plastic recycling process design imperative [1]. Chemical recycling and upcycling strategies not only reduce the mismanaged plastic waste, but also has the potential to reduce the carbon footprint to meet sustainable goals [2].

One challenge in plastic recycling process design arises from uncertain feedstock compositions. The type and proportion of plastic waste can exhibit variations influenced by factors like geographic location, resulting in substantial differences in both economic and energy values [3]. In addition, feedstock compositions affect the strategies for plastic recycling. For instance, one advantage of the pyrolysis process is that it can easily handle a mixed plastic waste feedstock with different ratio, especially noncatalytic pyrolysis unit. However, most of other chemical recycling technologies, including hydrogenolysis, typically requires relatively pure feedstock after careful sorting or impurity removal to ensure good catalyst performance [4,5].

Polyolefins (PO), including polypropylene (PP) and polyethylene (PE), are main source of plastic waste. Their inert carbon-carbon backbones make it challenging to breakdown the long chains and produce valuable products [6]. Many reaction pathways have been developed recently to effectively deconstruct PO, among which pyrolysis and hydroconversion (i.e., hydrocracking and hydrogenolysis) have shown promising potentials. Thermal pyrolysis reactions operate at elevated temperature and shorter residence time, which generates products a wider distribution and more gas that are most useful as fuels [7]. Hydroconversion, on the other hand, operates at milder conditions and produces liquid hydrocarbon within the fuel or lubricant ranges [6,8].

Depending on the feedstock composition, different technologies operate at different conditions to produce different products [3,9]. As the product selectivity and the use of catalyst depend on the plastic waste type [3], feedstock variability could affect not only the process performance but also its feasibility.

Existing studies in process design of plastic recycling typically focus on a particular recycling strategy and

its operating conditions [9]. For instance, Hernandez *et al.* compared the costs and emissions for four waste LDPE treatment processes – gasification, pyrolysis, hydrogenolysis [4]. Bora *et al.* performed life cycle assessment and techno-economic analysis on waste PP treatment processes and demonstrated that chemical recycling had low emissions but only profitable at large scales [10]. Zhao and You utilized the superstructure framework to optimize the net present value and greenhouse gas emissions of monomers, aromatic mixtures, and fuels production from waste HDPE [11].

An integrated plastic waste recycling technology selection and product separation provide opportunities for performance improvement. Moreover, it is important to consider the feedstock variability when designing such integrated chemical recycling facilities to ensure feasible operations among each connecting subunit. Robust optimization has been established as a computationally efficient framework to incorporate uncertainties and improves process performance [12]. Li *et al.* applied robust optimization to refinery production planning problem considering yield, cost, and price uncertainties [13].

This study proposes a methodology for the development of a chemical recycling facility for plastic waste. The design involves the selection of reaction technology and product separation guided by an optimization model. Since feedstock composition variability and product yield distribution are unavoidable in waste plastic treatment and affects the separation efficiencies, it is vital to guarantee feasible operation and good performance under these uncertainties. Rigorous process flowsheet simulations in Aspen Plus (Aspen Technology) [14] were carried out to obtain surrogate models of separation processes. Design decisions include chemical treatment technologies, distillation column design and unit connectivity. A robust optimization model is formulated to maximize profit under the worst case and ensure process feasibility (e.g., normal process operation) for all scenarios [6]. This robust optimization model will improve the chemical recycling process feasibility and performance under the worst uncertain case than the traditional deterministic optimization model [2,13]. Applying robust optimization instead of stochastic programming will also largely reduce the computational complexity, especially in the process design problem with high dimensionality arising from feedstock variability feedstock variability, yield uncertainty, and price fluctuation [15,16].

## MODEL FORMULATION

### Deterministic Superstructure Optimization

The superstructure elements for the chemical recycling facility includes plastic waste, other feedstocks, products, reaction technologies, and separation alternatives (Figure 1). In this study, decisions are made on three

levels. First, the combination of reaction or separation technologies is selected. Second, we determine the connectivity among feed, technologies, and products. Third, we decide the exact realization of a technology (e.g., distillation column operation conditions) by choosing an option. On this third level, incompatible connections, such as a liquid/solid stream entering a gas separator or an inappropriate reactant used by a particular reactor, could occur. Consequently, the connectivity of the superstructure elements is sorted *a priori* to eliminate those unproductive links. This step reduces the overall model size without cutting off potential candidate solutions [17].

As illustrated in Figure 2, the superstructure is connected with the inlet mixer and outlet splitters. The superstructure mass balance is established with the equations (1-4) for flow rate of each process stream going from  $i'$  to  $i$  for species  $k$  ( $F_{i',i,k}^I$ ). Equation (1) specifies the mixer balance for  $F_{i,k}^I$ , the flow rate of an inlet stream into a superstructure element  $i$  for species  $k$ . The splitter balance is enforced in equation (2) for the outlet stream of a superstructure element  $F_{i,k}^O$ . While a splitter with split ratio to multiple outlet stream is possible, we choose to maintain the linearity of the problem by allowing exactly one destination  $i'$  for each superstructure element  $i$  with the binary variable  $\eta_{i,i'}$ . To ensure the feasibility of each unit, a capacity limit is imposed as shown in (5-7). The total flow rate  $F_i^T$  is decided from (5), and the technology capacity  $F_i^{CAP}$  is enforced with (6). A big M constraint (7) is used to ensure if a technology is not selected ( $y_i = 0$ ), the capacity is 0.

$$F_{i,k}^I = \sum_{i' \in \text{inlets}(i)} F_{i',i,k}^I \quad \forall i \in I^{TECH} \forall k \quad (1)$$

$$F_{i,k}^O = \sum_{i' \in \text{outlets}(i)} F_{i,i',k}^O \quad \forall i \in I^{TECH} \forall k \quad (2)$$

$$\sum_k F_{i,i',k}^O \leq M \cdot \eta_{i,i'} \quad \forall i \in I^{TECH} \forall i' \in \text{outlets}(i) \quad (3)$$

$$\sum_{i' \in \text{outlets}(i)} \eta_{i,i'} = 1 \quad \forall i \in I^{TECH} \quad (4)$$

$$F_i^T = \sum_{k \in K} F_{i,k}^I \quad \forall i \in I^{TECH} \quad (5)$$

$$F_i^T \leq F_i^{CAP} \quad \forall i \in I^{TECH} \quad (6)$$

$$F_i^{CAP} \leq M \cdot y_i \quad \forall i \in I^{TECH} \quad (7)$$

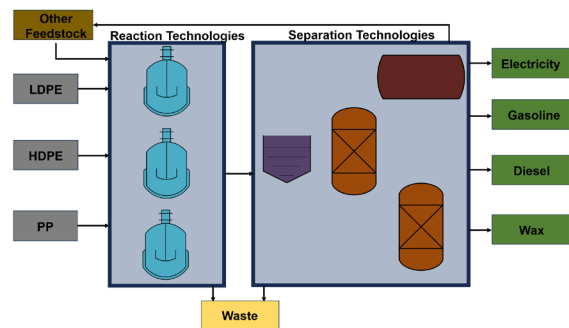
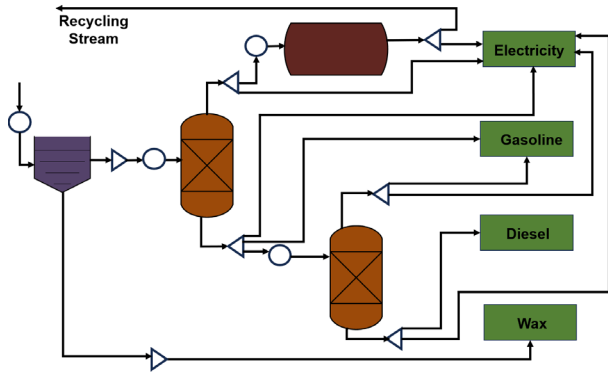


Figure 1. Elements of the chemical recycling plant superstructure.



**Figure 2.** Possible connectivity of chemical recycling plant's separation units

The reaction conversion and selectivity of the chemical recycling technologies are taken from the literature to model stoichiometric reactor units in the superstructure. A basis in the reaction feed stream  $F_i^{Ref}$  and a conversion coefficient  $\epsilon_{i,k}$  is used to determine the composition of the product streams as shown in (8). For some reactions, especially the hydroconversions [6,8], some undesired solid (e.g., coke) and gas formation are not well-characterized, leading to inaccurate estimation product yields thus violations of mass balance. To close the mass balance gap, we make a conservative assumption by including a waste stream that is not usable in the downstream operations (9). Admittedly, this assumption may not reflect the actual reaction. Thus, the effects of wax yield uncertainty are addressed in the robust optimization.

$$F_{i,k}^l + \epsilon_{i,k} \cdot F_i^{Ref} = F_{i,k}^p \quad \forall i \in I^{RXN} \quad \forall k \in Products(i) \quad (8)$$

$$\sum_k F_{i,k}^l - \sum_{k \neq WASTE} F_{i,k}^p = F_{i,WASTE}^p, \quad \forall i \in I^{RXN} \quad (9)$$

All separation units in this study are designed to achieve sharp separation (i.e., separators always isolate nearly all the light species to the light outlet stream  $F_{i,k}^{OL}$  and the heavy species to the heavy outlet stream  $F_{i,k}^{OH}$ ). The gas and fuel range products are separated in distillation columns with a sequence based on the target boiling points. To ensure that sharp separation is attainable with our design, surrogate models (12) are built to estimate the reflux ratio needed for 99% purity and the associated utilities from rigorous Aspen Plus process simulation. In practice, the distillation columns cannot deviate too much from the nominal reflux ratio once designed and installed. To ensure the operating feasibility, different design options are provided for the same distillation technology but at most one will be active as shown in (13). A 30% flexibility around the reflux ratio at nominal conditions is allowed for each design option  $p$  as presented in (13)-(15). When an option is selected ( $w_{i,p} = 1$ ), two big M constraints are used for to ensure the reflux ratio does

not deviate from the design value for more than 30% (14,15).

$$F_{i,k}^l = F_{i,k}^{OL} \quad \forall i \in I^{SEP} \quad \forall k \in light(i) \quad (10)$$

$$F_{i,k}^l = F_{i,k}^{OH} \quad \forall i \in I^{SEP} \quad \forall k \in heavy(i) \quad (11)$$

$$f_i(Light, Heavy) = (RR_i, Q_i^{cool}, Q_i^{heat}) \quad \forall i \in I^{SEP} \quad (12)$$

$$y_i = \sum_p w_{i,p} \quad \forall i \in I^{SEP} \quad (13)$$

$$-M(1 - w_{i,p}) + 0.7RR_p^{nom} \leq RR_i \quad \forall i \in I^{SEP} \quad (14)$$

$$RR_i \leq M(1 - w_{i,p}) + 1.3RR_p^{nom} \quad \forall i \in I^{SEP} \quad (15)$$

A supply constraint is imposed to the plastic waste feedstock in (16).

$$\sum_{i' \in outlets(i)} F_{i,i',k} \leq F_{j,k}^{SUP} \quad \forall i \in I^{Plastic} \quad \forall k \in plastic \quad (16)$$

For fuel range products, we impose a maximum for the olefin content [17] as a product requirement (17). In addition, we specify the limit of product quantities below or above the typical boiling point ranges for fuel product such as gasoline and diesel (18,19). Equations (18,19) represent a linear simplification of the blending rule by only considering boiling point and olefin content, although more rigorous blending rule can be incorporated.

$$\sum_{k \in K^{olefin}} F_{i,k}^{prod} \leq \phi^{olefin} \sum_{k' \in K} F_{i,k'}^{prod} \quad \forall i \in I^{Fuel} \quad (17)$$

$$\sum_{k \in K^{BPLO}} F_{i,k}^{prod} \leq \phi_k^{klo} \quad i \in I^{Fuel} \quad (18)$$

$$\sum_{k \in K^{BPUP}} F_{i,k}^{prod} \geq \phi_k^{kup} \quad i \in I^{Fuel} \quad (19)$$

The general deterministic formulation of this chemical recycling process design is presented as a mixed-integer linear programming (MILP) problem in equation (20):

$$(20) \quad \max (R^{PROD} + R^{ELEC} - C^{CAPEX} - C^{OPEX})$$

s. t. Superstructure Mass Balance  
Technology Constraints  
Product Rules  
Plastic Waste Feed Constraints  
Process Economics

## Surrogate Model for Distillation Columns

The distillation column surrogate model is an essential part of the superstructure optimization to create a mapping between the inlet flow rates for both the light and heavy components and the actual reflux ratio, condenser duty, and reboiler duty. Artificial neural network (ANN) is selected as surrogate models given its excellent fitting performance and its ability to maintain the linearity of the problem. We refer to our previous work and use Rectified Linear Unit (ReLU) as the activation function with a feed-forward ANN [18]. The predicted value  $a_m^k$  of layer  $k$  and node  $m$  is calculated using a linear combination of the values from the previous layer as shown in

(21). The activation function ReLU,  $z_n^k = \max(0, a_n^k)$ , is modelled using the big M constraints as shown in (22a-22d)

$$a_m^k = W_{n,m}^{k-1} z_n^{k-1} + b_m^{k-1} \quad (21)$$

$$-M \cdot (1 - \delta_m^k) \leq a_m^k \quad (22a)$$

$$a_m^k \leq M \cdot \delta_m^k \quad (22b)$$

$$0 \leq z_m^k \leq M \cdot \delta_m^k \quad (22c)$$

$$a_m^k \cdot (1 - \delta_m^k) \leq z_m^k \leq a_m^k + M \cdot (1 - \delta_m^k) \quad (22d)$$

## Robust Optimization

The uncertainty in this study arises from three sources. The first source is the dynamic supply of sorted bale plastic waste, which leads to uncertain feedstock flowrates. The second source arises from the product yield. For instance, not all solid products of hydroconversion of recycled plastic could be treated as waxes [6]. Instead, depending on the reaction condition, some solid plastic waste could remain unreacted or form coke. To reflect this, we incorporate a process yield uncertainty for waxes in hydrocracking and hydrogenolysis. Furthermore, the product selling price uncertainty is also included to reflect the price volatility of the petroleum refinery products. The fuel range product tends to be more volatile than waxes [19]. Therefore, a larger deviation should be considered.

To formulate the robust counterpart of the problem, we refer to the work of Li. *et al.* [13], in which the robust counterpart formulations have been derived for linear and mixed-integer linear programming with different uncertainty set. In this study, we consider box uncertainty set for all uncertain parameters. The robust counterpart for an inequality constraint (23) when considering left-hand-side (L.H.S.) uncertainty is presented in (24), and right-hand-side (R.H.S.) uncertainty in (25).

$$\sum_j a_{i,j} x_j \leq b_i \quad (23)$$

$$\sum_j a_{i,j} x_j + \sum_j \widehat{a}_{i,j} |x_j| \leq b_i \quad (24)$$

$$\sum_j a_{i,j} x_j + \widehat{b}_i \leq b_i \quad (25)$$

In this study, uncertainty parameters for process yield and product sale price occurs in equations instead of inequalities, which drastically restricts the feasible space of mathematical model and often causes infeasibility [16]. Therefore, we use the inequality constraints instead for process yield and sale price as presented in previous work in petroleum refinery by Leiras *et al.* [15] as shown in (26,27). Using this upper bound on wax produced and product sales, we now have L.H.S. uncertainty instead. While we acknowledge this reformulation creates a relaxation of the original problem mathematically, the profit objective typically encourages producing more products when possible. Therefore, these reformulated inequalities (26,27) are likely to be active at optimal

solution. The robust formulation is obtained by applying the properties of box uncertainty set as presented in (24,25).

$$F_{i,WAX}^{FEED} + \epsilon_{i,WAX} \cdot F_i^{Ref} \geq F_{i,WAX}^{PROD} \quad \forall i \in I^{RXN} \quad (26)$$

$$Revenue^{PROD} \leq Price \cdot F^{PROD} \quad (27)$$

## ILLUSTRATIVE EXAMPLE

In this example, we considered three categories of reaction technologies (i.e., hydrocracking [6], hydrogenolysis [8], and pyrolysis [7]), three types of plastic waste streams (i.e., PP, LDPE, and HDPE), one solid separation, two distillation (each with five options), and a pressure-swing absorption unit for hydrogen recovery. The details for the technologies are summarized in Table 1-3. For reaction technologies that allow mixture plastic waste, we enforced the feed composition to be the same as reported from experiments in the literature. Gasoline, diesel, and waxes are selected as the product. Additionally, another stream for electricity generation is available for mixture outlet streams that do not meet the product requirement for fuels.

**Table 1:** Reactions considered

Code	Reaction Type	Feed
R1A	hydrocracking	LDPE
R2A	hydrocracking	HDPE
R3A	hydrocracking	PP
R1B	hydrogenolysis	PP
R2B	hydrogenolysis	PP/PE(1:1)
R1C	pyrolysis	PE
R2C	pyrolysis	PE/PP(3:2)
R3C	pyrolysis	PE/PP(2:3)
R4C	pyrolysis	PP

**Table 2:** Separations considered

Code	Separator Type	Separation Task
FL1A	solid/fluid	fluid/wax
DT1A	distillation	C4/C5
DT2A	distillation	C12/C13
PSA1A	pressure swing absorption	hydrogen/alkanes

**Table 3:** Reflux ratios at nominal operating conditions for different options

Option	1	2	3	4	5
Reflux Ratio(molar)	1	1.83	3.43	6.37	11.83

The data for the surrogate model training for distillation columns were obtained by running sampled simulations in Aspen Plus [14] with Python interface. The

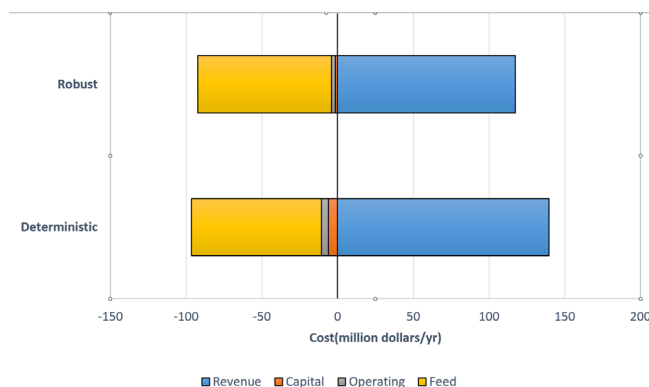
model was trained in PyTorch [20]. The ANN models for both distillation columns had 4 layers with 15, 18, 12 neurons in the hidden layers and 3 neurons at the output layer. The  $R^2$  of the ANN for DT1A was 0.9586 and for DT2A, 0.9675.

The operational basis of this example was based on the previous techno-economic analysis [4,11] with a supply of 12 tonne/hr PP, 10 tonne/hr HDPE, and 2.5 tonne/hr LDPE. The cost for the plastic collection and sorting depends on the population density and geographic variation [9]. The plastic waste collection and sorting cost was estimated in the work of Hernandez *et al.* [4] to be in the range of 250 to 700 \$/tonne. In this example, we assumed an overall cost for the sorted plastic to be \$500/tonne. The price for waxes was estimated to be \$1000/tonne [21]. Moreover, the fuel price at nominal condition was estimated to be \$960/tonne for gasoline [19] and \$920/tonne for diesel [19]. A wax yield of 0.5 was assumed for hydroconversions. The associated uncertainty type and deviations from the nominal value considered in this example is shown in Table 4.

**Table 4:** Uncertain parameters for illustrative example

Uncertain Parameter	Type	Deviation (%)
Plastic supply	RHS	10
Wax yield	LHS	40
Fuel price	LHS	30
Wax price	LHS	10

All optimization models were implemented in Pyomo [22] and solved with Cplex 22.1 solver [23] on a computer with Intel Xeon E-2274G CPU @ 4.00GHz 32 GB RAM. The deterministic model was solved under nominal condition, and the robust formulation was applied to include the uncertainty set in Table 4.

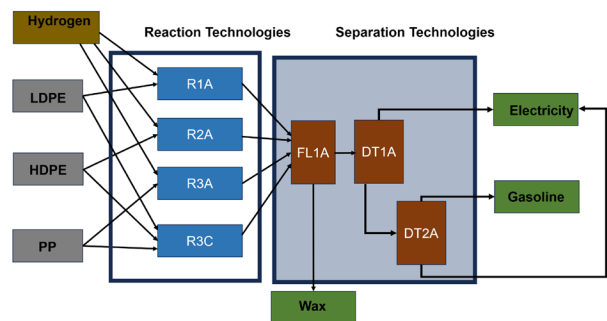


**Figure 3.** Cost breakdown of the plastic chemical recycling system.

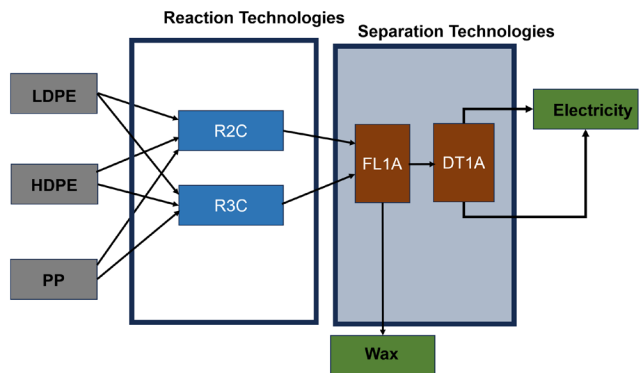
As shown in Figure 3, under nominal condition, the optimal integrated chemical recycling process has a

revenue of 139.58 million dollars/year and a total cost of 96.40 million dollars/year. The feedstock cost comprises of 89% of the total cost, which accounts for all upstream cost for plastic waste collection and sorting. The profit of the process is 43.18 million dollars/year. However, when the uncertainties are considered in the robust formulation, the total revenue of the best design dropped to 117.45 million dollars/year with a total cost of 92.5 million dollars/year. This leads to a profit of 25.95 million dollars/year.

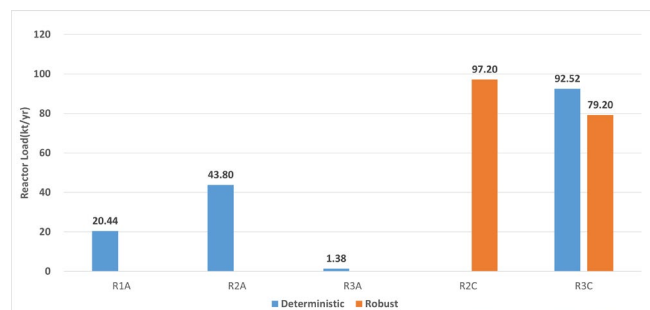
Under the nominal condition, the distillation column DT1A operates at a 6.64 molar reflux ratio, and operating condition option 4 is selected. For distillation column DT2A, the actual molar reflux ratio is 2.24, and option 2 is selected. When robust optimization is performed, the actual reflux ratio for DT1A becomes 0.07, and option 1 is selected. The technology DT2A is not selected at all as no gasoline product is pursued as shown in Figure 5.



**Figure 4.** Superstructure under the nominal condition

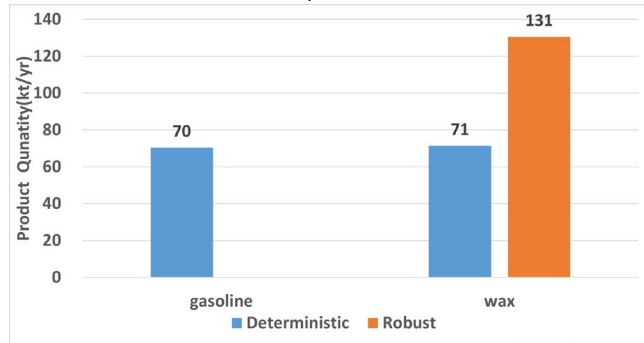


**Figure 5.** Superstructure under the robust formulation





**Figure 6.** Reactor loads for process designs from both deterministic and robust optimization



**Figure 7.** Product quantities for process designs from both deterministic and robust optimization.

As illustrated in Figure 4, the reaction of a combination of hydrocracking (R1A, R2A, R3A) and mixed olefin pyrolysis (R3C) is selected in the nominal case, and the reactor load is presented in Figure 6. When uncertainties are considered, only mixed olefin pyrolysis is chosen (R2C, R3C). The product produced changes from a combination of gasoline and wax to only wax. This change occurs primarily as a result of the price volatility of fuel products, which makes the wax product relatively more profitable than the nominal condition. In addition, the hydrocracking reactions are less economically favorable after the uncertainty in the wax yield of hydroconversions is included.

## CONCLUSIONS

In this work, we have proposed the robust optimization framework for designing plastic waste valorization system. We used an illustrative example to demonstrate the how the feedstock supply availability, process yield, and product price uncertainties affect the optimal process design and operation. In the future, this model could be extended to include more considerations including more types of products, upgrading technologies, upstream sorting process, reaction technologies, and more kinds of plastic waste.

## ACKNOWLEDGEMENT

This work is financially supported by the National Science Foundation under Grant No. 2134471.

## REFERENCES

- Geyer R, Jambeck JR, Law KL. Production, use, and fate of all plastics ever made. *Sci. Adv.*;3(7) (2017).
- Zhao X, You F. Consequential life cycle assessment and optimization of high-density polyethylene plastic waste chemical recycling. *ACS Sustain. Chem. Eng.*;9(36):12167–84 (2021).

- Li H, Aguirre-Villegas HA, Allen RD et al. Expanding plastics recycling technologies: chemical aspects, technology status and challenges. *Green Chem.* 24(23):8899–9002 (2022).
- Hernández B, Kots P, Selvam E et al. Techno-Economic and life cycle analyses of thermochemical upcycling technologies of low-density polyethylene waste. *ACS Sustain. Chem. Eng.* 11 (18), 7170–7181 (2023).
- Luo Y, Selvam E, Vlachos DG et al. Economic and Environmental Benefits of Modular Microwave-Assisted Polyethylene Terephthalate Depolymerization. *ACS Sustain. Chem.* 11 (10), 4209–4218 (2023).
- Rorrer JE, Troyano-Valls C, Beckham GT et al. Hydrogenolysis of polypropylene and mixed polyolefin plastic waste over RU/C to produce liquid alkanes. *ACS Sustain. Chem.* 9(35):11661–6 (2021).
- Predel M, Kaminsky W. Pyrolysis of mixed polyolefins in a fluidised-bed reactor and on a pyro-GC/MS to yield aliphatic waxes. *Polym. Degrad.* 70(3):373–85. (2000)
- Liu S, Kots PA, Vance BC et al. Plastic waste to fuels by hydrocracking at mild conditions. *Sci. Adv.* 7(17) (2021)
- Milbrandt A, Coney K, Badgett A et al. Quantification and evaluation of plastic waste in the United States. *Resour Conserv Recycl.* 183:106363(2022).
- Bora RR, Wang R, You F. Waste polypropylene plastic recycling toward climate change mitigation and Circular Economy: Energy, environmental, and Technoeconomic Perspectives. *ACS Sustain. Chem.* 8(43):16350–63. (2020)
- Zhao, X.; You, F., Waste high-density polyethylene recycling process systems for mitigating plastic pollution through a sustainable design and synthesis paradigm. *AIChE J.* 67 (4). (2021).
- Ben-Tal A, Goryashko A, Guslitzer E et al. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2), 351–376. (2004)
- Li Z, Ding R, Floudas CA. A comparative theoretical and computational study on robust counterpart optimization: I. Robust linear optimization and robust mixed integer linear optimization. *Ind. Eng. Chem. Res.* 50(18):10567–603. (2011)
- Haydary J. Chemical Process Design and Simulation: Aspen Plus and Aspen Hysys Applications. Hoboken, NJ: American Institute of Chemical Engineers (2019).
- Leiras A, Hamacher S, Elkamel A. Petroleum refinery operational planning using robust optimization. *Eng. Optim.* 42(12):1119–31. (2010)

16. Gorissen BL, Yanıkođlu İ, den Hertog D. A practical guide to robust optimization. *Omega*. 53:124–37. (2015)
17. Restrepo-Flórez JM, Maravelias CT. Advanced fuels from ethanol – a superstructure optimization approach. *Energy Environ. Sci*. 14(1):493–506. (2021)
18. Luo Y, Ierapetritou M. Multifeedstock and multiproduct process design using neural network surrogate flexibility constraints. *Ind. Eng. Chem. Res*. 62(5):2067–79. (2023)
19. <https://businessanalytiq.com/procurementanalytics/index>
20. Imambi S, Prakash KB, Kanagachidambaresan GR. Pytorch. *Prog. TensorFlow*. 87–104. (2021)
21. <https://www.petronaftco.com/paraffin-wax-price/>
22. Hart WE. Pyomo-- optimization modeling in python. New York: Springer. (2012)
23. Bussieck MR, Meeraus A. General algebraic modeling system (GAMS). *Appl. Optim*. 137–57. (2004)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Part 1

## Peer-Reviewed Articles

Section 3: Design and Emerging Fields



# A GRASP Heuristic for Solving an Acquisition Function Embedded in a Parallel Bayesian Optimization Framework

R. Cory Allen<sup>a\*</sup>, Youngdae Kim<sup>b</sup>, and Dimitri J. Papageorgiou<sup>a</sup>

<sup>a</sup> ExxonMobil Technology and Engineering Company, Modeling Optimization and Data Science, Spring, TX, USA

<sup>b</sup> ExxonMobil Technology and Engineering Company, Research-Energy Sciences, Annandale, NJ, USA

\* Corresponding Author: [richard.c.allen@exxonmobil.com](mailto:richard.c.allen@exxonmobil.com)

---

## ABSTRACT

Design problems for process systems engineering applications often require multi-scale modeling integrating detailed process models. Consequently, black-box optimization and surrogate modeling have continued to play a fundamental role in mission-critical design applications. Inherent in surrogate modeling applications, particularly those constrained by “expensive” function evaluations, are the questions of how to properly balance “exploration” and “exploitation” and how to do so while harnessing parallel computing in techniques. We devise and investigate a one-step look-ahead GRASP heuristic for balancing exploration and exploitation in a parallel environment. Computational results reveal that our approach can yield equivalent or superior surrogate quality with near linear scaling in the number of parallel samples.

**Keywords:** Optimization, Parallelization, Surrogate Model, Derivative Free Optimization, Machine Learning

---

## INTRODUCTION

In recent years, there has been an increased utilization of high-fidelity digital twins (HFDTs) in the design, development, and/or operation of complex systems and enterprise-wide supply chains. These digital twins typically originate from data-driven models, physics-based models, or some combination thereof. Depending on their physical location, which can lead to data latency issues, scale, underlying structure, evaluating them can be computationally expensive given a set of static and/or dynamic input parameters and they tend to come with only zeroth order information. This can be problematic for practitioners who would like to utilize them directly or within larger workflows to optimize and/or conduct sensitivity analyses on the system or supply chain in which they were originally intended to simulate [1].

These challenges have led practitioners to utilize black-box optimization (also known as derivative-free optimization) methods and/or surrogate modeling techniques to achieve their desired goals. Excellent overviews and reviews of derivative free optimization methods and surrogate modeling techniques are presented in [2], [3], [4], and references therein. Even with the recent

advances in the literature, there are still many open questions, such as how to properly balance “exploration” and “exploitation” and how to harness parallel computing in techniques, such as Bayesian optimization (BO), that were originally designed to be executed serially [5].

To this end, we present a BO procedure to allow the practitioner to be judiciously sample the design space in parallel. At the heart of the procedure lies a non-convex acquisition function that generates a batch of  $B$  different points that are intended to be evaluated by the HFDT parallel. Due to the difficulty in solving the non-convex acquisition function, we have created a greedy randomized adaptive search procedure (GRASP) that is quickly able to generate quality solutions via a construction heuristic and then subsequently improve on these initial solutions via a local search procedure – for a detailed treatise on GRASPs, specifically for problems with a continuous design space, please see [6] and references therein.

The remainder of this work is organized as follows: i) first, we present a brief literature review on BO; ii) next, we present the problem statement; iii) we then present our workflow; iv) next, we conduct computational experiments on 6 standard benchmark test functions to illustrate the effectiveness of our workflow; and v) finally we offer conclusions and future research directions.

## LITERATURE REVIEW

Several research groups have attempted to accelerate Bayesian optimization for the setting known as “batch Bayesian optimization,” i.e., when a “batch” of costly-to-evaluate function calls can be performed in parallel. As noted in one of the earliest works on batch BO [7], sequential BO (i.e., the traditional approach in which a single function evaluation is made per iteration) has a fundamental advantage over batch BO because, when applying the former, each function evaluation is immediately used to obtain a more accurate posterior. This improved posterior should, in theory, lead to a better, or more information-rich, sequence of function evaluations that accelerate convergence to a global optimum. Thus, researchers have attempted to design batch BO methods to compete with one-sample-per-iteration approaches in terms of information gain per sample.

One key computational challenge in batch BO is to find a jointly optimized batch of inputs. This challenge is mainly because of the lack of analytical formula and differentiability of the underlying selection criteria, i.e., expected improvement, for a batch of size greater than 2. To resolve this issue, Daxberger and Low [8] noted that there are two extremes to designing batch BO methods. On one end of the spectrum, greedy batch BO algorithms, e.g., [7], rely on the current (and therefore a single) posterior distribution to select the batch of points to sample. This single-step lookahead approach eases scalability, but each prospective sample in the batch does not account for the potential impact of the other prospective samples. On the other end of the spectrum, one can attempt a multi-step lookahead batch selection by considering the joint impact of all prospective samples, as is done in parallel predictive entropy search [9], and the parallel knowledge gradient method [10]. The downside is that these approaches tend to scale poorly in the batch size. To mitigate the scalability issue, [11] developed a stochastic gradient method for multipoints expected improvement (q-EI) computation, however, it requires many Monte Carlo samples for estimating a stochastic gradient and q-EI. These could be much more expensive than evaluating a single point EI (with  $q=1$ ) using its closed form solution.

De Ath et al. [12] introduced an epsilon-shotgun method that “leverages the model’s prediction, uncertainty, and the approximated rate of change of the landscape to determine the spread of batch solutions to be distributed around a putative location”. The initial target location is selected either in an exploitative fashion on the mean prediction, or - with probability  $\epsilon$  - from elsewhere in the design space. This results in locations that are more densely sampled in regions where the function is changing rapidly and in locations predicted to be good,

i.e., close to predicted optima, with more scattered samples in regions where the function is flatter and/or of poorer quality.” Kandasamy et al. [13] analyzed a randomized selection procedure known as Thompson sampling, which selects a point to evaluate by maximizing over a random sample from the posterior distribution and find that “making  $n$  evaluations distributed among  $M$  workers is essentially equivalent to performing  $n$  evaluations in sequence” [13]. Thompson sampling, however, is less common/popular than popular algorithms that choose evaluation points based on expected improvement (EI) or upper confidence bounds (UCB). Wang et al. survey recent advances in BO [14].

As the algorithm’s name suggests, Bayesian optimization is primarily, if not entirely, dedicated to the task of optimizing an unknown function. It is also of interest to generate a high-fidelity surrogate, which is locally or globally valid, for design applications. One can control the degree of exploration in BO by modifying the acquisition function. Chen et al. [15] after choosing multiple acquisition functions, identify Pareto-optimal solutions (based on a so-called minimum-diverse-exploitative strategy) of a multi-objective optimization problem to select candidate sample points.

## PROBLEM STATEMENT

Consider a decision maker (DM) who would like to utilize a HFDT to design, develop, and/or operate a complex system and/or enterprise-wide supply chain. The relationship between the inputs and the scalar output of the HFDT is given by  $f: \mathcal{X} \mapsto \mathbb{R}^1$ , where  $\mathcal{X}$  is a compact subset of  $\mathbb{R}^N$ . The scalar output,  $y$ , of the computationally intensive function can be noisy, consequently it is assumed that  $y = f(x) + \epsilon$ , where  $x \in \mathcal{X}$  and  $\epsilon \sim N(0, \sigma^2)$  i.i.d.

The DM would like to utilize BO, with a batch size of  $B$ , to optimize and/or create a surrogate model of the HFDT. It is assumed that the DM has access to  $B$  different machines, such that each sample point,  $b \in \mathcal{B}$ , in the batch,  $\bar{x}_{k,b} \in \mathcal{X}$ , can be evaluated by the HFDT in parallel at each major iteration,  $k \in \mathcal{K}$ , of the BO procedure, where  $\mathcal{B} := \{1 \dots B\}$  and  $\mathcal{K} := \{1 \dots K\}$ . Therefore, the goal of the DM is to ensure that the points generated by the acquisition function are judiciously selected, such that whether “exploration” or “exploitation” is preferred, the points will provide valuable “learnings”.

## METHODOLOGY

In this section, we present the BO procedure, which allows the practitioner to judiciously sample the design space in parallel. We also present a novel GRASP, that generates locally optimal solutions to the acquisition function, given by Eq. 1, quickly, by first randomly generating quality solutions via a construction procedure and



then subsequently improving them via a local search procedure.

$$\max_{\{\bar{x}_b \in \mathcal{X}\}_{b \in \mathcal{B}}} \sum_{b \in \mathcal{B}} \alpha(g(\bar{x}_b), \kappa) + \gamma \cdot \sum_{(b, \bar{b}) \in \mathcal{B} \times \mathcal{B}} \|\bar{x}_b - \bar{x}_{\bar{b}}\| \quad (1)$$

The first summation in the objective function of the optimization problem given in Eq. 1 sums the weighted outputs of the Gaussian process model for each of the selected points,  $\{\bar{x}_b^* \in \mathcal{X}\}_{b \in \mathcal{B}} - \alpha(\tilde{y}, \kappa)$  is equivalent to  $\mathbf{mean}(\tilde{y}) + \kappa \cdot \mathbf{var}(\tilde{y})$ , where  $\mathbf{mean}(\cdot)$  is a function that returns mean value of the Gaussian process model evaluated at  $\tilde{x} \in \mathcal{X}$  and  $\mathbf{var}(\cdot)$  is a function that returns variance of the Gaussian process model evaluated at  $\tilde{x} \in \mathcal{X}$ . The second summation sums the weighted Euclidian distances between the selected points,  $\{\bar{x}_b^* \in \mathcal{X}\}_{b \in \mathcal{B}}$ , where  $\gamma$  is a regularization term – this indirectly ensures that the selected points are sufficiently separated in the design space from each other.

The pseudo-code for the BO procedure is explicitly given in Algorithm 1. The procedure is a function of: i)  $f(\cdot)$ , the HFDT; ii)  $S$ , the number of random samples generated in the construction procedure of the GRASP; iii)  $\mathcal{K}$ , the set of major iterations for batch BO procedure; iv)  $\mathcal{B}$ , the set of batches; v)  $\mathcal{D}$ , initial set of inputs and outputs of the HFDT; vi)  $\mathcal{X}$ , continuous set of feasible inputs for the HFDT; vii)  $\alpha(\cdot, \cdot)$ , function that returns a weighted scalar output of the Gaussian process model parameterized by  $\kappa$ ; viii)  $\kappa$ , weighting parameter that balances “exploration” and “exploitation”; ix)  $\delta$ , minimal scaled distance between any two points in a batch; x)  $\sigma$ , variance for the noise utilized in generating the sorted restricted candidate list of selectable points in the construction procedure; and xi)  $\epsilon$ , incremental decrease in the minimal scaled distance between points in successive iterations of the construction procedure.

**Algorithm 1:** BO procedure.

```

BBOSM( $f, S, \mathcal{B}, \mathcal{K}, \mathcal{D}, \mathcal{X}, \alpha, \kappa, \delta, \sigma, \epsilon$ )
1 # initialization phase
2  $g \leftarrow \mathbf{ConstructGaussianProcess}(\mathcal{D})$ 
3 # iteration phase
4 for  $k \in \mathcal{K}$  do
5    $\{\bar{x}_b\}_{b \in \mathcal{B}} \leftarrow \mathbf{GRASP}(g, S, \mathcal{B}, \mathcal{X}, \alpha, \kappa, \delta, \sigma, \epsilon)$ 
6   parallel for  $b \in \mathcal{B}$  do;  $y_b^* \leftarrow f(\bar{x}_b^*)$ ; end;
7    $\mathcal{D} \leftarrow \mathcal{D} \cup \{(\bar{x}_b^*, y_b^*)\}_{b \in \mathcal{B}}$ 
8    $g \leftarrow \mathbf{ConstructGaussianProcess}(\mathcal{D})$ 
9 end
10 return  $g$ 

```

In the initialization phase of the procedure, a Gaussian process model,  $g(\cdot)$ , is created using the initial set of inputs and outputs,  $\mathcal{D}$ . In the iteration phase of the procedure, a set of  $|\mathcal{K}| \cdot |\mathcal{B}|$  points are judiciously selected to be evaluated. In each iteration,  $|\mathcal{B}|$  points are generated by the GRASP, and subsequently passed to the HFDT – the pseudo-code for the GRASP is given in Algorithm 2.

Once the points are evaluated, they are appended to the data set  $\mathcal{D}$ . Then the new updated data set,  $\mathcal{D}$ , is then utilized to create a new Gaussian process model and then the iterative procedure is continued for another  $|\mathcal{K}| - 1$  iterations.

**Algorithm 2:** GRASP to produce a high quality local optimal solution for the non-convex acquisition function.

```

GRASP( $g, S, \mathcal{B}, \mathcal{X}, \alpha, \kappa, \delta, \sigma, \epsilon$ )
1 # construction phase
2  $\{\bar{x}_b\}_{b \in \mathcal{B}} \leftarrow \mathbf{Construction}(g, S, \mathcal{B}, \mathcal{X}, \alpha, \kappa, \delta, \sigma, \epsilon)$ 
3 # local search phase
4  $\{\bar{x}_b\}_{b \in \mathcal{B}} \leftarrow \mathbf{LocalSearch}(g, S, \mathcal{B}, \mathcal{X}, \alpha, \kappa, \epsilon)$ 
5 return

```

In the first phase of the GRASP, an initial quality solution is generated by a construction procedure which randomly selects discrete points in the input space. In the second phase, the initial solution produced by the construction procedure is passed to a local search procedure to be improved, it should be noted that the solution is improved on in the continuous design space. The pseudo-codes for the construction procedure and the local search procedure can be seen in Algorithm 3 and Algorithm 4 respectively.

At the beginning of the construction procedure, two sets are created: i) the first set,  $\mathcal{S}$ , stores all the indices of all the selectable points; and ii) the second set,  $\bar{\mathcal{S}}$ , stores the indices of the selected points, since no indices have been selected the set is empty. Then  $|\mathcal{S}|$  random points are uniformly sampled from  $\mathcal{X}$  and subsequently evaluated by  $\alpha(\cdot, \cdot)$ . After the initialization phase, an iteration phase commences that selects  $|\mathcal{B}|$  candidate points. At the beginning of the iteration phase, in line 7 of Algorithm 3, a sorted list of unselected indices is generated by the function  $\mathbf{sort}(\cdot, \cdot)$ , which sorts the indices,  $\mathcal{S} \setminus \bar{\mathcal{S}}$ , in descending order based upon the value  $z_s + N(0, \sigma \cdot \mathbf{range}(\bar{z}))$ , where  $\mathbf{range}(\cdot)$  is a function that returns the maximum value of its input minus the minimum value of its input. Once the list of sorted indices is generated, a for loop is initiated and the Euclidian distance of the sample point corresponding to the index,  $s \in \mathcal{S} \setminus \bar{\mathcal{S}}$ , is compared to all the currently selected indices,  $\bar{\mathcal{S}}$ , and if the distance is greater than  $\delta$  away from all the currently selected points, the index is appended to the set  $\bar{\mathcal{S}}$ . This continues until  $|\bar{\mathcal{S}}| = |\mathcal{B}|$ . If by chance,  $|\bar{\mathcal{S}}| \neq |\mathcal{B}|$  then  $\delta$  is increased by  $\epsilon$  and the iteration phase continues until  $|\bar{\mathcal{S}}| = |\mathcal{B}|$ . We note that since we control the variability of samples in line 8 of Algorithm 3, we would prefer random uniform sampling over other sampling methods such as Latin hypercube sampling so that we can balance between exploration and exploitation.

**Algorithm 3:** Construction procedure to generate an initial set of candidate points.

**Construction**( $g, \mathcal{S}, \mathcal{B}, \mathcal{X}, \alpha, \kappa, \delta, \sigma, \epsilon$ )

```

1 # initialization phase
2  $\mathcal{S}, \bar{\mathcal{S}} \leftarrow \{1 \dots S\}, \emptyset$ 
3  $\{\bar{x}_s\}_{s \in \mathcal{S}} \leftarrow |S|$  random uniformly sampled points in  $\mathcal{X}$ 
4 for  $s \in \mathcal{S}$  do;  $z_s \leftarrow \alpha(g(\bar{x}_s), \kappa)$ ; end;
5 # iteration phase
6 while  $|\bar{\mathcal{S}}| < |\mathcal{B}|$  do
7   for  $s \in \text{sort}(\delta \setminus \bar{\mathcal{S}}, s: z_s + \mathcal{N}(0, \sigma \cdot \text{range}(\bar{\mathcal{Z}})))$  do
8     if  $\text{all}(\{\|\bar{x}_s - \bar{x}_{\bar{s}}\| \geq 1/(\delta \cdot |\mathcal{B}|)\}_{\bar{s} \in \bar{\mathcal{S}}})$  do
9        $\bar{\mathcal{S}} \leftarrow \bar{\mathcal{S}} \cup \{s\}$ 
10    end if
11    if  $|\bar{\mathcal{S}}| = |\mathcal{B}|$  do; break; end;
12  end
13   $\delta \leftarrow \delta + \epsilon$ 
14 end
15 return  $\{\bar{x}_s\}_{s \in \bar{\mathcal{S}}}$ 

```

Once the initial of points are generated in the construction procedure, they are passed to a local search procedure. In the local search procedure, each of the points are refined by solving the optimization problem given in line 2 of Algorithm 4, where  $\mathcal{N}(\cdot, \cdot, \cdot) \subseteq \mathcal{X}$  is a convex neighborhood in the output space of  $\alpha(\cdot, \cdot)$  evaluated at the point  $\bar{x} \in \mathcal{X}$ .

**Algorithm 4:** Local search procedure to refine the initial set of candidate points.

**LocalSearch**( $g, \mathcal{B}, \mathcal{X}, \alpha, \kappa, \{\bar{x}_b\}_{b \in \mathcal{B}}$ )

```

1 for  $b \in \mathcal{B}$  do
2    $\bar{x}_b \leftarrow \max_{\bar{x} \in \mathcal{N}(\bar{x}_b, \alpha, g, \kappa)} \alpha(g(\bar{x}), \kappa)$ 
3 end
4 return  $\{\bar{x}_b\}_{b \in \mathcal{B}}$ 

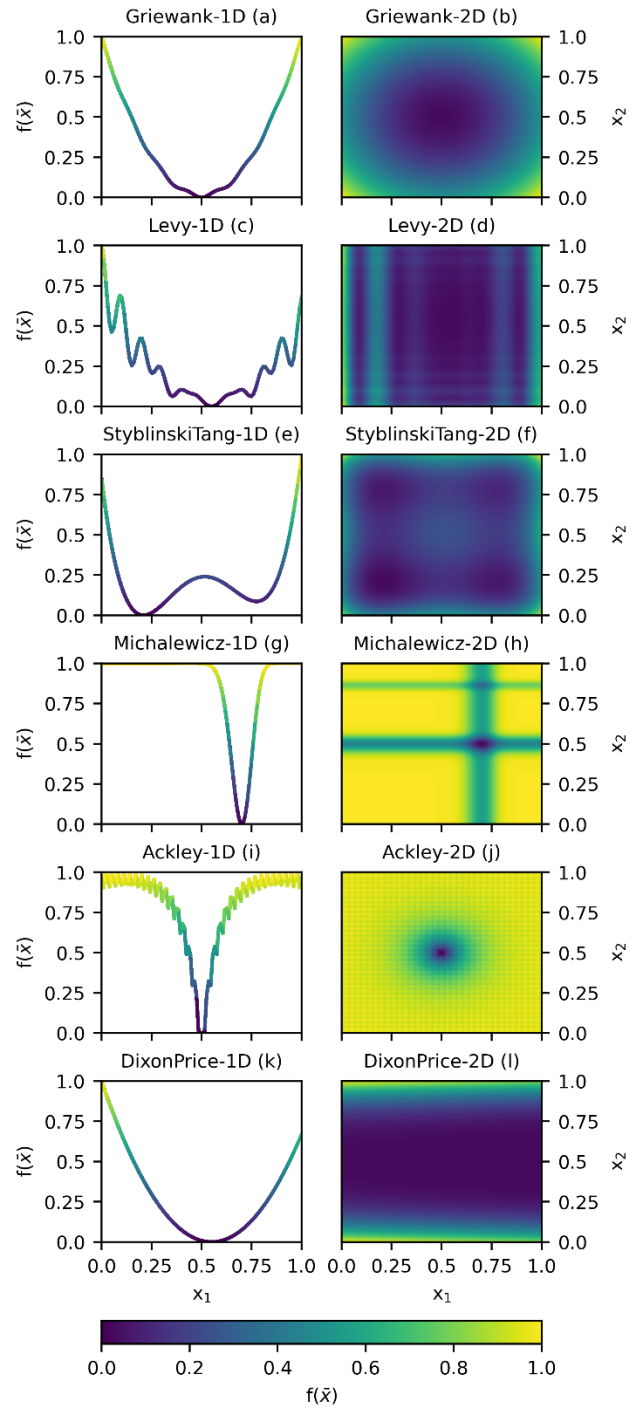
```

## COMPUTATIONAL EXPERIMENTS

This section presents the results of our computational experiments and highlights the ability of our parallel BO procedure to generate quality batches of sample points. The effectiveness of our BO procedure is compared against standard BO utilizing a batch size of one. In the first subsection, we present the benchmark test instances utilized in the computational experiments and the parameters utilized in our BO procedure. In the second subsection, we illustrate how our BO procedure generates a batch of points over sequential iterations for one of the benchmark test functions in two dimensions. In the final subsection, we illustrate the effectiveness of our parallel BO procedure, which produces a batch of samples to be evaluated in parallel per iteration, against standard BO, which evaluates a single point per iteration.

### Experimental set up

#### Benchmark test functions



**Figure 1.** Gallery of 6 test functions plotted in two input domains,  $\mathbb{R}^1$  (1D) and  $\mathbb{R}^2$  (2D).

We illustrate the effectiveness of our methodology on 6 standard benchmark test functions (StyblinskiTang, Griewank, Levy, Michalewicz, Ackley, and DixonPrice) across 4 input domains —  $\mathbb{R}^1$ ,  $\mathbb{R}^2$ ,  $\mathbb{R}^4$  and  $\mathbb{R}^6$ . The analytical form of the test instances were taken from the BoTorch library [16]. DixonPrice is the only strictly convex function and is, in general, thought to be “easy” to solve.

The remaining functions are non-convex and are thought to be “hard” to solve.

Figure 1 depicts each of the test functions in two input domains,  $\mathbb{R}^1$  and  $\mathbb{R}^2$ . For sake of illustration, the input domains and the output domains have been scaled on the unit cube, i.e. linearly spaced between 0 and 1. Subfigures a, c, e, g, i, and k are line plots and have an input domain of  $\mathbb{R}^1$ , while subfigures b, d, f, h, j, and l are contour plots and have an input domain of  $\mathbb{R}^2$ . For the figures with an input domain of  $\mathbb{R}^1$ , the input,  $x_1 \in \mathcal{X} \subset \mathbb{R}^1$ , is on the x-axis and the output,  $f(\bar{x})$ , is on the y-axis of the figure; the line color also corresponds to  $f(\bar{x})$ . For the figures with an input domain of  $\mathbb{R}^2$ , the inputs,  $(x_1, x_2) \in \mathcal{X} \subset \mathbb{R}^2$ , are on the x- and y-axis respectively and the color at two corresponding points represents the output,  $f(\bar{x})$ .

### Setting up the computational experiments

Here we briefly present the parameters passed to our BO procedure for conducting our computational experiments. We also describe our scaling procedure to ensure that exploration and exploitation are properly weighted. Table 1 presents the parameters passed to our BO procedure for each of the test instances.

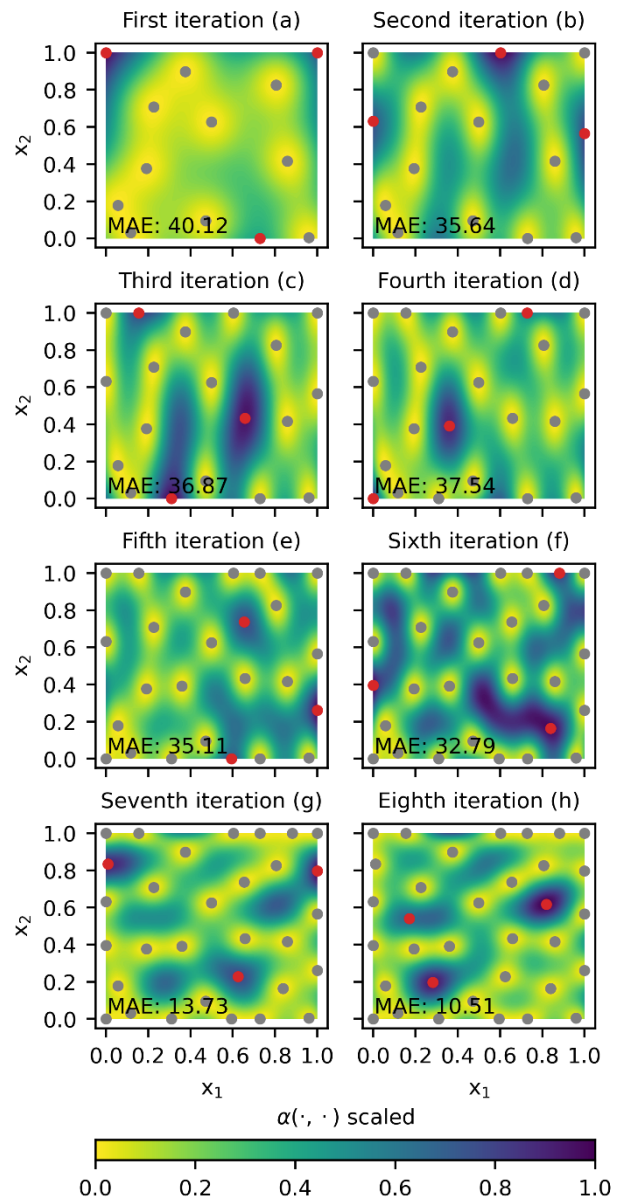
**Table 1:** Parameters passed to our BO procedure for each of the test instances.

Parameter	Value
$S$	300
$\kappa$	3
$\delta$	1
$\sigma$	0.1
$\epsilon$	0.1

To ensure that “exploration” and “exploitation” is properly scaled within  $\alpha(\cdot, \cdot)$ , we ensure that when a vector of design points are passed to the function their corresponding mean and variance are scaled to the unit cube i.e. all mean values are spaced linearly between 0 and 1 and all variance values are scaled between 0 and 1.

### Illustrative example

In this subsection, we use the StyblinskiTang test function with an input domain of  $\mathbb{R}^2$  to illustrate how the GRASP within our BO procedure generates new points in the design space to be evaluated by the HFDT in sequential iterations. For this illustrative example, 10 points in the design space were randomly sampled to build the initial set of ground truth data, subsequently these points were evaluated by the HFDT, and finally utilized to construct the initial Gaussian process model, in conjunction with the outputs from the HFDT. The batch size was set to 3 and the number of major iterations was limited to 8, giving rise to a total of 24 (plus the original 10) HFDT function evaluations.



**Figure 2.** Illustrative example showing how the GRASP in our BO procedure generates points to be sampled in the design space. Gray dots indicate points in the design space that have already been sampled. Red dots indicate points that the GRASP generated, which will be evaluated before the subsequent iteration commences.

Figure 2 depicts the results to the illustrative example. There are 8 subfigures, each representing a major iteration in our BO procedure. Each subfigure contains a surface plot and a scatter plot. The x- and y-axis in each subfigure represents the first input and second input variable to the HFDT respectively. The surface plot color represents the output of  $\alpha(\cdot, \cdot)$  scaled to the unit cube. The mean absolute error (MAE) of the Gaussian process model compared to the actual HFDT for 2000 randomly

selected points is stated in the bottom left corner of each subfigure. From inspection of each of the subfigures it is clear the GRASP within our BO procedure produces points to be evaluated by the HFDT with high values computed by the function,  $\alpha(\cdot, \cdot)$ , while simultaneously ensuring that these points are significantly far apart from each other, such that the “information gain” per iteration of the BO procedure is maximized. For instance, in the final iteration of the BO procedure, as illustrated in Fig 2.h, the 3 points generated to be sampled are conveniently located at what appears to the naked eye, as the three global maximums of the function  $\alpha(\cdot, \cdot)$ .

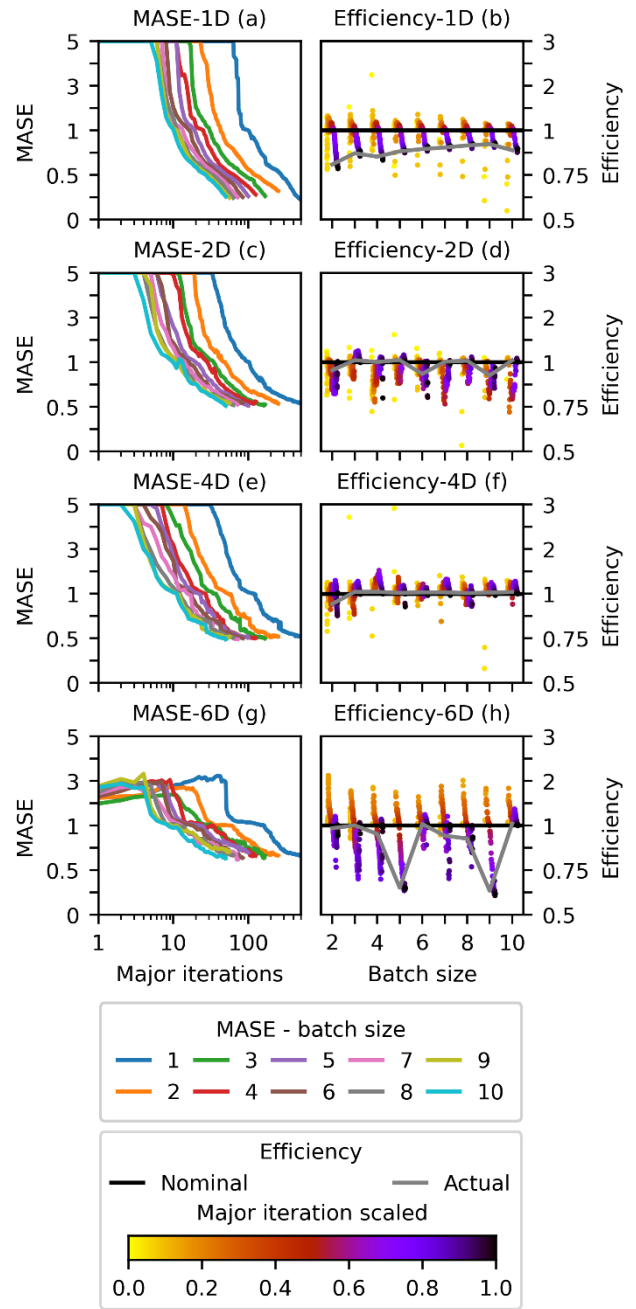
### Results from the computational experiments

To illustrate the benefits of our parallel computation framework, we present the computational results over our 6 test functions in various dimensions. In Figure 3, we show the aggregated mean absolute scaled error (MASE) and efficiency on the left and right, respectively. For each instance and batch size, we first compute its MASE defined in Equation (1), where the numerator is the MAE of the prediction of a Gaussian process (GP) model denoted by  $y_i^{GP}$  measured over  $N$ , 2000, ground truth values denoted by  $y_i^{GT}$ , and the denominator represents the MAE of the prediction of a reference GP model denoted by  $y_i^{GP,ref}$ . The MASE represents the relative performance of a surrogate model compared a reference model. We note that  $y_i^{GP,ref}$  is obtained from the BO after 100 function evaluations with a batch size of 1. Therefore, the denominator value does not change over iterations, whereas  $y_i^{GP}$  can change as we improve a GP model through iterations so can the numerator value. For fair comparison, the number of function evaluations is set to 500 across different batch sizes. Once MASE’s are computed for all instances, we aggregate them by taking their average to display the results of all instances in a single figure.

$$\text{MASE} := \frac{\frac{1}{N} \sum_{n \in \{1 \dots N\}} |y_i^{GP} - y_i^{GT}|}{\frac{1}{N} \sum_{n \in \{1 \dots N\}} |y_i^{GP,ref} - y_i^{GT}|} \quad (1)$$

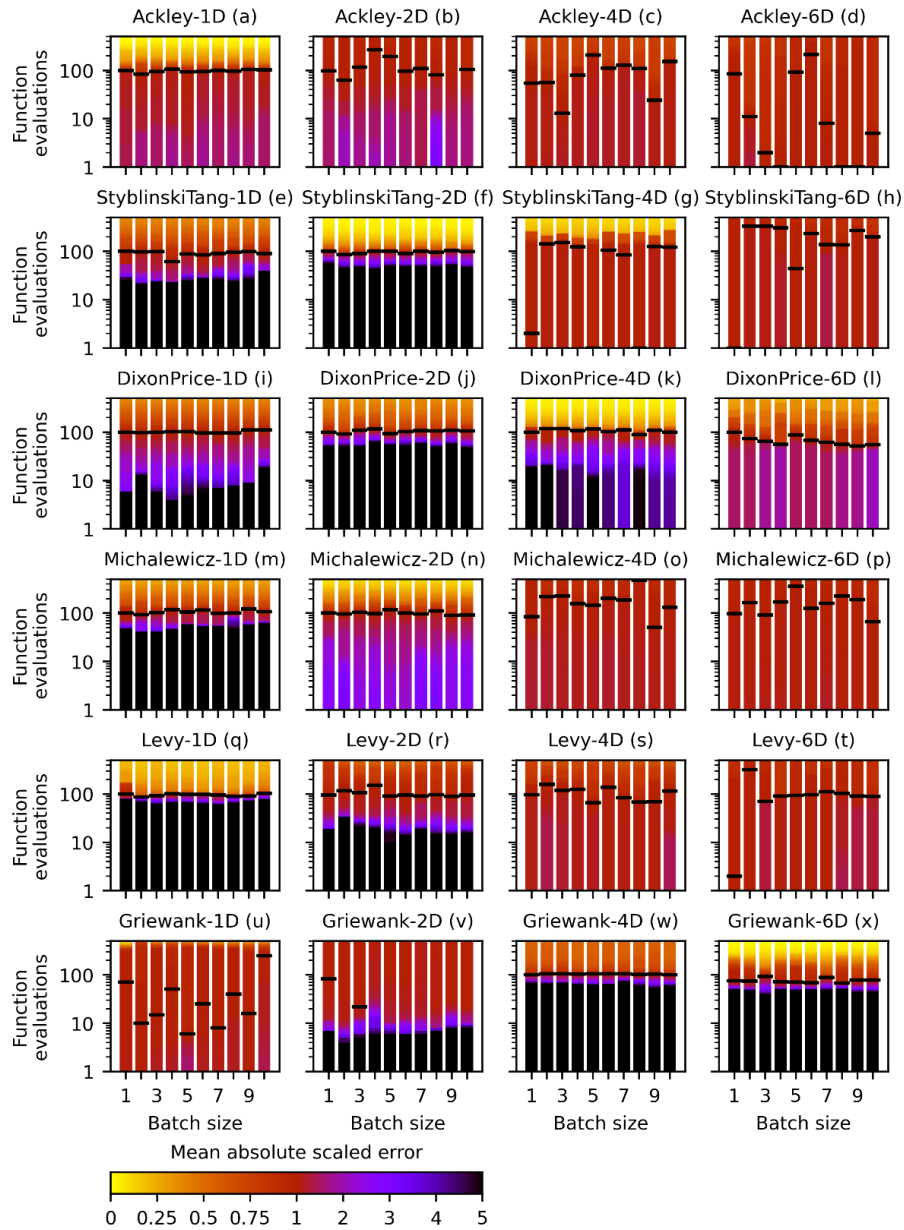
From the figure we see that the larger the batch size the quicker it reaches a MASE of 1 with a fewer number of iterations. For example, it took about 10 iterations for batch size 10 to reach a MASE of 1, while 100 iterations were needed for a batch size of 1 to arrive at that value. This result implies that we can gain a significant speed-up in computing a solution at high quality from parallel function evaluations.

As we increase dimensions, we observe that the MASE value tends to be smaller in early iterations. We think this is because the problem becomes harder as its dimension increases, leading to a smaller performance difference between the reference model and a model with less than 100 function evaluations.



**Figure 3.** Mean absolute scaled error (MASE) and efficiency of the aggregated test functions that were grouped by major iteration, dimension and batch size in subfigures a, c, e, and g; as well as the scaled efficiency of the aggregated test functions grouped by major iteration and batch size in subfigures b, d, f, h.

Whereas **Figure 3** illustrates surrogate quality aggregated over all test functions, **Figure 4** depicts the mean absolute scaled error for each test function in each



**Figure 4.** Mean absolute scaled error (MASE) as a function of the batch size and number of function evaluations for each function in 1, 2, 4, and 6 dimensions.

dimension of interest. Consistent with our intuition, for each batch size shown on the x axis, the MASE decreases as the number of function evaluations increases. As before, MASE is scaled relative to default BO with 100 function evaluations and a batch size of 1, which is shown in the first column of each subplot. In each column, a black horizontal line shows the number of function evaluations required for each batch size to attain an MASE of 1. Thus, the black horizontal line in the first column of each subplot is positioned at 100 function evaluations since this

column corresponds to default BO. As the batch size increases, however, the number of function evaluations required to attain the same MASE as default BO may differ. For example, in 1, 2, and 4 dimensions, the DixonPrice function indicates that, for all batch sizes, 100 function evaluations are needed. However, in 6 dimensions, fewer than 100 function evaluations are needed, suggesting that our approach is superior to default BO. In summary, from inspection of **Figure 4**, it is evident that in general the batch size does not affect the efficiency of our approach, i.e., the efficiency is roughly the same for a batch



of 1, which corresponds to sequential, to a batch size of 10. Consequently, our approach allows for increased parallelization, with minimal loss of efficiency.

## CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

Expensive-to-evaluate black-box functions often appear in challenging design problems involving multi-scale model integration. We devised and analyzed a GRASP heuristic embedded within a parallel Bayesian optimization framework to accelerate surrogate construction without sacrificing surrogate quality. Computational results revealed that our approach can yield equivalent or superior surrogate quality with near linear scaling in the number of samples evaluated in parallel.

As for future research directions, it would be interesting to investigate multi-fidelity BO, which makes use of lower fidelity observations available with a lower sampling cost. Specifically, it may be possible to query a lower-fidelity model when constructing a surrogate, which then introduces the additional complexity regarding how to sample: Take high- or low-fidelity samples in parallel.

## REFERENCES

1. T. R. Wanasinghe, W. Leah, B. K. Peterson, R. G. Gosine, L. A. James, O. De Silva, G. K. I. Mann and P. J. Warrian, "Digital twin for the oil and gas industry: overview, research trends, opportunities, and challenges," *IEEE Access*, vol. 8, pp. 104175-104197, 2020.
2. A. Bhosekar and M. Ierapetritou, "Advances in surrogate based modeling, feasibility analysis, and optimization: A review," *Computers & Chemical Engineering*, vol. 108, pp. 250-267, 2018.
3. J. Larson, M. Menickelly and S. M. Wild, "Derivative-free optimization methods," *Acta Numerica*, pp. 287-404, 2019.
4. L. M. Rios and N. V. Sahinidis, "Derivative-free optimization: a review of algorithms and comparison of software implementations," *Journal of Global Optimization*, pp. 1247-1293, 2013.
5. L. D. González and V. M. Zavala, "New paradigms for exploiting parallel experiments in Bayesian optimization," *Computers & Chemical Engineering*, vol. 170, p. 108110, 2023.
6. M. G. C. Resende and C. C. Ribeiro, *Optimization by GRASP*, New York: Springer Science+ Business Media, 2016.
7. J. Azimi, A. Fern and X. Z. Fern, "Batch Bayesian Optimization via Simulation Matching," in *Proceedings of the 23rd International Conference on Neural Information Processing Systems*, 2010.
8. E. A. Daxberger and B. K. H. Low, "Distributed Batch Gaussian Process Optimization," in *Proceedings of the 34th International Conference on Machine Learning*, 2017.
9. A. Shah and Z. Ghahramani, "Parallel Predictive Entropy Search for Batch Global Optimization of Expensive Objective Functions," in *Proceedings of the 28th International Conference on Neural Information Processing Systems*, 2015.
10. J. Wu and P. Frazier, "The Parallel Knowledge Gradient Method," *Advances in neural information processing systems*, vol. 29, 2016.
11. J. Wang, S. C. Clark, E. Liu and P. I. Frazier, "Parallel Bayesian global optimization of expensive functions," *Operations Research*, pp. 1850-1865, 2020.
12. G. De Ath, R. M. Everson, J. E. Fieldsend and A. A. M. Rahat, "?-Shotgun: ?-Greedy Batch Bayesian Optimisation," in *Proceedings of the 2020 Genetic and Evolutionary Computation Conference*, 2020.
13. K. Kandasamy, A. Krishnamurthy, J. Schneider and B. Póczos, "Parallelised Bayesian Optimisation via Thompson Sampling," in *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*, 2018.
14. X. Wang, Y. Jin, S. Schmitt and M. Olhofer, "Recent Advances in Bayesian Optimization," *ACM Computing Surveys*, vol. 55, no. 13, pp. 1-36, 2023.
15. J. Chen, F. Luo, G. Li and Z. Wang, "Batch Bayesian optimization with adaptive batch acquisition functions via multi-objective optimization," *Swarm and Evolutionary Computation*, vol. 79, p. 101293, 2023.
16. M. Balandat, B. Karrer, D. Jiang, S. Daulton, B. Letham, A. G. Wilson and E. Bakshy, "BoTorch: A framework for efficient Monte-Carlo Bayesian optimization," *Advances in neural information processing*

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Learn-To-Design: Reinforcement Learning-Assisted Chemical Process Optimization

Eslam G. Al-Sakkari<sup>a,b</sup>, Ahmed Ragab<sup>a,b\*</sup>, Mohamed Ali<sup>c</sup>, Hanane Dagdougui<sup>a</sup>, Daria C. Boffito<sup>d</sup>, and Mouloud Amazouz<sup>b</sup>

<sup>a</sup> Polytechnique Montréal, Department of Mathematics and Industrial Engineering, Montréal, Québec, H3T 1J4, Canada

<sup>b</sup> CanmetENERGY Varennes, Varennes, Québec, J3X 1P7, Canada

<sup>c</sup> CanmetENERGY Devon, Devon, Alberta, T9G 1A8, Canada

<sup>d</sup> Polytechnique Montréal, Department of Chemical Engineering, Montréal, Québec, H3T 1J4, Canada

\* Corresponding Author: [ahmed.ragab@nrcan-rncan.gc.ca](mailto:ahmed.ragab@nrcan-rncan.gc.ca)

## ABSTRACT

This paper proposes an AI-assisted approach aimed at accelerating chemical process design through causal incremental reinforcement learning (CIRL) where an intelligent agent is interacting iteratively with a process simulation environment (e.g., Aspen HYSYS, DWSIM, etc.). The proposed approach is based on an incremental learnable optimizer capable of guiding multi-objective optimization towards optimal design variable configurations, depending on several factors including the problem complexity, selected RL algorithm and hyperparameters tuning. One advantage of this approach is that the agent-simulator interaction significantly reduces the vast search space of design variables, leading to an accelerated and optimized design process. This is a generic causal approach that enables the exploration of new process configurations and provides actionable insights to designers to improve not only the process design but also the design process across various applications. The approach was validated on industrial processes including an absorption-based carbon capture, considering the economic and technological uncertainties of different capture processes, such as energy price, production cost, and storage capacity. It achieved a cost reduction of up to 5.5% for the designed capture process, after a few iterations, while also providing the designer with actionable insights. From a broader perspective, the proposed approach paves the way for accelerating the adoption of decarbonization technologies (CCUS value chains, clean fuel production, etc.) at a larger scale, thus catalyzing climate change mitigation.

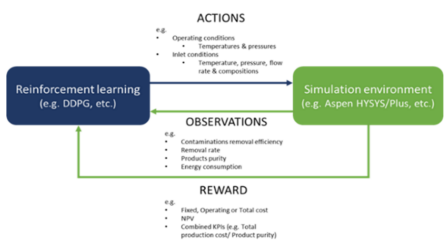
**Keywords:** Carbon Capture, Process Design, Optimization, Artificial Intelligence, Reinforcement Learning, Machine Learning, Simulation-based Optimization

## INTRODUCTION

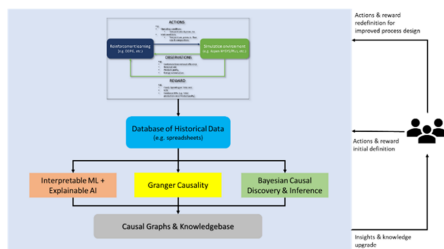
Climate change has become a crucial problem, where massive efforts are underway to mitigate its disastrous outcomes [Ref. Paris Agreement]. In this regard, there is a major trend to develop efficient carbon capture technologies, recognized as one of the promising solutions to the challenges posed by climate change. Furthermore, the research is accelerating in this hot field due to the clear and prioritized need for developing cost-efficient and environmentally friendly carbon capture processes. Various methods based on mechanistic models and combined with traditional operation research methods, including linear and non-linear programming, are

used to optimize existing carbon capture processes. Nevertheless, despite their robustness, most existing methods for optimizing process design suffer from certain limitations that hinder the development and optimization of carbon capture systems. These limitations include time consumption, high computational costs, and a lack of adaptability and transferability.

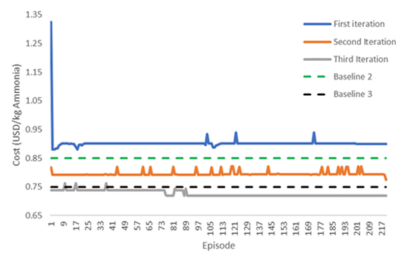
Therefore, there is a pressing need for the development of new methodologies to help overcome the above-mentioned limitations. Artificial intelligence (AI) and its related machine learning (ML) technologies holds significant promise to address this critical gap. Specifically, reinforcement learning (RL), as a prescriptive machine learning technique, offers an effective approach due to



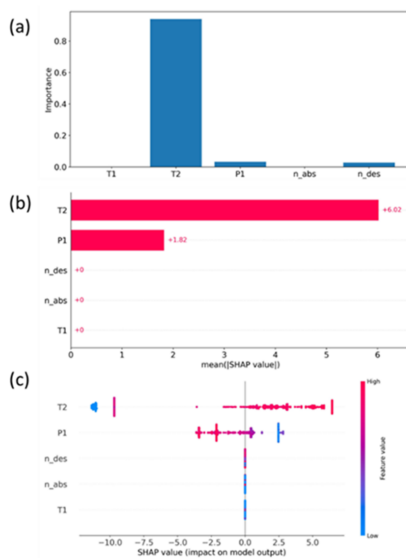
**Figure 1.** General simulation-based optimization approach.



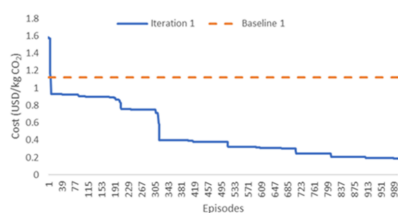
**Figure 2.** Summary of Causality Analysis Methodology



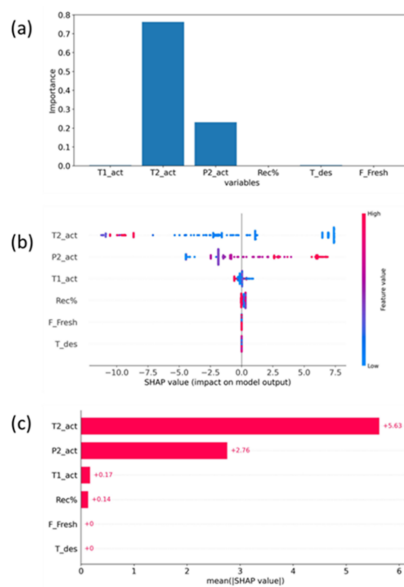
**Figure 4.** Optimization results (cost reduction along iterations and episodes) for ammonia-water system.



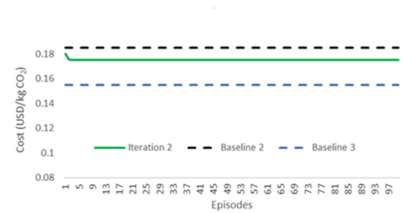
**Figure 3.** Results of IML+XAI (Dependency graphs) for ammonia-water system.



**Figure 6.** Optimization results for CO<sub>2</sub>-MEA system.



**Figure 5.** Results of IML+XAI (Dependency graphs) for CO<sub>2</sub>-MEA system.



being a learnable optimizer. In addition, its application has been successful in various fields, including process control [1], [2] which underscores its potential for accelerating chemical process design. RL encompasses various types, ranging from model-free to model-based algorithms, as well as policy-based and value-based agents. Examples of policy-based RL agents include deep Q-network (DQN) [3] and proximal policy optimization (PPO) [4]. Besides, there are agents that combine both policy-based and value-based calculations, such as deep deterministic policy gradient (DDPG) [5]. These diverse types offer versatile tools for enhancing chemical process design acceleration.

Recently, several research studies have employed RL within the realm of chemical engineering to optimize process sequences and to suggest new process flow diagrams for simple processes [6]–[12]. For instance, in [10], the authors used a hierarchical deep RL (DRL) agent to automate the synthesis of flowsheets for the ethyl tert-butyl ether (ETBE) production process. The study adopted the *SynGameZero* approach outlined in [13]. The Soft Actor Critic (SAC) algorithm was used in [9] to

design a distillation train to separate a mixture of benzene, toluene, and *p*-xylene. The SAC agent interacts with a Distillation Gym environment that interfaces with *COCO* open-source simulator. The paper [14] conducted a survey on the RL-assisted flow sheet synthesis, highlighting the interaction with other simulators, i.e. DWSIM, and the concept of transfer learning for process design [8]. Nonetheless, some of these attempts are still at early stages as the agents developed suggest technically infeasible flowsheets in some cases due to the lack of reasoning for the trained agent. Besides, one major limitation of these studies is the absence of causality. Analyzing the historical data gathered during the optimization process can offer new insights to the designers, enabling them to identify the causal variables that impact selected key performance indicators (KPIs), e.g. costs. Extracting knowledge from causality analysis part will lead to more generalizable results and provide actionable insights into hidden information and relationships that can be transferred to enhance other processes.

The novelty of the current work lies in the introduction of causal DRL to accelerate the design of chemical

processes. Two systems were considered for illustration: 1) ammonia-water system (hypothetical physical absorption-desorption process as a simplified example) and 2) mono-ethanolamine (MEA)-based carbon capture process. Introducing the concept of causal reinforcement learning to process design, specifically in applications related to solvent-based capture processes, can enhance the agent's reasoning capabilities, and enrich the knowledge of design experts with new insights. This approach can be generalized to other systems, under the full supervision of human experts, to streamline and accelerate the design process. The specific research objectives of this study are outlined as follows: 1) Develop an incremental DRL (IDRL) simulation-based optimization methodology to accelerate the design of steady state processes; 2) Integrate a newly developed causality analysis methodology, employing various algorithms with the IRDL to capture the underlying cause-effect relationships among different process variables and the specified KPI(s); 3) Deploy the developed IDRL coupled with causality analysis methodology to optimize different processes, including absorption-based carbon capture process utilizing amine solvents.

## METHODOLOGY

### General DRL Simulation-based Optimization Approach

The general simulation-based optimization approach and its underlying methodology achieving its objectives is depicted in **Figure 1**. A DRL agent is linked with a simulation environment in an interactive and learnable manner that minimizes the search space. Within this approach, the term "rewards" is used to denote the key performance indicators, whereas the term "states" is used to represent the observations defining the current condition of the process. The agent provides "actions" (representing design variables including equipment sizing and operating conditions) to the simulation environment based on the current state and the observed reward. The agent tries to find the optimal point without sticking in local minima where it may follow the shortest path, depending on factors such as problem complexity, reward formulation, the RL algorithm selected, and available computational resources. Unlike most of the traditional optimizers, DRL-based optimization is driven by the concept of reward and penalty. The agent receives a reward when following the right path of optimizing the objective function, i.e. minimizing the process costs, or minimizing its environmental impact. Conversely, a penalty is incurred if the agent followed the wrong direction, where the actions result in cost or environmental impacts exceeding specified values. To avoid settling in local minima, the agent poses the ability to explore new scenarios, aiming to identify paths leading to the global minimum point. Balancing

between exploitation and exploration is crucial to speed up achieving the optimal solution. Additionally, achieving an accelerated optimization requires proper formulation of reward function and optimization of the agent's hyperparameters.

#### DRL and problem formulation as Markov Decision

**Process.** From a mathematical standpoint, when employing RL, the problem is typically formulated as a *Markov Decision Process* (MDP). This formulation involves a set of observations that characterize the state of the environment, along with a distribution of initial states. Additionally, it entails a set of actions, whether continuous or discrete, that can modify the state of the environment based on rewards received by the agent. An MDP comprises transition dynamics or probabilities, which describe the relation between states and actions. Two other important components are the immediate rewards and the discount factor (which ranges from 0 to 1), a lower discount factor places greater emphasis on the immediate reward. For a more detailed explanation of the mathematical formulation of RL and MDP, refer to this study [3]. It is worth noting that one of the primary differences between dynamic and steady-state problems lies in the definition of discount factor during MDP formulation. The discount factor is an optional element in defining an MDP and can be disregarded (assigned a zero value) when considering the optimization of steady-state processes, where immediate rewards take precedence over long-term ones. Disregarding the discount factor in case of steady-state processes facilitates the design optimization problem definition/formulation. Nevertheless, finding the optimal policy based on the cumulative reward for the RL agent remains relevant. This design problem adheres to the *Markov Property*, resulting in a memoryless agent, where the current state depends only on the previous state and influences the selection of the subsequent actions that modify the successive states. Consequently, the transition dynamics or probabilities over time and the whole states history will not be taken into consideration. The transition probabilities between two successive states based on the proposed or taken actions are primarily deterministic. In chemical process design, these transitions may involve thermodynamic/equilibrium relations, kinetic/rate equations, or heuristic rules.

**RL Algorithm selection & development of the simulation environment.** Various RL algorithms can be developed, optimized, and then utilized for process design optimization. As previously mentioned, the PPO and SAC algorithms were employed process flowsheet synthesis. In this work, we consider the DDPG algorithm [5], which is one of the promising algorithms. It combines the value-based and policy-based optimization principles.

Upon developing the simulation environment, the potential actions (design variables) under consideration include equipment sizing and process operating

conditions. Additionally, parameters such as recycle rate, purge percentage and flowrates of inlet streams are deemed significant actions to be prescribed or optimized. Equipment sizing varies according to the equipment type. For instance, the number of stages is a crucial variable for separation columns, defining their heights and costs. Continuous stirred reactors (CSTRs) are sized by volume, which is also related to the percentage conversion of limiting reactants. Heat exchange equipment sizing is determined by the heat exchange area, which is also related to other process operating conditions. Process operating conditions encompass different streams/equipment temperatures and pressures. Key observations describing the simulation environment's state include the purities of specified product streams/components, recovery rates of components of interest, and utilities consumption. These states and actions vary from a problem to another depending on the specific requirements of the problem and the specifications set by the process expert.

**Reward definition (formulation).** The reward function can take various forms, depending on the specified objective defined by the process expert. For instance, it can be formulated as the relative production cost, i.e. production cost per kg of product or material. Alternatively, it could represent the total capital investment (TCI), net present value (NPV) or any other form of economic KPIs. Furthermore, it might be a combined value of both techno-economic KPIs such as relative production cost per material/product recovery or purity. One objective function that can serve as a highly relevant reward function is the total environmental impact or the relative emissions per kg of the product. This function holds significant importance in current and future research studies for process design optimization, enabling the assessment of processes' environmental footprints and promoting the development of eco-friendly processes.

**Incremental learning.** It is worth noting that our approach for the DRL-based optimization is built upon the concept of incremental learning. Specified costs, environmental impacts and actions ranges change from one case to another until the maximum allowable performance is achieved, i.e. lowest possible relative production cost and minimum acceptable environmental impact. More specifically, the actions ranges are narrowed or redefined after analyzing the historical data collected from previous RL-assisted optimization cases. In addition, reward function and its value are redefined based on this analysis. For example, the specified cost value may be reduced to explore the agent's ability to surpass it and reach the minimum allowable cost.

## Towards causality analysis for causal and rational DRL

Following the incremental learning-based optimization, historical data representing the simulation runs

resulting from the agent-simulation interactions are stored in spreadsheets for the causal analysis, where each run represents a data point (observation). This is to gain more actionable insights into the root causes and the importance of design variables impacting the KPIs. Three methods were employed to perform a preliminary causality analysis based on observations. These methods include combined interpretable ML (IML)-Explainable AI (XAI), Granger causality (GC), and Bayesian causal discovery & inference.

**Combined IML-XAI.** Initially, the decision tree classification algorithm (i.e. IML) is combined with XGBoost and SHAP explanation algorithm (i.e. XAI) to extract the patterns, assess variables importance and analyze the relationships governing the dependence of the specified KPI (reward) and observations (state) on the design variables (actions). The IML extracts the patterns (ranges of design variables) that lead to processes with reasonable feasibility. Additionally, it assesses the variables importance, determining the most significant ones. The XAI gives relative magnitudes and directions of the effect of each input variable on the final KPI (process overall production cost). This will help afterwards in the incremental learning procedure, helping to select/redefine the relevant input variables ranges for guided optimization.

**Granger causality (GC).** This method is used preliminary to validate the relations between the actions (especially the most significant ones identified by the combined IML-XAI method) and the KPI. Since actions are provided continuously by agent and they depend on the previous set of actions and states, consequently actions, states, and rewards/KPIs are treated as a form of time series data, leading to the application of GC. As mentioned earlier, in case of optimizing steady-state processes, the current state (representing the current process configuration) is solely influenced by the preceding one, rendering longer-term relationships negligible. The dependence between each successive pair of states ( $S_t$  &  $S_{t+1}$ ) and their associated set of actions based on the reward provided to the agent, justifies the utilization of approximated time series assumption to apply GC.

**Bayesian causal discovery & inference.** This method is used to generate causal graphs to discover the root causes that significantly affect changes of the KPI. It also reveals the interdependence among actions, discovering the root cause design variable that affects both other design variables and KPI. **Figure 2** summarizes the proposed causality analysis methodology. Further details regarding the combined IML-XAI method can be found in our recent study [15]. The insights gained from this causality analysis will upgrade/enrich the experts' knowledge, where they can redefine the baselines/rewards and the range of actions to retrain the agent, thereby fostering further improvement in design optimization.



This methodology can be applied in flowsheet synthesis to enhance the DRL agent's reasoning capability by providing it with the extracted knowledge and causal graphs. This empowers the agent to operate as a rational causal DRL agent that avoids the synthesis of technically infeasible flowsheets/processes, thus accelerating the synthesis and design process. In addition to the other design variables, equipment and their sequencing will be added to the list of actions. Consequently, the causal graphs will account for these two crucial categorical variables and quantify their impact on the defined KPI/reward. Furthermore, integrating domain knowledge from process designers and operators as constraints into the causal DRL agent's structure will enhance its rationality. These domain knowledge-informed constraints ensure the avoidance of invalid streams/equipment selection, placement, and sequencing, e.g. separation train sequencing [16] and chemical reactors sequencing [17]. This will further shorten the time needed for the agents to converge on the optimal new process, thus ensuring the desired ML-assisted design acceleration.

## SIMULATION CASE STUDIES

Aspen HYSYS was selected as the simulation software for constructing the process flow diagrams and conducting the process design. Two simulation case studies were developed to validate the proposed CIRL approach. The simplified process flow diagrams of the two case studies are presented in the supplementary materials and their details are summarized in the following subsections.

### Case study 1: Ammonia water system

An ammonia-water absorption-desorption system was developed with a stream of air contaminated with ammonia is introduced to the absorption column, where pure water is utilized as a physical solvent to separate ammonia from air. The captured ammonia is then removed from the water solvent using steam in a re-boiled stripping column. Heat exchange between the stream entering the stripping column and its bottoms product stream enables heat integration, minimizing the reboiler heat load. A condensation and a flash separation are applied for the top product stream of stripping column to remove any water associated with the separated ammonia, thus enhancing its purity and recovering any escaped solvent. Recycle was not considered in this simplified case study. The NRTL simulation fluid package was used [18]. The actions considered include number of stages in both absorption and stripping columns, temperatures of fresh solvent and inlet contaminated gaseous stream, and the pressure of the inlet contaminated gaseous stream.

### Case study 2: Mono-ethanolamine (MEA)-based carbon capture system

This case considers chemical absorption (CO<sub>2</sub>-MEA system), where flue gas enters the absorption column after solids removal, pressurization, and cooling to the desired operating temperature. The absorption-desorption setup is like the case of ammonia (case study 1). However, in this case, lean MEA is recirculated, after carbon dioxide stripping, to the absorption column and mixed with a makeup stream to enhance the process feasibility. A purge stream is used to avoid excessive accumulation of carbon dioxide in the system and the recycled lean MEA stream. Effect of temperature and pressure of inlet flue gas stream and temperature of makeup solvent were considered as actions. Effect of recycle percentage and temperature of stream entering the stripping column were also considered for optimization. The *Acid Gas - Chemical Solvents* Aspen built-in fluid package was used to perform the simulation calculations. The stream of flue gases comes from a power plant, where all conditions defined according to that study [19]. A baseline/reference for the carbon dioxide recovery cost of 1.12 USD/kg was adopted from a published study [20] that ultimately achieved the lowest cost after several iterations.

### Reward definition

Important observations that describe the state of simulation are purity of purified gas stream, recovery of contaminants (for ammonia and carbon dioxide), purity of the product stream, and utilities consumption. The reward function is defined as the relative removal cost, i.e. removal cost per kg of captured material. The procedure for calculating the removal cost and other related costs follows the guidelines outlined in this textbook [21] and these two studies [22], [23]. Utilities and MEA prices were sourced from a study published in [24]. The reward was determined as follows: if the agent progresses in the specified direction (new cost  $\leq$  specified cost), it receives a value equivalent to the total production cost (unity factor); otherwise, it receives a zero value. Furthermore, if the new cost in current step is lower than the cost of previous step, the agent receives the full total production cost value; otherwise, it receives zero.

## RESULTS AND DISCUSSION

### Optimized DRL agent

The agent developed and adopted in this study is based on DDPG, a policy-based model-free reinforcement learning algorithm. It comprises two neural networks, one for the actor and one for the critic. For a deeper understanding of the mathematical underpinnings of the DDPG, interested readers can refer to [1], [25], [26]. The structure and crucial hyperparameters of

the optimized DDPG agent are summarized in **Table 1**.

**Table 1:** DDPG optimized hyperparameters.

Hyperparameter	Value
Action noise standard deviation	0.2
Actor learning rate	0.001
Critic learning rate	0.002
Discount factor (gamma)	0
Tau parameter used to update target networks	0.005
Total number of episodes	100-1000
Activation function for the critic neural network dense layers	ReLU
Activation function for the actor neural network inputs dense layers	ReLU
Activation function for the actor neural network outputs dense layer	Tanh

The *Tanh* activation function was selected to standardize the range of actions between -1 to 1, which are then converted back to their real values before being used in the simulation.

## Process design optimization

### Ammonia-water system

For the ammonia-water absorption system, the initial range for both inlet solvent and flue gas streams temperatures was set between 20 to 100 °C, with pressure ranging from 600 to 950 kPa. It is important to mention that the absorption column pressure was set equal to the inlet pressure of flue gas. This simplification facilitates the manipulation of the absorber pressure in the simulation environment by the RL agent. It is important to note that the inlet flue gas pressure must be higher than that of the absorber to ensure proper flow inside the column. Hence, a pressure drop value will be added to specify the pressure of inlet flue gas to the designer to incorporate it in real life operations. Additionally, the range of number of stages for both absorption and desorption columns was from 3 to 15. The initial desired cost (baseline) had the value of 2 USD/kg of ammonia based on a simulation of base case. As a result, after 220 episodes, a cost of almost 0.880 USD/kg of ammonia was obtained initially, which is significantly lower than the specified baseline. Accordingly, following the proposed incremental learning method, this value was given as a new baseline to the agent for further improvement of the process economic optimization, where the ranges of variables were the same. Surprisingly, the agent could reach another lower production/recovery cost of 0.775 USD/kg of ammonia. The preliminary causality analysis methodology was applied to extract some other knowledge about the most effective variables and the patterns to achieve lower process costs. Thus, the historical data from these two first

optimization iterations was collected, preprocessed, and analyzed. The preprocessing was done through dividing the observations to high cost (above 0.9 USD/kg of ammonia) observations and labeled as Class (0). The other portion was labeled as Class (1), which represented the low cost (below 0.9 USD/kg of ammonia) observations. **Figure 3** represents the results of employing IML+XAI to get the dependency graphs.

As shown, the most effective two variables are T2 (temperature of flue gases entering the absorption column) and the pressure of the solvent stream. These results were validated through performing the GC analysis to test the dependency of the continuous cost change corresponding to the continuous actions given by the agent, in which the new actions depend on the previous actions, state and reward. *p*-value of the dependence between T2 and cost was < 0.001 which indicates high GC. Besides, according the XAI dependency graphs, higher values of T2 lead to lower costs (Class (1)). The variable P1 possesses clearly lower significance, where, after data preprocessing/analysis, the lower values in the range of 830-900 kPa (around 850 kPa) lead to lower costs (Class (1)). These results of the current case mean that the operating costs dominate the removal cost. In addition, upon analyzing the available data, to achieve higher purities and recoveries of ammonia, it is preferred to utilize absorption columns with low number of stages and stripping column with a higher number of stages. The causal incremental learning methodology aims at minimizing the search space to accelerate the process design optimization. Hence, after obtaining the dependency graphs, the ranges of the variables were redefined and narrowed to see if further improvement can be obtained. The new ranges are (85-99 °C) for T2, (25-41 °C) for T1, (800-900 kPa) for P1, (3-5) for absorption number of stages and (10-15) for stripping column number of stages. In addition, a new baseline was set to 0.75 USD/kg of ammonia. As a result, a new lower production cost of 0.719 USD/kg of ammonia was obtained at the new conditions of T1 = 40.5 °C, T2 = 99 °C, P1 = 900 kPa, absorption number of stages = 4.8 (almost 5) and stripping column number of stages = 15. At these conditions, the purity of outlet air and ammonia recovery were 97.2% and 96.3%, respectively. **Figure 4** summarises the results of the proposed causal incremental learning methodology for process design optimization. The Bayesian causal graph showed that the temperature of the flue gas stream is one of the root causes of changing the KPI (removal cost) and can even affect other design variables (absorption column stages (*n\_abs*) & stripping column stages (*n\_des*)) and state. This graph supports the results of the first two methods and is provided in the supplementary materials.

### MEA-based Carbon Capture system

In the case of MEA-based carbon capture system, the first ranges, according to literature [19], were inlet solvent and flue gas temperatures (T1 and T2) of 30-90 °C and inlet flue gas pressure (P1) of 150-500 kPa. Besides, recycle percentage ranged initially between 90% and 98%, while the range of temperatures of rich MEA stream entering the stripping column was 77-85 °C. As in the case of ammonia-water system, the same assumption of considering the absorption column pressure as the inlet flue gas pressure is still applicable. After using the first baseline of 1.12 USD/kg CO<sub>2</sub>, this initial optimization led to a lower production cost of almost 0.190 USD/kg CO<sub>2</sub>. The collected data was then preprocessed and labeled to the IML-XAI based low level causality analysis of the historical simulation data collected from this iteration show that the most significant variables are flue gas inlet temperature and its pressure. The GC analysis confirmed this where the  $p$ -value in the case of cost dependence on T2 did not exceed 0.0001. It is worth mentioning that when the test was done in the reverse direction the  $p$ -value exceeded 0.9; this means that T2 causes the change in the cost but not the vice versa. In addition, the  $p$ -values in the case of the other variables T1, Rec%, T before stripping column and the fresh solvent flowrate exceeded 0.1, which indicates low to no significance of these variables compared to T2 in the current study within the specified range. Figure 5 shows the results of the combined IML+XAI approach in the current case of MEA-based carbon capture system. According to these results the ranges of the variables were narrowed/redefined for the sake of exploration for further optimization through retraining the agent considering the extracted causal knowledge, which gives it more reasoning. This reasoning/causality accelerates the retraining step and the subsequent design optimization. In this regard, the ranges of the actions were narrowed/redefined to become inlet solvent and flue gas temperatures (T1 & T2) of 23-35 °C and inlet gas pressure (P1) of 250-400 kPa in the second iteration. In addition, the new range of the recycle percentage was 85%-95% and accordingly the range of temperatures of rich amine stream entering the stripping column changed to 71-81 °C. Furthermore, the baseline was decreased to 0.185 USD/kg CO<sub>2</sub>. A third baseline, based on the expertise of carbon capture and process engineering experts, was put at 0.155 USD/kg CO<sub>2</sub> for comparison and more improvement in future research. Figure 6 summarizes the results of the two training iterations with their corresponding baselines. As shown, the minimum relative removal cost obtained was about 0.175 USD/kg CO<sub>2</sub>. This cost was obtained after 25 episodes during the second iteration (agent retraining step). The conditions that achieved this value were T1 of 24 °C, T2 of 25 °C, P2 of 270 kPa, recycle percentage of 85 % and temperature of the rich amine entering the desorption column of 71 °C, respectively. Additionally, the

carbon capture percentage at these conditions was 92.5% with a carbon dioxide purity of 98%. The causal graph based on Bayesian framework showed that one of the clear root causes of changing the KPI and can even affect other design variables and state is the temperature of the flue gas stream in agreement with what was obtained by the first two methods. This graph is put as supplementary material.

## RECOMMENDATIONS FOR FUTURE WORK

In forthcoming research, we aim to incorporate emissions and other environmental impacts as crucial optimization objectives. In particular, the net emissions should be considered as a reward, where the emissions of the capture process itself should be minimized as possible. We advocate for exploring the application of transformers-based RL for flowsheet synthesis and process design acceleration. This will allow adding more reasoning to the causal incremental learning-based RL agent to help in generating/suggesting new technically feasible process flow diagram. This will be done also through defining constraints based on process design experts' domain knowledge and rules as a practical application of the interesting concept of RL with human preferences. This will guarantee the acceleration of process design under the full supervision of human experts to always insure simultaneous technical and economic feasibility of the generated flowsheets. Uncertainty of inlet streams composition and energy/fuel prices will be considered in our future research as well which is one of the advantages of RL to deal with uncertainties.

## CONCLUSION

This study introduced and employed a causal Incremental Reinforcement Learning (CIRL) agent as a learnable optimizer with reasoning capabilities to accelerate efficient process design. Two case studies were considered to validate the approach, physical absorption (ammonia-water system) and chemical absorption (CO<sub>2</sub>-MEA system). In the case of chemical absorption, the causal DRL agent reached an optimal cost after only 25 episodes. The causality analysis revealed that key variables, such as inlet flue gas temperature and absorption pressure, significantly influence the CO<sub>2</sub>-MEA system's production cost. For instance, it showed that the relative production cost of the process is inversely proportional to the flue gas temperature fed to the absorber. Feeding the agent with the extracted knowledge and causal information obtained from the causality analysis helped it to achieve more improved process design. These findings hold relevance beyond the studied systems and offer actionable insights for future absorption system designs.

## REFERENCES

1. K. Nadim et al., "Learn-to-supervise: Causal reinforcement learning for high-level control in industrial processes," *Eng. Appl. Artif. Intell.*, vol. 126, p. 106853, (2023)
2. P. Daoutidis et al., "Machine learning in process systems engineering: Challenges and opportunities," *Comput. Chem. Eng.*, vol. 181, p. 108523, (2024)
3. K. Arulkumaran et al., "Deep reinforcement learning: A brief survey," *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 26–38, (2017).
4. C. Yu et al., "The surprising effectiveness of ppo in cooperative multi-agent games," *Adv. Neural Inf. Process. Syst.*, vol. 35, pp. 24611–24624, (2022)
5. T. P. Lillicrap et al., "Continuous control with deep reinforcement learning," *arXiv Prepr. arXiv1509.02971*, (2015)
6. L. Stops et al., "Flowsheet generation through hierarchical reinforcement learning and graph neural networks," *AIChE J.*, vol. 69, no. 1, p. e17938, (2023)
7. Q. Göttl et al., "Deep reinforcement learning uncovers processes for separating azeotropic mixtures without prior knowledge," *arXiv Prepr. arXiv2310.06415*, (2023)
8. Q. Gao et al., "Transfer learning for process design with reinforcement learning," *arXiv Prepr. arXiv2302.03375*, (2023)
9. L. I. Midgley, "Deep reinforcement learning for process synthesis," *arXiv Prepr. arXiv2009.13265*, (2020)
10. Q. Göttl et al., "Automated flowsheet synthesis using hierarchical reinforcement learning: proof of concept," *Chemie Ing. Tech.*, vol. 93, no. 12, pp. 2010–2018, (2021)
11. S. J. Plathottam et al., "Solvent extraction process design using deep reinforcement learning," *J. Adv. Manuf. Process.*, vol. 3, no. 2, p. e10079, (2021)
12. A. A. Khan and A. A. Lapkin, "Designing the process designer: Hierarchical reinforcement learning for optimisation-based process design," *Chem. Eng. Process. Intensif.*, vol. 180, p. 108885, (2022)
13. Q. Göttl et al., "Automated synthesis of steady-state continuous processes using reinforcement learning," *Front. Chem. Sci. Eng.*, pp. 1–15, (2021)
14. Q. Gao and A. M. Schweidtmann, "Deep reinforcement learning for process design: Review and perspective," *arXiv Prepr. arXiv2308.07822*, (2023)
15. E. G. Al-Sakkari et al., "Machine learning-assisted selection of adsorption-based carbon dioxide capture materials," *J. Environ. Chem. Eng.*, p. 110732, (2023)
16. E. Marcoulaki et al., "Design of separation trains and reaction-separation networks using stochastic optimization methods," *Chem. Eng. Res. Des.*, vol. 79, no. 1, pp. 25–32, (2001)
17. R. Chebbi, "Optimizing reactors selection and sequencing: minimum cost versus minimum volume," *Chinese J. Chem. Eng.*, vol. 22, no. 6, pp. 651–656, (2014)
18. Simoni, L. D., Lin, Y., Brennecke, J. F., & Stadtherr, M. A. Modeling liquid– liquid equilibrium of ionic liquid systems with NRTL, electrolyte-NRTL, and UNIQUAC. *Industrial & engineering chemistry research*, 47(1), 256–272. (2008)
19. L. E. Øi, "Aspen HYSYS simulation of CO<sub>2</sub> removal by amine absorption from a gas-based power plant," in *The 48th Scandinavian Conference on Simulation and Modeling (SIMS 2007)*, pp. 73–81 (2007)
20. J. Chen and F. Wang, "Cost reduction of CO<sub>2</sub> capture processes using reinforcement learning based iterative design: A pilot-scale absorption–stripping system," *Sep. Purif. Technol.*, vol. 122, pp. 149–158, (2014)
21. M. S. Peters et al., *Plant design and economics for chemical engineers*, vol. 4. McGraw-hill New York, (1968)
22. E. G. Al-Sakkari et al., "Comparative Technoeconomic Analysis of Using Waste and Virgin Cooking Oils for Biodiesel Production," *Front. Energy Res.*, p. 278, (2020)
23. M. M. Naeem et al., "Single-stage waste oil conversion into biodiesel via sonication over bio-based bifunctional catalyst: optimization, preliminary techno-economic and environmental analysis," *Fuel*, vol. 341, p. 127587, (2023)
24. N. Wang et al., "MEA-based CO<sub>2</sub> capture: a study focuses on MEA concentrations and process parameters," *Front. Energy Res.*, vol. 11, (2023)
25. D. Mehta, "State-of-the-art reinforcement learning algorithms," *Int. J. Eng. Res. Technol.*, vol. 8, pp. 717–722, (2020)
26. Y. Hou et al., "A novel DDPG method with prioritized experience replay," in *2017 IEEE international conference on systems, man, and cybernetics (SMC)*, pp. 316–321 (2017)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Cost-optimal Selection of pH Control for Mineral Scaling Prevention in High Recovery Reverse Osmosis Desalination

Oluwamayowa O. Amusat<sup>a,\*</sup>, Alexander V. Dudchenko<sup>b</sup>, Adam A. Atia<sup>c,d</sup>, and Timothy Bartholomew<sup>c</sup>

<sup>a</sup> Lawrence Berkeley National Laboratory (LBNL), 1 Cyclotron Rd, Berkeley, CA 94720, USA

<sup>b</sup> SLAC National Accelerator Laboratory, 2575 Sand Hill Rd, Menlo Park, CA 94025, USA

<sup>c</sup> National Energy Technology Laboratory (NETL), Pittsburgh, PA 15236, USA

<sup>d</sup> NETL Support Contractor, Pittsburgh, PA, 15236, USA

\* Corresponding Author: [ooamusat@lbl.gov](mailto:ooamusat@lbl.gov)

## ABSTRACT

Explicitly incorporating the effects of chemical phenomena such as chemical pretreatment and mineral scaling during the design of treatment systems is critical; however, the complexity of these phenomena and limitations on data have historically hindered the incorporation of detailed water chemistry into the modeling and optimization of water desalination systems. Thus, while qualitative assessments and experimental studies on chemical pretreatment and scaling are abundant in the literature, very little has been done to assess the technoeconomic implications of different chemical pretreatment alternatives within the context of end-to-end water treatment train optimization. In this work, we begin to address this challenge by exploring the impact of pH control during pretreatment on the cost and operation of a high-recovery desalination train. We compare three pH control methods used in water treatment ( $H_2SO_4$ , HCl, and  $CO_2$ ) and assess their impact on the operation of a desalination plant for brackish water and seawater. Our results show that the impact of the acid choice on the cost can vary widely depending on the water source, with  $CO_2$  found to be up to 11% and 49% more expensive than HCl in the seawater and brackish cases, respectively. We also find that the acid chemistry can significantly influence upstream processes, with use of  $H_2SO_4$  requiring more calcium removal in the softening step to prevent gypsum scaling in HPRO system. Our work highlights why incorporating water chemistry information is critical when evaluating the key cost and operational drivers for high-recovery desalination treatment trains.

**Keywords:** Technoeconomic Analysis, Optimization, Surrogate Model, Water, Pretreatment, Reverse Osmosis

## INTRODUCTION

There is growing interest in transforming the linear water economy to a circular water economy, where wastewater is treated for reuse while contaminants are recovered as valuable products [1]. This paradigm shift would significantly reduce freshwater withdrawals and wastewater discharges to the environment and could enable distributed treatment, minimizing transportation costs and externalities. However, technological advances are needed to reduce the energy use and cost of high-recovery treatment trains, which are essential for circular water treatment and minimizing brine generation in inland water reuse applications.

Modeling high-recovery treatment trains is challenging due to complex water chemistry phenomena. As more water is recovered, the concentration of dissolved solids will eventually reach saturation, and the solids will precipitate out of solution. This precipitation can result in mineral scaling on the surfaces of equipment and damage them [2, 3]. While mineral scaling can be prevented or delayed through pretreatment, traditionally with chemical precipitation or the addition of antiscalants, it is critical to explicitly incorporate precipitation and mineral scaling predictions in technoeconomic assessments in order to relate the extent of pretreatment with the water recovery of the system. However, the complexity of these phenomena and limitations on data have



historically hindered the incorporation of detailed chemistry into process-scale water treatment models, particularly for novel, high-recovery systems.

We previously developed a framework for integrating detailed water chemistry into the design of water treatment systems using data-driven surrogates [4]. We demonstrated that scaling mitigation using chemical addition is a substantial component of process cost and a requirement for achieving high recoveries in desalination treatment trains, and that chemical pretreatment must be explicitly accounted for in process design. However, despite a significant body of work on scaling and chemical pretreatment [5, 6], little has been done on assessing the techno-economic implications of different chemical pretreatment alternatives within the context of end-to-end water treatment train optimization. This gap in the literature hinders the development of comprehensive optimization strategies that can effectively integrate chemical pretreatment with other treatment processes to achieve cost-effective desalination.

In this work, we begin to address this knowledge gap by exploring the techno-economic implications of one important chemical pretreatment decision in high-recovery desalination: pH control. The pH of water entering a reverse osmosis (RO) treatment train impacts both mineral-scale formation and membrane longevity, making pH control critical. Using WaterTAP [7], an open-source tool for modeling desalination and water treatment systems, we conduct an in-depth operational and economic assessment of post-softening pH control using acidification in water treatment. We compare three commonly used acids,  $H_2SO_4$ ,  $HCl$ , and  $CO_2$ , for pH control based on system cost, system operation, and safety for an emerging high-recovery desalination technology. Our results show that the choice of pH control acid can significantly impact costs and should account for the feedwater and acid chemistries, highlighting why water chemistry incorporation at the design stage of water treatment systems is critical. Through our analysis, we provide qualitative insights to guide decision-making on achieving high-recovery desalination.

## METHODOLOGY

This section describes the desalination train, our modeling approach, and our optimization problem formulation for cost-optimal system design.

### System Description

We investigate the impact of pH control choice on high-recovery treatment trains for brackish and seawater sources. We consider a medium-sized plant (5000  $m^3/day$ ); feedwater compositions are provided in Table 1.

Fig. 1 shows the schematic for our desalination train. Our proposed high-recovery treatment train uses high-

pressure reverse osmosis (HPRO) technology, an emerging process for water desalination that allows RO systems to operate at significantly higher pressures than the 85-bar limit of traditional seawater RO, thus enabling higher recoveries [8, 9]. Recent work has shown that HPRO can be more energy efficient and cheaper than other high-recovery alternatives [8, 10]. For our case study, we explore the viability of HPRO technology with chemical pretreatment.

Current seawater RO systems are limited to 50% recovery [11]; pretreatment allows the desalination system to increase water recovery before mineral scaling occurs. Our chemical pretreatment is done in two steps: softening followed by acidification [5]. First, soda ash ( $Na_2CO_3$ ) addition softens the water by removing calcium ions as  $CaCO_3$ . Next, acid addition lowers the pH. To assess the impact of pH control on HPRO performance, we consider three acidification alternatives:  $CO_2$ ,  $H_2SO_4$ , and  $HCl$ .

The pretreated water goes through RO/HPRO, where clean water is produced. Energy is recovered from the high-pressure waste via an energy recovery device (ERD), improving the energy efficiency of the process.

**Table 1:** Plant feedwater conditions

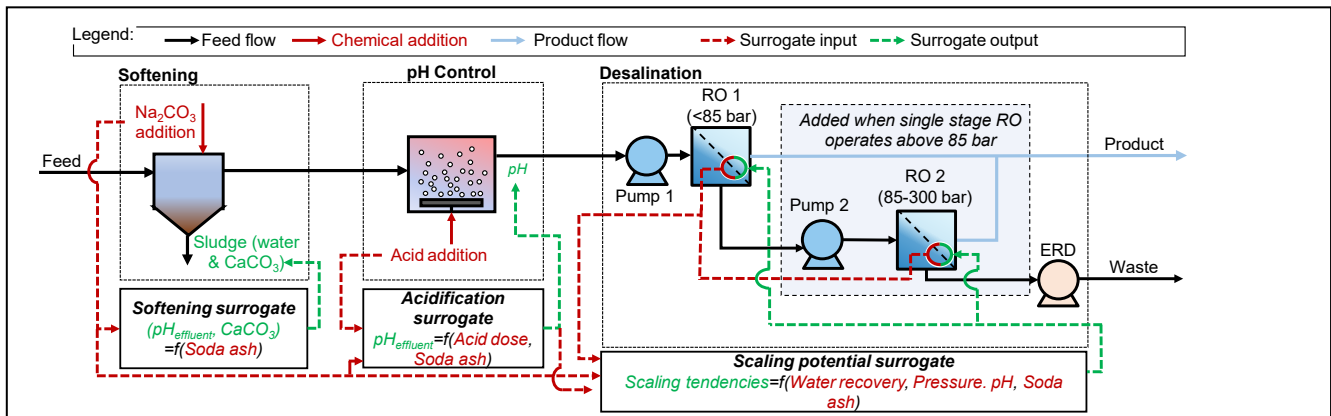
	Brackish	Seawater
Capacity ( $m^3/d$ )	5000	5000
Temperature ( $^{\circ}C$ )	25	25
Pressure (bar)	1.013	1.013
$Na^+$ (mg/L)	739	10556
$K^+$ (mg/L)	9	380
$Ca^{2+}$ (mg/L)	258	400
$Mg^{2+}$ (mg/L)	90	1262
$Cl^-$ (mg/L)	870	18980
$SO_4^{2-}$ (mg/L)	1011	2649
$HCO_3^-$ (mg/L)	385	140
pH	7.07	7.56

### HPRO Modeling

The HPRO system model was developed using WaterTAP, an open-source Python framework for the simulation and optimization of water treatment trains, built upon the Pyomo-based IDAES Integrated Platform [12]. We develop the HPRO system as a greybox model, combining mechanistic models for desalination components with surrogate models for water chemistry. The surrogates compute functions for the desired chemical phenomena that are added to the EO model formulation.

Physics-based models of the desalination components (RO, pumps, and ERD) are implemented using WaterTAP's model library which includes unit and costing models for the components. Detailed equations for each of the components may be found in Atia et. al. [13].

Detailed water chemistry and thermodynamic information are necessary to predict the impact of chemical



**Figure 1:** Schematic of the high-recovery treatment train showing surrogate insertion points. For softening and acidification, the surrogates predict the effluent pH and solids concentration directly from the decision variables (pretreatment chemical doses). For scaling, the surrogates predict scaling tendencies at the outlets of the RO units, with the inputs based on the upstream conditions and the RO operating conditions (pressure, recovery).

**Table 3:** Cost parameters for softening and acidification.

Chemical	Capital cost	Operating costs	Source(s)
$\text{Na}_2\text{CO}_3$	$12985x^{0.5901}$ ; $x = \text{lb/day}$	Chemical cost: $\$0.19/\text{kg}$	[16, 17]
$\text{H}_2\text{SO}_4$	$-0.0029x^2 + 48.434x + 22,648$ ; $x = \text{gpd}$	Chemical cost: $\$0.12/\text{kg}$	[18-20]
HCl	$-0.0029x^2 + 48.434x + 22,648$ ; $x = \text{gpd}$	Chemical cost: $\$0.17/\text{kg}$	[18, 21]
$\text{CO}_2$	$\text{CO}_2$ liq. feed system: $9 \times 10^{-8}x^3 - 0.001x^2 + 42.578x + 130,812$ ; $x = \text{lb/day}$	Chemical cost: $\$0.24/\text{kg}$	[18-19, 22]
	$\text{CO}_2$ Basin: $4 \times 10^{-9}x^3 - 2 \times 10^{-4}x^2 + 10.027x + 19,287$ ; $x = \text{ft}^3$	Electricity: $0.11 \text{ kWh}/\text{kg CO}_2$	

pretreatment in membrane systems. We incorporated the required water chemistry information into our WaterTAP model using surrogate models. As shown in Fig. 1, surrogates are used to represent three types of chemical phenomena in our process:

1. Softening – softening precipitates  $\text{CaCO}_3$  from solution, changing the feedwater pH. The surrogate predicts the  $\text{CaCO}_3$  concentration and pH as functions of the  $\text{Na}_2\text{CO}_3$  concentration added to the feedwater. Distinct surrogates were developed for brackish water and seawater.
2. Acidification – the acidification process regulates the pH for desalination. For each acid ( $\text{CO}_2$ ,  $\text{H}_2\text{SO}_4$ , HCl), we develop surrogates that predict post-acidification pH based on  $\text{Na}_2\text{CO}_3$  and acid dose.
3. Mineral scaling – mineral scale formation limits the water recovery attainable in membrane systems. We develop surrogates to represent the scaling tendency (ST), which is the thermodynamic driving force for precipitation [2] and predicts whether scale formation will occur under specific conditions. A

value of  $\text{ST} > 1$  indicates that scaling will occur. We monitor scaling at the RO outlet, with the scaling tendency dependent on the pretreatment conditions ( $\text{Na}_2\text{CO}_3$  and acid doses) and RO operation (recovery, RO pressure). The scalants of interest are calcite ( $\text{CaCO}_3$ ), gypsum ( $\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$ ) and  $\text{CaSO}_4$ .

Water chemistry data is obtained from OLI [14], a water chemistry software that provides first-principles-based equilibrium calculations and chemistry models for electrolyte and non-electrolyte streams. To generate surrogate training data, we sweep across the decision variables for each acid; ranges for the decision variables for each acid and water source are shown in Table 2.

We used the PySMO software [15], part of the IDAES-IP, to generate surrogates for the chemical phenomena of interest. PySMO provides several surrogate methods, including radial basis functions (RBFs), Gaussian process models, and polynomial regression. For this application, we used RBFs because they had the best performance and are, in our experience, robust.

## Optimization Problem

For each water source and acid choice, the objective is to minimize the levelized cost of water (LCOW) of the system:

$$\min_d LCOW = \frac{CRF * \text{capital cost} + \text{annual operating cost} [\$/\text{yr}]}{\text{Treated water volume} [m^3/\text{yr}]} \quad (1)$$

where  $d$  is the vector of decision variables ( $\text{Na}_2\text{CO}_3$  dose, acid dose, water recovery, and RO pressure), and CRF is the capital recovery factor (set at 0.1).

Constraints on the optimization problem include:

- equality constraints representing the mechanistic models of the HPRO train components (pumps, RO, ERD).
- inequality constraints representing operational limitations of RO systems. The inlet crossflow velocity is limited to a maximum of 25 cm/s, the minimum allowable salt rejection in the RO is constrained to 98%, and maximum operating pressure of RO1 is limited to 85 bar.
- inequality constraints that ensure that no scale formation occurs based on the scaling tendency surrogates  $\overline{ST}_s$ :

$$\overline{ST}_s(d) \leq 1 \quad \forall s \in [\text{CaSO}_4, \text{CaCO}_3, \text{CaSO}_4 \cdot 2\text{H}_2\text{O}] \quad (2)$$

The optimization problems are solved with IPOPT [26] using the MA27 linear solver in WaterTAP.

**Table 2:** Decision variable ranges and surrogate metrics. For the softening and acidification surrogates, we present the  $R^2$  and maximum absolute errors.

Variable	Range	
	Brackish	Seawater
$\text{Na}_2\text{CO}_3$ , mg/L	0-750	0-1200
Pressure, bar	10-110	50-300
Recovery	50-90%	50-87%
$\text{CO}_2$ , mg/L	0-300	0-50
$\text{H}_2\text{SO}_4$ , mg/L	0-150	0-50
HCl, mg/L	0-150	0-50
Surrogates	Metrics ( $R^2$ , MaxAE)	
Softening pH	1.00, 0.03	1.00, 0.02
Softening $\text{CaCO}_3$ , mg/L	1.00, 7.59	1.00, 5.07
Acidification pH ( $\text{CO}_2$ )	0.99, 0.01	0.99, 0.05
Acidification pH ( $\text{H}_2\text{SO}_4$ )	0.99, 0.01	0.99, 0.04
Acidification pH (HCl)	0.99, 0.01	0.99, 0.03
Min ST classification accuracy (%)	>99.2	>99.2

Cost parameters for the different pretreatment options are shown in Table 3. The capital cost of  $\text{CO}_2$  is made up of two components: the cost of the feed equipment and the cost of the recarbonation basin. All costs

were converted to 2018 dollars using the CEPCI cost index. Acid purities of 99.5%, 93%, and 30% are considered for  $\text{CO}_2$ ,  $\text{H}_2\text{SO}_4$  and, HCl respectively [6, 16, 23], while the liquid  $\text{CO}_2$  is assumed to be at  $-18^\circ\text{C}$  and 20.4 atm [5]. It should be noted that the capacities in our optimization problem are treated in a continuous manner, with the installed capacities for pretreatment assumed to be equal to the minimum capacities required for daily operation.

## RESULTS

### Surrogate Performance

Table 2 summarizes our surrogate performance. For the softening process, RBF models with 13 terms were found to sufficiently capture the curvilinear relationship between the input ( $\text{Na}_2\text{CO}_3$  dose) and outputs ( $\text{CaCO}_3$  precipitation and pH), with  $R^2 \approx 1$  and the mean absolute errors below 1 mg/L and 0.004 for both water sources. The maximum prediction error is 7.59 mg/L for precipitation and 0.03 for pH.

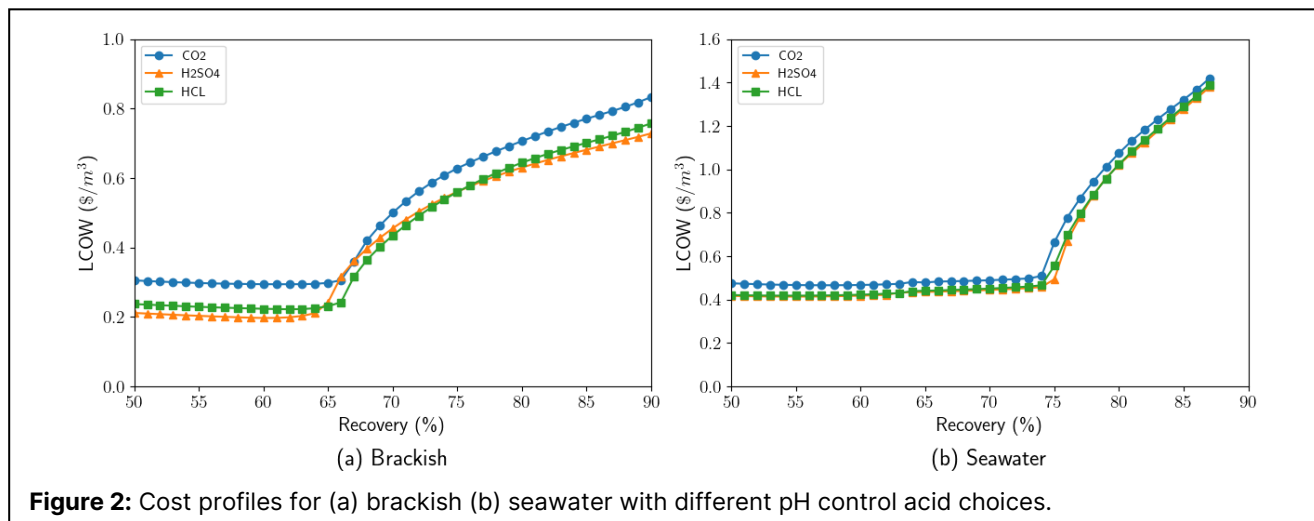
For post-acidification pH, the final RBF surrogates are trained on 100 training points, with the models predicting within  $\pm 0.05$  (less than 1%) of the true pH values observed from OLI. This level of accuracy makes the surrogates useful for comparing directly between the pH values of the different acids considered.

For mineral scaling, the most important factor is accurately determining whether or not scaling will occur. Thus, the critical performance metric for the scaling tendency surrogates is the accuracy in classifying potential operating scenarios into scaling or non-scaling. As shown in Table 2, our scaling tendency surrogates have >99% accuracy in discriminating between scaling/non-scaling cases. This means that our surrogates are accurate enough to provide the optimization model with accurate information about the operational feasibility.

### Impact of Acid Choice on Cost

Figure 2 shows the optimal cost profiles for the different acids across a range of water recoveries for the two water sources. Our model results show that (a) there is a marked increase in cost beyond 66% and 74% recovery for brackish water and seawater, respectively; (b) the choice of acid for pH control has a greater impact on brackish water than seawater.

The cost profiles follow the same pattern: a regime where the cost profile is essentially flat, followed by a regime where the costs increase with recovery in a curvilinear manner. The flat regime corresponds to a region in which only pH control is required to prevent mineral scaling, and no softening is required ( $\text{Na}_2\text{CO}_3$  dose=0). As recovery increases, however, pH control becomes insufficient to prevent mineral scaling, and Ca removal using softening becomes necessary. At this point, the LCOW increases rapidly, with softening costs dominating the



**Figure 2:** Cost profiles for (a) brackish (b) seawater with different pH control acid choices.

cost profile. Softening begins at about 64-66% and 74% for brackish and seawater, respectively.

CO<sub>2</sub> is the most expensive acid choice for both water cases. For brackish water, the cost difference between CO<sub>2</sub> and H<sub>2</sub>SO<sub>4</sub> is as high as 49% in the no-softening region. For seawater, the difference between the costs of the three acids is much lower, particularly at the higher recoveries. CO<sub>2</sub> is still the most expensive, being almost 11.5% higher than H<sub>2</sub>SO<sub>4</sub> in the no-softening regime and up to 5% higher in the high recovery, softening regime. The difference between the magnitudes of the effect of pH control on costs in our two water sources emphasizes the fact that feed water chemistry plays an important role in the acid choice for pH control.

Comparing the cost profiles for HCl and H<sub>2</sub>SO<sub>4</sub> for the Brackish water case yields interesting results. We find that H<sub>2</sub>SO<sub>4</sub> is cheaper than HCl at low recoveries (when softening is not needed) and high recoveries (> 76%). This finding agrees with the literature [6]. However, our results also show that in the mid-recovery regime of 66-76% (at low soda ash dosing requirements), pH control with HCl is cheaper. This results from sulfate-based scale formation that is promoted by the addition of H<sub>2</sub>SO<sub>4</sub> and is discussed in detail in the section below.

It should be noted that while our results consistently show CO<sub>2</sub> as the most expensive of the three acidification options, we do not consider the cost of handling, monitoring, and extra maintenance costs necessary when using H<sub>2</sub>SO<sub>4</sub> and HCl, which may be significant due to their corrosiveness. These additional costs, when combined with the uncertain nature of CO<sub>2</sub> costs [4], suggest that CO<sub>2</sub> may be cost-competitive with HCl and H<sub>2</sub>SO<sub>4</sub> under best-case cost scenarios and thus should remain under consideration as an acidification option.

## Impact on HPRO Operation

### Softening

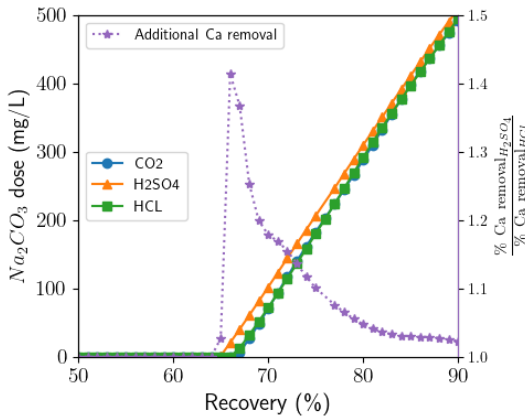
Fig. 3 shows the optimal Na<sub>2</sub>CO<sub>3</sub> doses for the

brackish water case. The softening profiles mirror the two-regime nature of the cost profiles – a flat regime with no softening, followed by a regime where more Na<sub>2</sub>CO<sub>3</sub> is increasingly required to reach higher recoveries. CO<sub>2</sub> and HCl require the same extent of water softening (same soda ash dose, resulting in the same amount of calcium removal), while use of H<sub>2</sub>SO<sub>4</sub> requires a higher degree of softening. This is because H<sub>2</sub>SO<sub>4</sub> increases the sulfate concentration of the stream, which favors sulfate-based scaling. Thus, additional calcium ion removal is required to counteract the sulfate ion increase and prevent gypsum scaling in the RO membrane. The additional calcium removal required when using H<sub>2</sub>SO<sub>4</sub> depends on the recovery, varying from 2-41% across the recovery range. The challenge with gypsum scaling when using H<sub>2</sub>SO<sub>4</sub> also means that Na<sub>2</sub>CO<sub>3</sub> dosing for gypsum scaling control starts earlier than with the other acids. This finding agrees with the literature which suggests that using HCl may be preferred when calcium sulfate or barium sulfate scaling is a concern [6].

The increased soda ash consumption to prevent gypsum scaling when using H<sub>2</sub>SO<sub>4</sub> for acid control is responsible for the high LCOW values (compared to HCl) observed with H<sub>2</sub>SO<sub>4</sub> between 66-76% for the brackish water case. At low Na<sub>2</sub>CO<sub>3</sub> doses, the additional softening costs incurred when using H<sub>2</sub>SO<sub>4</sub> have a significant impact on the LCOW. However, past 76%, the cost impact of the additional softening required with H<sub>2</sub>SO<sub>4</sub> is overshadowed by the overall cost of pH control pretreatment required to satisfy recovery targets, and H<sub>2</sub>SO<sub>4</sub> becomes the cheapest option due to its higher purity and lower operating costs relative to HCl.

Our results show that softening is primarily required to control gypsum scaling: the recovery at which Na<sub>2</sub>CO<sub>3</sub> dosing starts corresponds to when the gypsum scaling tendency constraint becomes active in the optimization problem ( $\widehat{ST}_{gypsum} = 1.0$ ). Our results also show that the softening step cannot be designed independently of the

downstream chemical operations (e.g. acidification). For the seawater case, softening begins at 74%, with  $\text{Na}_2\text{CO}_3$  consumption levels the same for all acids.



**Figure 3.**  $\text{Na}_2\text{CO}_3$  concentration profile (left axis) and additional Ca removal required when using sulfuric acid (purple, right axis) for Brackish case.

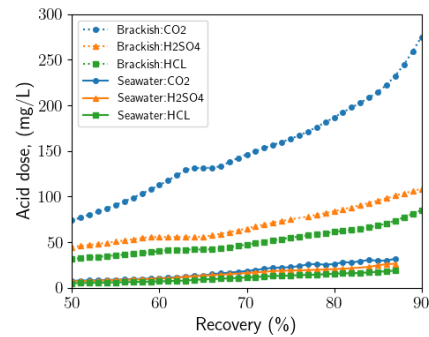
### pH control

Fig. 4 shows acid dose profiles for the two water sources. Seawater requires significantly less acid addition for pH control than brackish water despite its lower starting pH. The high acid requirement in the brackish case is due to its high bicarbonate concentration compared to the seawater source (see Table 1), which produces a buffering effect [24]. The cost-optimal acid doses required for the brackish water cases are between 4-6 times greater than that for seawater when using HCl/ $\text{H}_2\text{SO}_4$ , and 7-10 times greater with  $\text{CO}_2$ . The low acid consumption in the seawater case contributes to why the choice of acid has a smaller effect on the cost profiles. For both water sources, control with  $\text{CO}_2$  requires the highest acid dosing, requiring more than three times the HCl dosing at the highest recovery in the brackish case. The dosage requirements reflect the strengths of the acids: HCl and  $\text{H}_2\text{SO}_4$  are strong acids ( $K_a > 1$ ) and thus require lower acid dosing, while the carbonic acid formed by  $\text{CO}_2$  is a weak acid ( $K_a < 1$ ), meaning that higher concentrations are required. Our analysis of the scaling behavior indicates that pH adjustment is the preferred mechanism for controlling calcite scaling, with the calcite scaling tendency constraint active ( $\widehat{ST}_{\text{calcite}} = 1.0$ ) in the optimization problem over the entire recovery range.

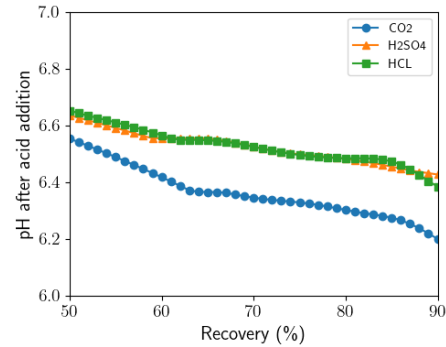
Fig 5. shows the pH exiting the pretreatment stage after acid addition for the Brackish case. Again, we find one acid behaving differently from the other two. HCl and  $\text{H}_2\text{SO}_4$  drop the pH to similar values across the range. However, using  $\text{CO}_2$  necessitates dropping the pH lower. We observed the same pattern in the seawater case (not shown). The reason for this is that  $\text{CO}_2$  addition increases the carbonate ion concentration of the stream, favoring

more calcite scaling. Since pH control is the mechanism for controlling calcite scaling in the system, the pH must be lowered more than in the HCl and  $\text{H}_2\text{SO}_4$  cases to combat the effect of the higher carbonate concentration.

The results of our analysis of the pretreatment process show that while choosing  $\text{H}_2\text{SO}_4$  for pH control impacts gypsum scaling and the operation of the softening step, choosing  $\text{CO}_2$  impacts calcite scaling and the pH of the feedwater to the desalination train. These findings on  $\text{CO}_2$  and  $\text{H}_2\text{SO}_4$  underscore why water chemistry incorporation at the design stage in water treatment is critical.



**Figure 4.** Optimal acid doses for brackish and seawater.



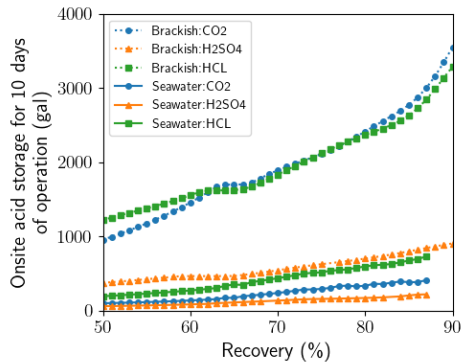
**Figure 5.** pH post-acidification for Brackish case.

### Safety considerations

Safety is a critical factor to consider when dealing with acids in water treatment due to the requirements for onsite storage. Fig. 6 shows onsite acid storage requirements for ten days of plant operation (ten-day storage assumption adopted from [18]). Despite requiring the lowest acid dosing requirements (Fig. 4), HCl requires the highest storage capacity in the seawater case and the joint highest in the brackish case. This is due to its low purity compared to the other two alternatives.  $\text{H}_2\text{SO}_4$  requires the lowest volume for onsite storage due to its high density and typical purity at 93%.  $\text{H}_2\text{SO}_4$  also causes less fuming to the atmosphere than HCl, which results in less corrosion to surrounding metal components [6]. Thus, while the LCOW results are comparable,  $\text{H}_2\text{SO}_4$  is safer than HCl due to its significantly lower storage



requirements and lower corrosiveness. However, of the alternatives considered, CO<sub>2</sub> is the safest choice for pH control. Unlike H<sub>2</sub>SO<sub>4</sub> and HCl, it is non-corrosive, reduces pipe clogging, and requires much lower handling and monitoring costs (not modeled in this work) [25].



**Figure 6.** Onsite acid storage for 10 days of operation.

## CONCLUSION

We present an assessment of the impact of pH control pretreatment choices on the cost and operation of a medium-sized high-recovery RO desalination plant by comparing three common acidification options (H<sub>2</sub>SO<sub>4</sub>, HCl, and CO<sub>2</sub>). We find that the impact of the acid choice on the cost can vary widely depending on the water source, with the seawater LCOW being significantly less sensitive to the acid choice than our brackish water (11.5% vs 49%). Our results show that the choice of pH control acid should account for the properties of the feedwater (e.g., buffer capacity), the chemistry of the acids (e.g., sulfate vs. bicarbonate anions), and the scalants of concern. Water sources with calcite scaling as a primary concern may prefer to avoid CO<sub>2</sub> acidification, while H<sub>2</sub>SO<sub>4</sub> may exacerbate sulfate scaling challenges. Our findings underscore why water chemistry incorporation at the design stage of water treatment systems is critical. The methodology used in this work can easily be applied to evaluate the benefits of process changes and innovations that allow higher maximum allowable scaling tendencies or reduce the cost of chemical precipitation pretreatment. This includes the effect of antiscalants, which can be incorporated into our methodology by increasing the tendency threshold (i.e., RHS of Eq. 2) beyond 1.0, as demonstrated in a previous work [4].

In the future, we will extend this analysis to include an assessment of alternative chemical softening options. We will also consider the impact of electrified pretreatment as an alternative to chemical pretreatment.

## ACKNOWLEDGEMENTS

This material is based upon work supported by the National Alliance for Water Innovation (NAWI), funded by

the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy (EERE), Advanced Manufacturing Office, under Funding Opportunity Announcement Number DE-FOA-0001905. The version of IPOPT implemented in this work was compiled using HSL, a collection of Fortran codes for large-scale scientific computation. See <http://www.hsl.rl.ac.uk>.

## DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the U.S. Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof or the Regents of the University of California.

## REFERENCES

1. M.S. Mauter, P.S. Fiske. Desalination for a circular water economy. *Energy Environ. Sci.* 13:3180–3184 (2020). <https://doi.org/10.1039/D0EE01653E>
2. A. Matin, F. Rahman, H.Z. Shafi, S.M. Zubair. Scaling of reverse osmosis membranes used in water desalination: Phenomena, impact, and control; future directions. *Desalination.* 455: 135–157 (2019). <https://doi.org/10.1016/j.desal.2018.12.009>
3. T. Tong, A.F. Wallace, S. Zhao, Z. Wang. Mineral scaling in membrane desalination: Mechanisms, mitigation strategies, and feasibility of scaling-resistant membranes. *J. Membr. Sci.* (2019). <https://doi.org/10.1016/j.memsci.2019.02.049>
4. O. Amusat, A. Atia, A. Dudchenko, T. Bartholomew. Modeling framework for cost-optimization of process-scale desalination systems with mineral scaling and precipitation. *ACS ES&T Eng.* <https://doi.org/10.1021/acsestengg.3c00537>
5. L.K Wang, J. S. Wu, N.K. Shammass, K. Nazih D.A. Vaccari. Recarbonation and Softening. *Physicochemical Treatment Processes. In: Physicochemical Treatment Processes. Handbook*

- of Environmental Engineering, vol 3. Ed: Wang, L.K., Hung, Y.T., Shammam, N.K. (eds) Humana Press (2005). <https://doi.org/10.1385/1-59259-820-x:199>
6. Nitto. Chemical pretreatment for RO and NF. Technical Application Bulletin , No. 111 (2022). <https://membranes.com/wp-content/uploads/Documents/application-report/TAB111.pdf>
  7. WaterTAP contributors, WaterTAP: An open-source water treatment model library. Version 0.6. Sponsored by California Energy Commission, National Alliance for Water Innovation, and USDOE., (n.d.). <https://github.com/watertap-org/watertap>.
  8. D.M. Davenport, A. Deshmukh, J.R. Werber, M. Elimelech. High-Pressure Reverse Osmosis for Energy-Efficient Hypersaline Brine Desalination: Current Status, Design Considerations, and Research Needs, Environ. Sci. Technol. Lett. 5: 467–475 (2018). <https://doi.org/10.1021/acs.estlett.8b00274>.
  9. J. Wu, B. Jung, A. Anvari, S. Im, M. Anderson, X. Zheng, D. Jassby, R.B. Kaner, D. Dlamini, A. Edalat, E.M.V. Hoek. Reverse osmosis membrane compaction and embossing at ultra-high pressure operation, Desalination. 537: 115875 (2022). <https://doi.org/10.1016/j.desal.2022.115875>.
  10. A.V. Dudchenko, T.V. Bartholomew, M.S. Mauter. High-impact innovations for high-salinity membrane desalination, Proceedings of the National Academy of Sciences. 118:e2022196118 (2021). <https://doi.org/10.1073/pnas.2022196118>
  11. T. Tong, M. Elimelech. The global rise of zero liquid discharge for wastewater management: drivers, technologies, and future directions. Environ. Sci. Technol. 50: 6846–6855 (2016).
  12. A. Lee, J.H. Ghouse, J.C. Eslick, C.D. Laird, J.D. Sirola, M.A. Zamarripa, D. Gunter, J.H. Shinn, A.W. Dowling, D. Bhattacharyya, L.T. Biegler, A.P. Burgard, D.C. Miller. The IDAES process modeling framework and model library—Flexibility for process simulation and optimization. J. Adv. Manuf 3: e10095 (2021). <https://doi.org/10.1002/amp2.10095>.
  13. A.A. Atia, J. Allen, E. Young, B. Knueven, T.V. Bartholomew, Cost optimization of low-salt-rejection reverse osmosis. J. Desal. 551:116407 (2023). <https://doi.org/10.1016/j.desal.2023.116407>
  14. P. Wang, A. Anderko, R.D. Young, A speciation-based model for mixed-solvent electrolyte systems, Fluid Phase Equilibria. 203: 141–176 (2002). [https://doi.org/10.1016/S0378-3812\(02\)00178-4](https://doi.org/10.1016/S0378-3812(02)00178-4)
  15. PySMO: Python-based Surrogate Modelling Objects — IDAES v1.13.2, (n.d.). [https://idaes-pse.readthedocs.io/en/stable/explanations/modeling\\_extensions/surrogate/api/pysmo/index.html](https://idaes-pse.readthedocs.io/en/stable/explanations/modeling_extensions/surrogate/api/pysmo/index.html). (Accessed November 16, 2023).
  16. McGivney W, Kawamura S. Cost Estimating Manual for Water Treatment Facilities. J.Wiley&Sons (2008).
  17. Miara A, Talmadge M, Sitterley K, Evans A, Huang Z, Macknick J, McCall J, Kurup P, Akar S, Van Allsburg K, Stokes-Draut J, Bartholomew T, Lee A, Gingerich D, WaterTAP3 (The Water Technoeconomic Assessment Pipe-Parity Platform), (2021). <https://doi.org/10.11578/DC.20210709.1>
  18. Sharma JR. (2010). Use of cost equations and cost indexes for preliminary cost estimates of water treatment plants.” M.S. thesis, Dept. of Civil Eng., Univ. of Texas at Arlington, Arlington, TX (2010).
  19. Sharma, JR, Najafi M/ Qasim SR. Preliminary Cost Estimation Models for Construction, Operation, and Maintenance of Water Treatment Plants, *J. Infrastruct. Syst.* Vol. 19, No. 4 :451-464 (2013). [10.1061/\(ASCE\)IS.1943-555X.0000155](https://doi.org/10.1061/(ASCE)IS.1943-555X.0000155)
  20. Intratec. <https://www.intratec.us/chemical-markets/sulfuric-acid-price>. Accessed 2 Aug. 2023.
  21. Intratec. <https://www.intratec.us/chemical-markets/hydrochloric-acid-price>.
  22. R.C. Gumerman, R.L. Culp, S.P. Hansen, Estimating water treatment costs: Volume 2. Cost curves applicable to 1 to 200 mgd treatment plants, Drinking Water Research Division, Municipal Environmental Research Laboratory, US EPA, EPA-600/2-79-162b. (1979).
  23. Intratec. Intratec Primary Commodity Prices Commodities Specifications. <https://cdn.intratec.us/docs/methodologies/ipcp-commodities-specifications.pdf>.
  24. L. Birnhack, N. Voutchkov, O. Lahav. Fundamental chemistry and engineering aspects of post-treatment processes for desalinated water—A review. Desalination 273: 6–22 (2011).
  25. I.S. Al-Mutaz, M.A. Al-Ghunaimi, pH control in water treatment plant by the addition of carbon dioxide, in: The IDA World Congress on Desalination and Water Reuse, Bahrain, 2001: pp. 1–12.
  26. A. Wachter, L.T. Biegler, L.T. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. Math. Prog. 106,25–57(2005).

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Hybrid Rule-based and Optimization-driven Decision Framework for the Rapid Synthesis of End-to-End Optimal (E2EO) and Sustainable Pharmaceutical Manufacturing Flowsheets

Yash Barhate<sup>a</sup>, Daniel Casas-Orozco<sup>a</sup>, Daniel J. Laky<sup>a</sup>, Gintaras V. Reklaitis<sup>a</sup>, Zoltan K. Nagy<sup>a\*</sup>

<sup>a</sup> Davidson School of Chemical Engineering, Purdue University, West Lafayette, IN, United States

\* Corresponding Author: [znagy@purdue.edu](mailto:znagy@purdue.edu).

## ABSTRACT

In this paper, a hybrid heuristic rule-based and deterministic optimization-driven process decision framework is presented for the analysis and optimization of process flowsheets for end-to-end optimal (E2EO) pharmaceutical manufacturing. The framework accommodates various operating modes, such as batch, semi-batch and continuous, for the different unit operations that implement each manufacturing step. To address the challenges associated with solving process synthesis problems using a simulation-optimization approach, heuristic-based process synthesis rules are employed to facilitate the reduction of the superstructure into smaller sub-structures that can be more readily optimized. The practical application of the framework is demonstrated through a case study involving the end-to-end continuous manufacturing of an anti-cancer drug, lomustine. Alternative flowsheet structures are evaluated in terms of the sustainability metric, E-factor while ensuring compliance with the required production targets and critical product quality attributes.

**Keywords:** Process Synthesis, Optimization, Modelling and Simulations, Derivative-Free Optimization, Industry 4.0

## INTRODUCTION

The ongoing technological transformations in pharmaceutical manufacturing, driven by emerging paradigms such as quality-by-design (QbD) and quality-by-control (QbC), underscore the importance of model-based digital design tools for informed decision-making in process design and operation [1], [2]. These tools play a crucial role in enabling the comprehensive analysis and optimization of various unit operations as well as integrated process flowsheets to improve manufacturing efficiency while ensuring regulatory compliance and product quality [3], [4]. As the pharmaceutical industry is increasingly leaning towards transitioning from batch to continuous operations, there is a growing demand for modeling tools capable of flexibly simulating diverse operating modes, such as end-to-end batch (E2EB) or continuous (E2EC) flowsheets, as well as hybrid processing schemes, which are often the desired alternatives to

achieve end-to-end optimal (E2EO) manufacturing. To that end, PharmaPy a user-friendly, open-source Python-based tool that has the capabilities of configuring and simulating various manufacturing setups, including batch, continuous, and hybrid systems (containing both batch and continuous unit operations) has been developed as reported in [5].

The optimization of flowsheets, particularly those involving unit operations with diverse operating modes, presents a significant challenge. For pharmaceutical process development, this is crucial, especially during the selection and design of an optimal manufacturing scheme where the design needs to be based on the consideration and analysis of various techno-economic, sustainability, and regulatory compliance considerations rather than relying on pre-existing assumptions [5].

Traditional approaches for solving process flowsheet synthesis problems involve the formulation of a superstructure network, typically solved using mixed-

integer nonlinear optimization (MINLP) or generalized disjunctive programming (GDP) [6], [7]. However, the prevalent use of equation-oriented GDP/mixed-integer optimization tools is often limited to flowsheets modeled with simplified unit operation models and approximate thermodynamic property relations [8]. For flowsheets modeled with rigorous unit operation models, a simulation-optimization-based framework is frequently adopted. Nevertheless, addressing complex process synthesis problems through a simulation-optimization approach presents challenges, either due to the need to exhaustively optimize each flowsheet resulting in high computational cost, or due to the requirement for specialized interfaces between MINLP optimization algorithms and process simulators [9].

As a viable alternative, in this study, a hybrid process-decision framework was devised for the optimization of end-to-end pharmaceutical manufacturing flowsheets, which combines a rule-based selection approach with deterministic optimization steps. By applying a set of heuristic process synthesis rules influenced by user preferences, the goal is to systematically condense the initial master superstructure, encompassing all conceivable alternatives, into smaller, more manageable sub-structures for subsequent optimization. This approach mitigates the computational complexity associated with simultaneously optimizing the entire array of flowsheets. By reducing the problem to smaller sub-structures, the framework enables the exploration of scenario-based practical situations, facilitating practical decision-making in the optimization of diverse manufacturing campaigns.

The structure of the paper is organized as follows: firstly, the framework for hierarchical process generation in the end-to-end pharmaceutical manufacturing routines and the enumeration of all possible process candidate flowsheets are presented. Subsequently, the rule-based decision system for reducing the master superstructure into smaller case-specific sub-structures is formally outlined and discussed. Finally, the efficacy of the devised framework is demonstrated through a case study focused on optimizing end-to-end pharmaceutical flowsheets for the production of the commercial active pharmaceutical ingredient (API), Lomustine.

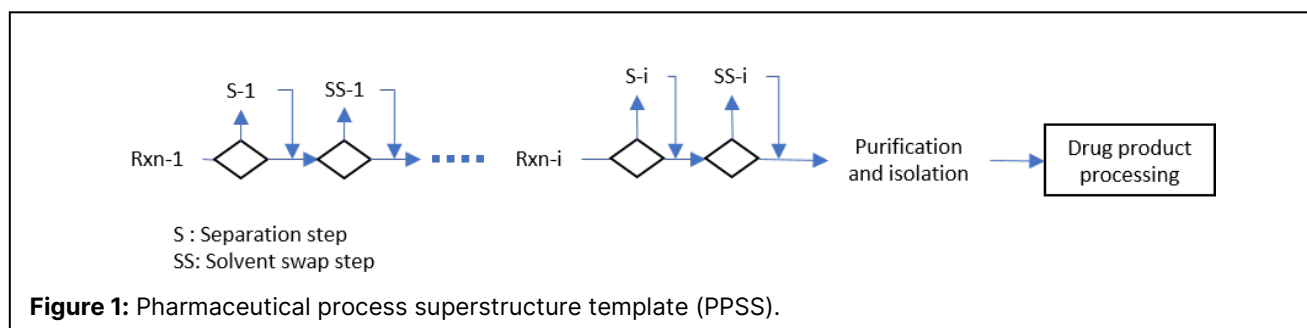
## METHODOLOGY

### Framework for process synthesis of small molecule pharmaceutical processes

The fundamental structure of a generic pharmaceutical process superstructure template (PPSS) (**Figure 1**) for the production of a small molecule drug substance in solid form consists of a sequence of processing steps including reactions, separations, and purifications before progressing to drug product manufacturing. Beginning with reaction steps, strategic decision points denoted by diamonds in **Figure 1** arise within the template, prompting consideration for the inclusion of separation or solvent swap steps before progressing to subsequent reactions. A separation step becomes essential when isolating specific reaction intermediates is required before moving to further reaction steps, while a solvent swap step is crucial when a change of solvent is needed before proceeding to subsequent reaction or purification steps. The latter part of the template addresses purification and isolation steps to obtain the desired API in solid form.

For each processing step within the generic superstructure, various alternatives exist in terms of unit operations. Furthermore, different operating mode choices such as batch, semi-batch, or continuous are also available for each unit operation. The exponential expansion of the search space over possible flowsheet alternatives becomes apparent with the increasing number of alternative unit operations and their operating modes for each step. Additionally, flowsheets featuring hybrid modes of operations necessitate the addition of supplementary units, such as holding tanks and batch-to-continuous connectors, to ensure seamless connectivity between the batch and continuous components of the flowsheets. This interplay of decision points, alternatives, and considerations underscores the complexity inherent in designing an effective process superstructure for the production of small molecule drug substances in solid form.

In combinatorial process synthesis methods, the superstructure is formulated by listing relevant alternatives for processing steps and connecting them in all possible ways. However, as previously mentioned, this leads to a vast search space for the design problem, resulting in numerous redundant and implausible configurations. The



expanded search space also increases the computational effort required for optimizing all these flowsheets [10].

In-sights-based synthesis methods, on the other hand, utilize a synthesis logic that integrates commercial and engineering insights as input to systematically generate a superstructure [10]. These insights, derived from the problem definition (which involves defining the scope in terms of raw materials, products, and process technologies) and process chemistry, guide the adaptation of the generic superstructure template into a process-specific superstructure containing well-defined processing steps relevant to the specific manufacturing process. Nevertheless, the design search space with the process-specific superstructure can still be considerable, particularly when considering alternative unit operations and their operating modes. This presents a significant computational bottleneck when optimizing all the flowsheets within the superstructure that are represented with rigorous models and optimized using a simulation-optimization framework. Herein, the heuristic rule-based synthesis logic becomes instrumental, facilitating the reduction of the superstructure into smaller, more manageable substructures that can be optimized with moderate computational efforts.

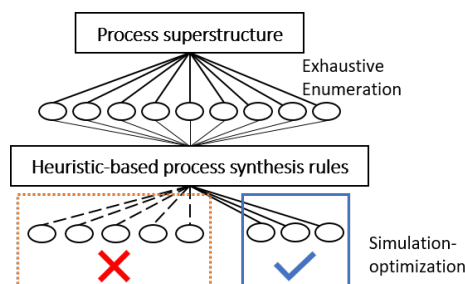
### Rule-based heuristic decision system

The heuristic process synthesis rules are formulated and generalized, taking into consideration common commercial and process engineering decision points in pharmaceutical manufacturing. The rules are categorized into five distinct groups: (1) process feasibility rules, (2) regulatory considerations, (3) equipment availability constraints, (4) experience or knowledge-based rules, and (5) scenario analysis.

*Regulatory considerations* are of utmost importance in pharmaceutical manufacturing, encompassing rules that ensure compliance with regulations and guidance set forth by agencies such as the FDA and EMA. *Equipment availability* constraints are included to anticipate practical limitations that will arise when transferring a conceptual design to existing manufacturing sites, such as the presence or absence of specific unit operations at

these sites. This category ensures that the synthesis rules are grounded in real-world scenarios. *Experience or knowledge-based* rules are formulated by integrating logical process engineering knowledge. Lastly, *scenario analysis* rules are designed to facilitate the comparison and optimization of various industrially relevant and intriguing manufacturing scenarios such as end-to-end continuous (E2EC) manufacturing, end-to-end batch manufacturing (E2EB), telescoped reaction synthesis, etc. While presenting a comprehensive list of all rules is beyond the scope of this work, **Table 1** provides illustrative examples from each category, offering a glimpse into the flexibility offered by the synthesis rules during the optimization of pharmaceutical flowsheets.

Within the framework of the synthesis rules outlined above, a specific set of rules is chosen based on user preferences and insights. These selected rules are then applied to the master superstructure to generate smaller-scale superstructure realizations as shown in Figure 2. Configurations that violate one or more synthesis rules are systematically eliminated from the search space. This exclusion of non-legal configurations streamlines the optimization process, by reducing the computational resources required to optimize different flowsheets. The explicit formulation of synthesis rules also provides a transparent and systematic procedure for both analyzing potential configurations and offering the rationale for the choices made [11].



**Figure 2:** Heuristics reduce the search space of valid superstructure layouts.

### CASE STUDY

**Table 1:** Process synthesis rules with their categories.

No.	Category	Example
1	Regulatory aspects	R-1: Isolation of ISO intermediate after Rxn1, increases drug safety
2	Feasibility rules	F-1: For semi-batch reactors, all reactions are assumed to happen in single reactor
3	Equipment availability constraints	A-1: Semi-batch evaporator is not available
4	Scenario Analysis	S-1: End-end continuous manufacturing S-2: Telescoped reaction synthesis
5	Experience or Knowledge-based rules	E-1: If conversion of reactant in step 1 is not complete, and this reactant is inert for the next reaction, then separation of the reactant is not necessary.



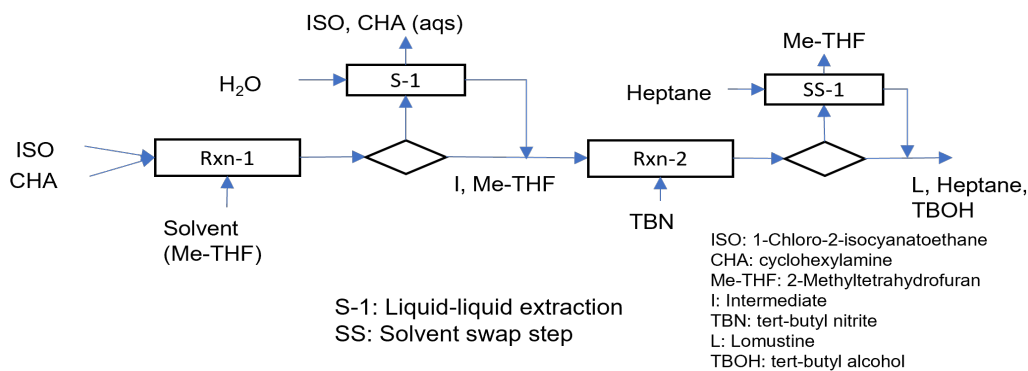


Figure 3: Process specific superstructure template for Lomustine.

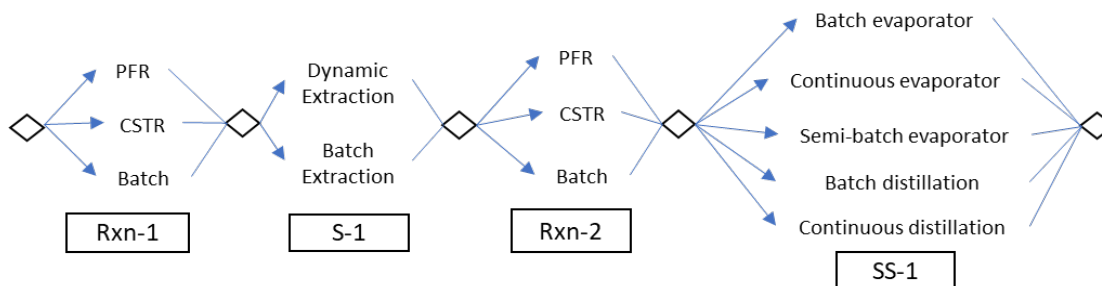


Figure 4: Exhaustive enumeration of possible flowsheet alternatives.

In this study, the above methodology is applied for the optimization of end-to-end flowsheets for the manufacturing of Lomustine. Recent research has identified efficient lab-scale continuous synthesis pathways for Lomustine, which allow replacing environmentally harmful solvents with more sustainable alternatives [12]. This case study aims to supplement the above studies, by designing end-to-end optimal processes using different separation technologies and comparing solvent use as a function of process configuration. Based on the proposed methodology, at first, the process-specific superstructure template was drafted using the process chemistry of Lomustine manufacturing as shown in Figure 3.

It should be noted that the scope of this case study is only limited to the analysis of synthesis, solvent swap, and separation processing steps. For conducting each of the processing steps shown in Figure 3, various alternative unit operations are available with varying operating modes. The process superstructure containing all possible flowsheet configurations is shown in **Figure 4**. An exhaustive enumeration of all potential flowsheets from this superstructure yields a total of 90 unique configurations. Since, the lab studies analyzed continuous manufacturing routes, from the heuristic process synthesis rule database regulatory rule R-1 and scenario analysis rule S-1 were applied to the above superstructure to screen out end-to-end continuous (E2EC) manufacturing flowsheets. With these rules, the search space is significantly narrowed down from 90 flowsheets to 8 E2EC

manufacturing flowsheets to be optimized. The next section provides a detailed description of the optimization problem aimed at optimizing the screened flowsheets of interest.

## Design Problem

The design problem is formulated as a nonlinear constrained optimization problem aimed at minimizing the overall waste produced in the process  $J(\mathbf{x})$ , which is the combination of total mass of solvent (Me-THF, Acetic acid,  $H_2O$ , and Heptane) and the total mass of unreacted reactants. Due to the sequential modular architecture of PharmaPy, solving the optimization problem required a simulation-optimization framework where obtaining gradient information is challenging. To address this, the problem was first transformed into an unconstrained optimization problem by integrating the nonlinear inequality constraints ( $g_i$ ) directly into the objective function via weighted penalties ( $\alpha_i$ ), as represented by Equation (1). This enables the solution of the problem using derivative-free or direct-search methods, and in this work, Nelder-Mead algorithm was used.

$$\min_{\mathbf{x}} J(\mathbf{x}) + \sum_{i \in I} \max(0, \alpha_i \cdot g_i)^2 \quad (1)$$

$$s. t. \text{ PharmaPy Process model}$$

$$\mathbf{x}_{lb} < \mathbf{x} < \mathbf{x}_{ub} \quad (2)$$

The process model consists of set of differential-

**Table 2:** List of inequality constraints.

$$\text{Weekly production constraint} \quad g_1(\mathbf{x}) = 1\text{kg} - \dot{F}_{SS} (C_{SS,API} - C_{solub}(273K)) \cdot t_{hor} \leq 0 \quad (3)$$

$$\text{Maximum precursor after extraction} \quad g_2(\mathbf{x}) = C_{extr,ISO} - 8\text{mol/m}^3 \leq 0 \quad (4)$$

$$g_3(\mathbf{x}) = \int_{t_{st}}^{t_f} (C_{SS,API}(t) - C_{solub}(T, C_{Me-THF}, C_{ac-acid})) dt \leq 0 \quad (5)$$

$$\text{Path constraints (Solvent switch)} \quad g_4(\mathbf{x}) = \int_{t_{st}}^{t_f} (T_{SS}(t) - T_{max}) dt \leq 0 \quad (6)$$

$$g_5(\mathbf{x}) = \int_{t_{st}}^{t_f} (0.7 - x_{SS,C7}(t)) dt \leq 0 \quad (7)$$

algebraic equations simulated within PharmaPy. The values of the weighted penalties were carefully selected based on the desired residual values for the constraints at convergence. Decision variables  $\mathbf{x}$  used for all flow-sheet optimizations are shown in **Table 3**.

**Table 3:** Decision variables used in design problem.

Step	Unit operation	Decision Variables
Rxn-1	PFR/CSTR (R01)	$C_{in,ISO}, C_{in,CHA}, V_{R01}, \tau_{R01}$
S-1	Dynamic extractor (EXTR01)	$\tau_{E01}, F_{H_2O,in}$
Rxn-2	PFR/CSTR (R02)	$C_{in,TBN}, \tau_{R02}$
SS-1	Continuous vaporizer (VAP01)	$\tau_{VAP01}, v, F_{Heptane,in}$
	Continuous distillation (DIST01)	$F_{Heptane,in}, x_{LK}, R$

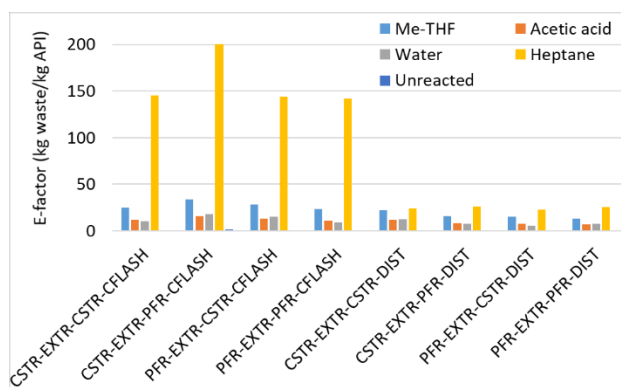
**Table 4:** List of all symbols.

Symbol	Meaning	Unit
$C_{in}$	Inlet concentration	mol/L
$V$	Volume of unit	$\text{m}^3$
$\tau$	Residence time	S
$P$	Pressure	bar
$v$	Vapor fraction	-
$F_{in}$	Inlet flowrate	mol/s
$x_{LK}$	Mole frac. of lowkey component	-
$R$	Reflux ratio	-
$F_{SS}$	Steady state output flowrate	$\text{m}^3/\text{s}$
$C_{SS}$	Steady state concentration	$\text{kg}/\text{m}^3$
$C_{solub}$	Solubility concentration	$\text{kg}/\text{m}^3$
$C_{EXTR}$	Concentration at EXTR01 outlet	mol/ $\text{m}^3$
$T_{SS}$	Steady state temperature	K
$T_{max}$	Maximum allowed temperature	K
$x_{SS,C7}$	Mole frac. of heptane at steady state	-

All the decision variables were scaled within a 0-1 interval. Furthermore, various manufacturing constraints such as weekly production targets, product quality attributes, and process operational constraints were formulated as inequality constraints listed in **Table 3**. A comprehensive list of all the symbols used can be found in **Table 4**.

## Results

Using the optimization framework outlined in the previous section, each flowsheet was optimized while satisfying all the operational constraints, product quality attributes and production targets. The comparative analysis of all optimal flowsheets in terms of waste minimization, is illustrated in Figure 5.

**Figure 5:** Bar graph showing E-factors for different flowsheets.

The bar graph shows the E-factors for various components including solvents and unreacted reactants, across all eight optimal flowsheets, calculated using following equation (8).

$$E - factor = \frac{\text{kg waste component}}{\text{kg Lomustine}} \quad (8)$$

The E-factors for the reaction section for lomustine

synthesis were like the typical E-factors encountered for synthesis steps in pharmaceutical manufacturing (25–80) [13]. Furthermore, it can be seen from the graph that flowsheets with flash separation have higher solvent usage, leading to significantly higher E-factors when compared with flowsheets containing distillation columns. This is because, the start-up times for flash-containing flowsheets were found to be larger than the flowsheets containing distillation columns, leading to a higher waste generation. Lastly, the analysis also prompts that further reduction in E-factor could be achieved by solvent recovery since solvent consumption contributes the highest among all towards waste generation.

## CONCLUSIONS

In this study, a heuristic rule-based process decision framework was developed for the optimization of end-to-end pharmaceutical flowsheets encompassing various operating modes for the different unit operations. The framework was also applied for the optimization and analysis of end-to-end continuous manufacturing flowsheets for the synthesis of Lomustine. The resulting flowsheets were effectively optimized and evaluated in terms of sustainability metric E-factor while ensuring compliance with respect to production targets and product quality attributes.

## ACKNOWLEDGEMENTS

Funding for this work was made possible, by the US Food and Drug Administration (FDA) through grant number U01FD006738. Views expressed herein do not necessarily reflect the official policies of the Department of Health and Human Services; nor does any mention of trade names, commercial practices, or organizations imply endorsement by the United States Government.

## REFERENCES

1. Lawrence X. Y., Amidon G., Khan M.A., Hoag S.W., Polli J., Raju G.K., Woodcock J. Understanding Pharmaceutical Quality by Design. *AAPS J*, 16(4), 771, (2014).
2. Su Q., Ganesh S., Moreno M., Bommireddy Y., Gonzalez M., Reklaitis G.V., Nagy Z.K. A perspective on Quality-by-Control (QbC) in pharmaceutical continuous manufacturing. *Comput Chem Eng*, 125, 216–231 (2019).
3. Casas-Orozco D., Laky DJ., Wang V., Abdi M., Feng X., Wood E., Reklaitis G.V., Nagy Z.K. Techno-economic analysis of dynamic, end-to-end optimal pharmaceutical campaign manufacturing using PharmaPy. *AIChE Journal*, 69(9), e18142, (2023).
4. Laky D.J., Casas-Orozco D., Laird C.D., Reklaitis

- G.V., Nagy Z.K. Simulation-Optimization Framework for the Digital Design of Pharmaceutical Processes Using Pyomo and PharmaPy. *Ind Eng Chem Res*, 61 (43), 16128–16140, (2022).
5. Casas-Orozco D., Laky DJ., Wang V., Abdi M., Feng X., Wood E., Laird C., Reklaitis G.V., Nagy Z.K. PharmaPy: An object-oriented tool for the development of hybrid pharmaceutical flowsheets. *Comput Chem Eng*, 153, 107408, (2021).
6. Chen Q., Grossmann I. E. Recent Developments and Challenges in Optimization-Based Process Synthesis. *Annu. Rev. Chem. Biomol. Eng.*, 8: 249–283, (2017).
7. Wu W. Z., Henao C. A., Maravelias C. T. A superstructure representation, generation, and modeling framework for chemical process synthesis. *AIChE Journal*, 62 (9), 3199–3214, (2016).
8. Navarro-Amorós M. A., Ruiz-Femenia R., Caballero J. A. Integration of modular process simulators under the Generalized Disjunctive Programming framework for the structural flowsheet optimization. *Comput Chem Eng*, 67, 13–25, (2014).
9. Corbetta M., Grossmann I. E., Manenti F. Process simulator-based optimization of biorefinery downstream processes under the Generalized Disjunctive Programming framework, *Comput Chem Eng*, 88, 73–85, (2016).
10. Quaglia A., Gargalo C. L., Chairakwongsa S., Sin G., Gani R. Systematic network synthesis and design: Problem formulation, superstructure generation, data management and solution. *Comput Chem Eng*, 72, 68–86, (2015).
11. Siirola J. J. Strategic process synthesis: Advances in the hierarchical approach. *Comput Chem Eng*, 20 (2), S1637–S1643, (1996).
12. Ewan H. S., Iyer K., Hyun S. H., Wleklinski M., Cooks R. G., Thompson D. H. Multistep Flow Synthesis of Diazepam Guided by Droplet-Accelerated Reaction Screening with Mechanistic Insights from Rapid Mass Spectrometry Analysis. *Org Process Res Dev*, 21 (10), 1566–1570, (2017).
13. Diab S., Mytis N., Boudouvis A. G., Gerogiorgis D. I. Process modelling, design and techno-economic Liquid-Liquid Extraction (LLE) optimisation for comparative evaluation of batch vs. continuous pharmaceutical manufacturing of atropine. *Comput Chem Eng*, 124, 28–42, (2019).

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Neural Networks for Prediction of Complex Chemistry in Water Treatment Process Optimization

Alexander V. Dudchenko<sup>a,\*</sup>, Oluwamayowa O. Amusat<sup>b</sup>

<sup>a</sup> SLAC National Accelerator Laboratory, 2575 Sand Hill Rd, Menlo Park, CA 94025, USA

<sup>b</sup> Lawrence Berkeley National Laboratory (LBNL), 1 Cyclotron Rd, Berkeley, CA 94720, USA

\* Corresponding Author: avd@slac.stanford.edu

## ABSTRACT

Water chemistry plays a critical role in the design and operation of water treatment processes. Detailed chemistry modeling tools use a combination of advanced thermodynamic models and extensive databases to predict phase equilibria and reaction phenomena. The complexity and formulation of these models preclude their direct integration in equation-oriented modeling platforms, making it difficult to use their capabilities for rigorous water treatment process optimization. Neural networks (NN) can provide a pathway for integrating the predictive capability of chemistry software into equation-oriented models and enable optimization of complex water treatment processes across a broad range of conditions and process designs. Herein, we assess how NN architecture and training data impact their accuracy and use in equation-oriented water treatment models. We generate training data using PhreeqC software and determine how data generation and sample size impact the accuracy of trained NNs. The effect of NN architecture on optimization is evaluated by optimizing hypothetical black-box desalination processes using a range of feed compositions from USGS brackish water data set, tracking the number of successful optimizations, and testing the impact of initial guess on the final solution. Our results clearly demonstrate that data generation and architecture impact NN accuracy and viability for use in equation-oriented optimization problems.

**Keywords:** Machine Learning, Water, Technoeconomic Analysis, Pyomo, Wastewater

## INTRODUCTION

Water chemistry plays a critical role in the design and operation of water treatment processes. Detailed chemistry modeling tools use a combination of advanced thermodynamic models and extensive databases to accurately predict phase equilibrium and reaction processes, such as those done by open-source PhreeqC software [1]. The complexity, formulation, and extensive database of these models preclude their direct integration in equation-oriented modeling platforms, making it difficult to use their capabilities for rigorous water treatment process optimization.

Neural networks (NN) can provide a pathway for integrating the predictive capability of chemistry software into equation-oriented models and enable optimization of complex water treatment processes across a broad range of conditions and process designs. A key challenge in developing a broadly applicable surrogate model for aqueous chemistry is the high non-linearity of the phenomena and high problem dimensionality. NNs have the potential to learn chemical phenomena and provide

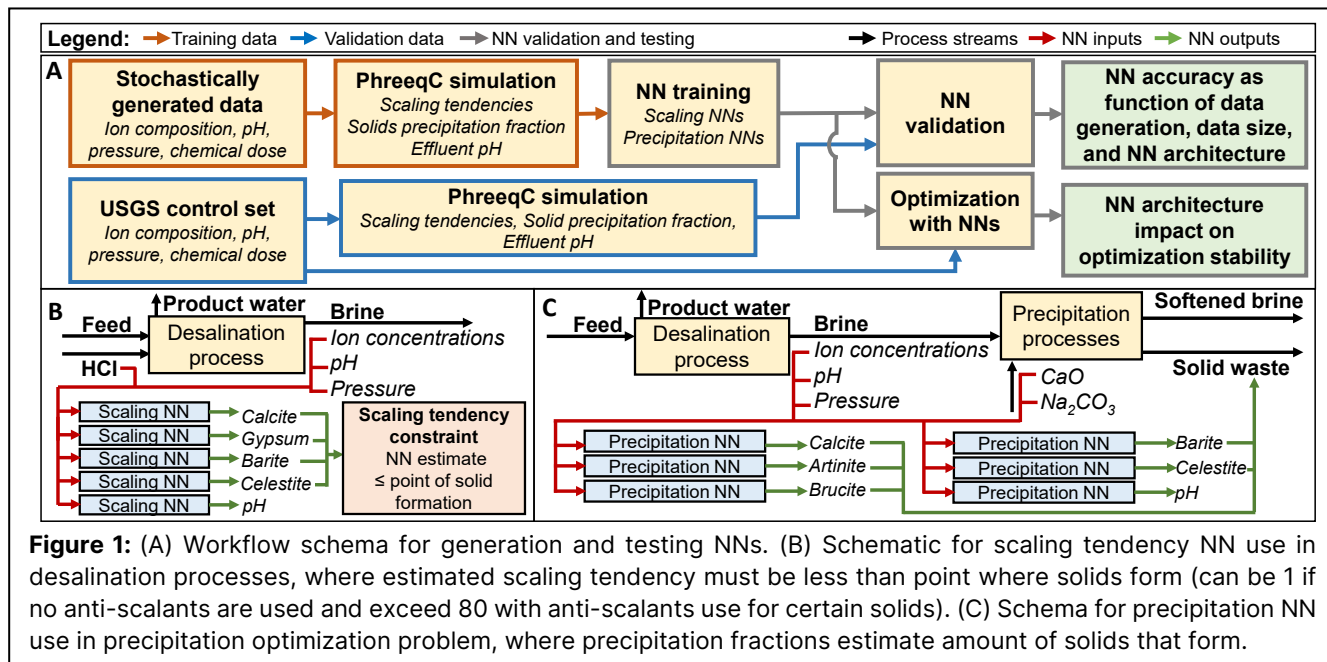
accurate estimates over a broad range of ion and reactant compositions [2].

NNs developed for water treatment optimization must accurately predict three critical chemistry phenomena: scaling tendencies, precipitation fractions, and changes in pH.

Modeling scaling tendency in desalination processes is critical for predicting solid formation, which occurs as ion concentration increases and results in process failure [3]. Typically, scaling is mitigated by reducing water recovery of desalination process, adding acids, which reduce scaling potential, and adding antiscalants that increase maximum allowable scaling tendency [4].

Modeling chemically driven precipitation is important for estimating the performance of pretreatment and softening processes, such as lime (CaO) and soda-ash (Na<sub>2</sub>CO<sub>3</sub>) softening [5]. These processes are commonly used to drive the formation of solids that remove divalent ions and reduce water hardness. Removing divalent ions minimizes the potential for scale formation in downstream processes.

Accounting for pH change due to chemical addition,



**Figure 1:** (A) Workflow schema for generation and testing NNs. (B) Schematic for scaling tendency NN use in desalination processes, where estimated scaling tendency must be less than point where solids form (can be 1 if no anti-scalants are used and exceed 80 with anti-scalants use for certain solids). (C) Schema for precipitation NN use in precipitation optimization problem, where precipitation fractions estimate amount of solids that form.

such as HCl and lime, is critical for modeling full treatment trains. The addition of HCl and lime will change the effluent stream pH and change requirements for chemical addition in downstream processes.

Herein, we investigate how the training data and NN architecture affect NN performance for the three chemistry phenomena (Figure 1). We train NNs using stochastically generated data, and validate NN accuracy and stability in equation-oriented optimization models using USGS data set [6] (Figure 1A). Stochastic data generation is used to simulate a broad range of chemistries and chemical additions, which are used as inputs for PhreeqC to predict chemical phenomena of interest. We investigate the impact of training data by considering the effect of data skewing to sample lower ion concentration distributions preferentially, which better captures real water compositions represented by the USGS brackish water data set. Finally, the developed NNs are integrated into equation-oriented models, and their stability is evaluated by tracking the number of successful solutions and the sensitivity to the initial guess.

## METHODS

### Data generation

We use PhreeqC with the included Pitzer database to estimate scaling tendencies, precipitation fractions, and effluent pH based on feed composition and chemical addition (Figure 1A). Feed is charge neutralized by adjusting Cl ion concentration, and charge neutral composition is used as input for NN.

For NN training, we stochastically generated ion compositions and chemical addition amounts (Table 1). We store PhreeqC result for feed composition without and with chemical addition (Table 2). The precipitation

fraction is the fraction of primary ion in the tracked phase that precipitates from the solution (e.g. precipitate fraction for  $\text{CaCO}_3$  is the ratio of precipitated Ca to total Ca in solution (feed and added reactant) before reaction).

The feed compositions, pressure, pH, and amount of chemical added is sampled using the standard Latin hypercube (LHS) method [7]. The LHS samples are scaled 0-1 and are rescaled to ion concentration, pressure, and chemical addition using the exponential skew function (eq. 1) and for pH using the log skew function (eq. 2). In these equations,  $x$  is the LHS sample,  $S$  is the skew magnitude,  $R_{high}$  is the high range of absolute value, and  $R_{low}$  is the low range for absolute value.

$$f_{exp-skew}(x) = (S^x - 1) * \frac{R_{high} - R_{low}}{S - 1} + R_{low} \quad (1)$$

$$f_{log-skew}(x) = \log_{10}(S * x + 1) * \frac{R_{high} - R_{low}}{\log_{10} S + 1} + R_{low} \quad (2)$$

The exponential skew function rescales the input value to be exponentially lower, such that an LHS sample of 0.1 (range 0-1) would be rescaled to 0.028, 0.006, 0.001, 0.0002 for skew magnitudes of 10, 100, 1,000, and 10,000, respectively. This skewing increases the sampling of low ion concentration distribution typically observed in real waters. The log skew function has an inverse relationship; for an LHS sample of 0.1 (range 0-1) the rescaled samples would be 0.15, 0.34, 0.5, and 0.6 for skew magnitudes of 10, 100, 1,000, and 10,000, respectively. The log skew function is used to increase the number of samples with basic pH (higher pH), where pH sensitive scalants have high scaling tendencies and are filtered out during data generation (described below).

We generate 2 million total samples using PhreeqC for NN training using a range of skew magnitudes only applied to ion composition, pressure, and chemical



addition, while pH skew magnitude is maintained at 10. The data is generated iteratively by taking 10,000 LHS samples at every step, applying the appropriate skew functions to the LHS samples, and keeping samples with total solids concentrations below 360 g per kg of water (g/kgw) and scaling tendency below 100. The procedure is repeated until 2 million samples are generated.

We further use the USGS brackish water data set to test NNs for real water conditions. The data set was filtered to isolate samples that exist in water-stressed regions using the method described by Ahban et al.[8]. This data set was filtered to remove water composition with high silica scaling potential by removing samples with Si concentrations above 50 ppm and  $H_4SiO_2$  concentration above >2 mM at 90% water recovery and pH of 6.5 [9]. This data is referred to as the USGS control set and contains 782 feed compositions.

**Table 1:** Feed composition and inputs in the stochastic data set. Temperature is fixed at 20°C.

<b>Feed composition in stochastic data</b>			
<i>Input</i>	<i>Low range</i>	<i>High range</i>	<i>Unit</i>
Na	0	136	
Cl	0	180	
Ca	0	10	
Mg	0	10	
HCO <sub>3</sub>	0	10	g/kgw
SO <sub>4</sub>	0	100	
K	0	40	
Sr	0	10	
Ba	0	0.1	
pH	4	12	pH

<b>Inputs for scaling tendency prediction</b>			
<i>Input</i>	<i>Low range</i>	<i>High range</i>	<i>Unit</i>
Pressure	1	401	atm
HCl	0	2000	PPM

<b>Inputs for precipitation fraction prediction</b>			
<i>Input</i>	<i>Low range</i>	<i>High range</i>	<i>Unit</i>
CaO	0	2000	PPM
Na <sub>2</sub> CO <sub>3</sub>	0	2000	

The USGS control data set is used to generate test data by imitating a black box desalination process within PhreeqC. The test data set for scaling tendencies includes process operating with water recoveries ranging from 0 to 90% in 10% steps and with HCl addition of 10, 50, 100, 1,000, and 1,500 ppm (N=38,357). The water recoveries for the precipitation fraction test data set were 0, 20, 60, and 90%, and CaO and Na<sub>2</sub>CO<sub>3</sub> addition was 10, 100, and 1,000 PPM (N=27,460). All data sets excluded samples that resulted in scaling tendencies above

100.

## Neural network training and validation

NNs were built and trained using Pytorch 2.0.0 on NVIDIA GPUs using standard CUDA implementation [10]; only stochastically generated data was used for training, with 10,000 samples set aside for testing. Throughout the paper, we only present results that use the USGS control set to quantify errors in NNs.

We build dense networks using 3 and 5 deep layers, with 30, 60, and 90 neurons with either sigmoid or tanh activation function applied to all layers except the output layer. All NNs use the ion composition and feed pH as inputs. Additionally, the scaling tendency NNs include HCl addition and pressure as inputs, while the precipitation NNs include CaO and Na<sub>2</sub>CO<sub>3</sub> as inputs. Each NN predicts only a single output, resulting in five scaling NNs and six precipitation NNs per architecture type. Additionally, each network is trained using two different weight decays of 0.001 and 10<sup>-6</sup>. Thus, a total of 24 NN architectures are evaluated (2 deep layers x 3 neuron types x 2 activation function x 2 weight decays).

**Table 2:** Output scaling tendencies ranges and precipitation fractions.

<b>Scaling tendency data set outputs</b>		
<i>Output</i>	<i>Low range</i>	<i>High range</i>
Calcite	0	98.19
Gypsum	0	14.61
Barite	0	99.03
Celestite	0	97.7
pH	0.56	11.96

<b>Precipitation fraction data set outputs</b>			
<i>Output</i>	<i>Low range</i>	<i>High range</i>	<i>Primary ion</i>
Calcite	0	1	Ca
Artinite	0	0.96	Mg
Brucite	0	1	Mg
Barite	0	1	Ba
Celestite	0	0.99	Sr
pH	4	12.8	N/A

We use Pytorch implementation of AdamW optimizer using cosine annealing with warm restart to train NNs [11]. We train with five cycles, that switch based on the number of gradient updates, performing ~2M gradient updates in total. The learning rates for sigmoid and tanh activation functions were 0.01 and 0.001, respectively, with the final learning rate of 10<sup>-6</sup>. We use mini-batches with a size of 4096 samples. At each epoch, the data is shuffled and sampled without replacement. Finally, we linearly scale all input and output data for training between 0 and 1.

To test the effect of skew magnitude and data size on NN accuracy, we use NN with 5 layers and 60 neurons (5x60) with a sigmoid activation function and trained with a weight decay of  $10^{-6}$  on Calcite scaling tendency, pH after HCl addition, and Calcite precipitation fraction.

We validate NN accuracy by comparing their predictions against the USGS control data sets and presenting errors in all figures as statistical distributions, showing 5<sup>th</sup>, 25<sup>th</sup>, median, 75<sup>th</sup>, and 95<sup>th</sup> percentile errors. The error for scaling tendencies and pH is the percent difference between NN prediction and ground truth (*GT*) generated using PhreeQC, as shown in equation 3.

$$\text{error} = \frac{NN_{\text{prediction}} - GT}{GT} * 100\% \quad (3)$$

We exclude any errors for scaling tendencies below 0.5, as in water treatment no scaling occurs below scaling tendencies of 1. For these samples, the absolute errors remain below 0.2 but can significantly shift error distribution (e.g. for a scaling tendency of 0.1, the prediction could be 0.3, resulting in 200% error but having no implication for optimization of the water treatment process). For precipitation fractions, we present the absolute difference between NN and ground truth in percent, as shown in equation 4.

$$\text{error} = NN_{\text{prediction}} - \text{ground}_{\text{truth}} * 100\% \quad (4)$$

## Neural network integration into waterTAP

The two desalination processes we consider (Figure 1B and 1C) are modeled in WaterTAP, a framework for technoeconomic assessment of water treatment systems. We integrate developed NNs into WaterTAP using the Optimization and Machine Learning Toolkit (OMLT) [12, 13]. We use the reduced smooth formulation in OMLT, which loads the NN as a single large expression into the model. Additionally, we add constraints that convert absolute values to scaled NN inputs and outputs.

We formulate two optimization problems that emulate black box desalination processes with scaling and precipitation NNs (Figure 1B and 1C). The desalination process is emulated using a single feed block that specifies the feed mass flow of ions and water. The ion mass flow is fixed to ion concentrations as specified by a sample from USGS data set. The mass flow of water is unfixed during optimization, imitating a desalination process that removes pure water from the feed block and increases ion concentration in the remaining brine. The difference between the initial and the optimized flow mass of water is equal to the amount of product water, while the optimized water flow mass is equal to mass flow of waste brine.

The scaling tendency problem, as shown in Figure 1B, is where scaling NNs are added to the desalination problem to predict scaling tendencies due to increased ion concentration in feed block, with HCl added as a scaling control mechanism. The scaling tendencies are

constrained to remain below 60 for Calcite, 2.3 for Gypsum, 60 for Barite, and 8 for Celestite, imitating a desalination process operating with anti-scalants [3].

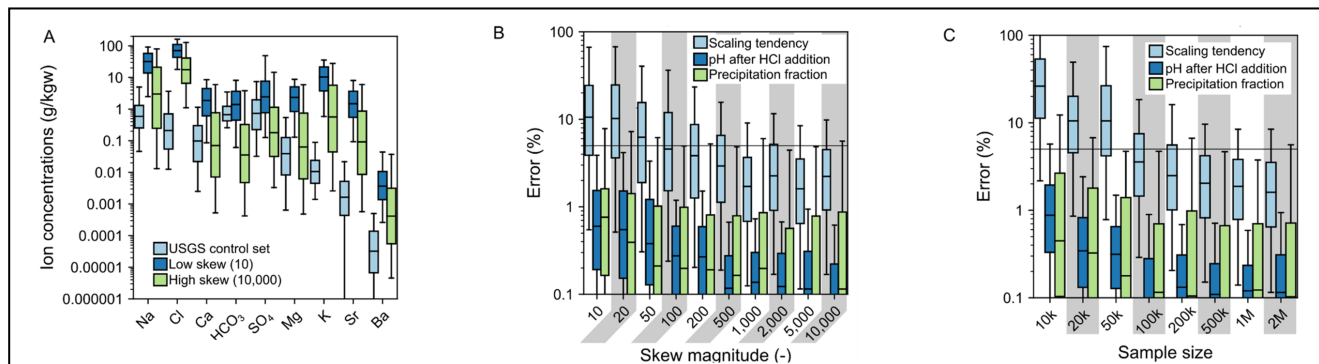
The precipitation problem, as shown in Figure 1C, is where precipitation NNs are added to the desalination problem to predict the removal of solids that form due to increased concentration of ions and chemical addition [5]. Here an additional constraint is added to calculate the final brine hardness after solid formation, which is constrained to be below or equal to 50 ppm of  $\text{CaCO}_3$  as shown in equation 5, where  $Ca_{f\text{-total}}$  is mole flow rate of Ca in feed and Ca added from CaO addition,  $Mg_f$  is mole flow rate of Mg in feed, and  $rm_{\text{Calcite}}$ ,  $rm_{\text{Artenite}}$ , and  $rm_{\text{Brucite}}$  are removal fraction for Calcite, Artinite, and Brucite respectively.

$$\begin{aligned} \text{Total - hardness} = \\ (|Ca_{f\text{-total}} - Ca_{f\text{-total}} * rm_{\text{Calcite}}| * 100.1 + |Mg_f - Mg_f * \\ (rm_{\text{Artenite}} + rm_{\text{Brucite}})| * 100.1) * 1000 \quad (5) \end{aligned}$$

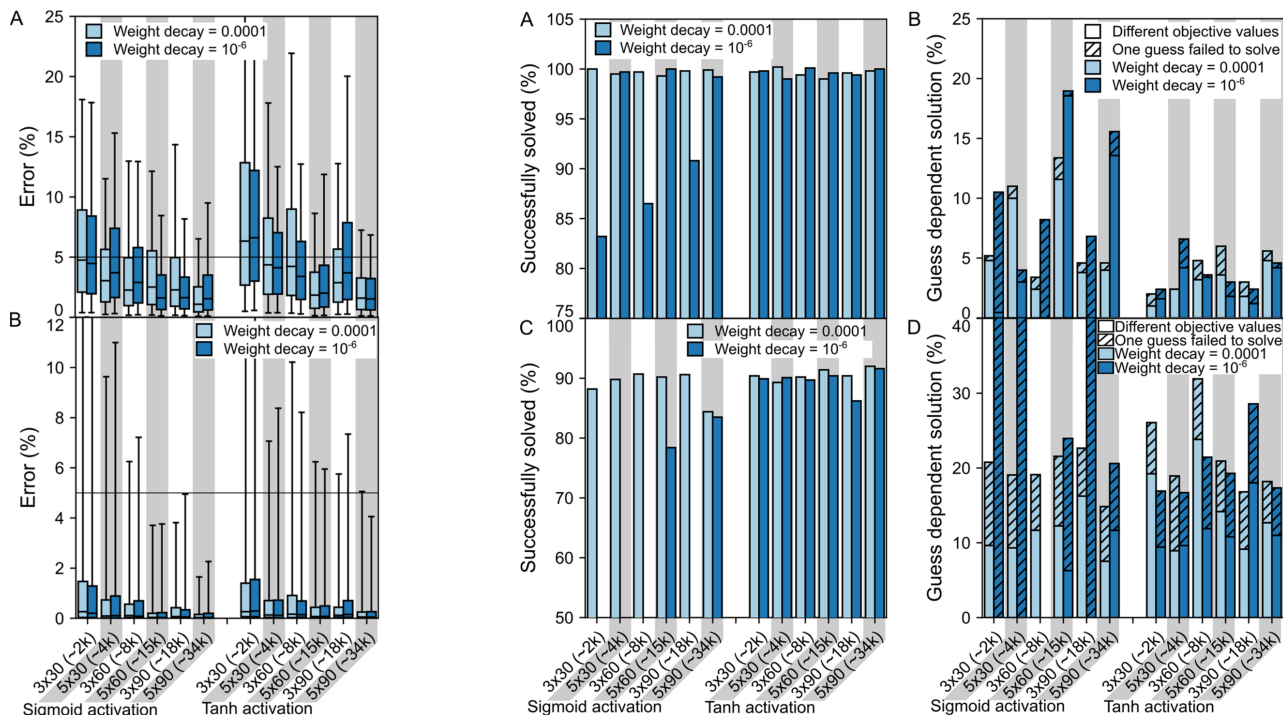
The optimization problems are formulated to maximize the value of produced water from black box desalination processes. In the optimization, the product water can be sold with a value of 0.5 units/kg, while the remaining waste brine is penalized with a cost of 10 units/kg, and the addition of chemicals (HCl, CaO, and  $\text{Na}_2\text{CO}_3$ ) is penalized with the value of 0.19 units/kg. For the scaling tendency problem, there are 2 degrees of freedom: the amount of water to remove and the amount of HCl to add. For the removal fraction problem, there are 3 degrees of freedom: amount of water to remove, amount of CaO to add, and amount of  $\text{Na}_2\text{CO}_3$  to add.

In these optimization problems, removing water increases ion concentration, which increases scaling tendencies and hardness. The scaling tendencies can be decreased for some scalants with HCl addition. Hardness can be reduced by adding CaO and  $\text{Na}_2\text{CO}_3$ , which increase precipitation fractions of Ca and Mg containing solids. Thus, the optimization balances the cost of producing water, disposing of brine, and satisfying scaling and hardness constraints through chemical addition.

We optimize the two problems by randomly drawing 500 feed compositions from the USGS control set, and providing two initial guesses. We provide one guess that is poorly posed where 90% water is removed and 1 PPM of chemicals is added, resulting in a feed composition that is likely violating the scaling tendencies and hardness constraints. The second guess is well-posed, where feed is diluted by 5 times and 1,000 PPM of chemicals are added, resulting in a feed composition that is unlikely to violate scaling tendencies and hardness constraints. The total solvability is evaluated by tracking the percent of total solved samples and guesses ( $N=1000$ ). The NN propensity to solve to a local minimum is quantified by tracking the number of solutions with different objective values caused by different initial guesses for the same feed composition and where one guess failed to solve. The



**Figure 2:** Effect of data generation on NN accuracy. (A) Effect of skew magnitude on ion sampling, (B) Effect of skew magnitude on NN accuracy, (C) Effect of sample size on NN accuracy. Error is shown for Calcite scaling tendency and precipitation fraction.



**Figure 3:** Effect of network architecture on NN error in prediction of (A) scaling tendencies and (B) removal fractions of Calcite. The network size is labeled as depth x width with values in parenthesis showing a rounded number of trainable parameters (e.g. 3x30 (~2k) indicates an NN with 3 deep layers and 30 neurons, with ~2,000 trainable parameters).

**Figure 4:** Effect of network architecture on optimization using IPOPT. (A) Percent of successful solves and (B) guess dependence for scaling problems. (C) Percent of successful solves and (D) guess dependence for precipitation problem. The network size is labeled as depth x width with values in parenthesis showing a rounded number of trainable parameters (e.g. 3x30 (~2k) indicates an NN with 3 deep layers and 30 neurons, with ~2,000 trainable parameters).

problems were solved using IPOPT with MA27 linear solver [14].

## RESULTS AND DISCUSSION

### Effect of data on Neural Network accuracy

The ultimate accuracy of NNs in predicting phenomena of interest depends on the statistical distribution of underlying training data and data size. In the case of aqueous chemistry used for water treatment

applications, a common objective is to predict the propensity of ions in solution to form solids as a function of ion composition, pH, temperature, and pressure. Due to the sheer number of dimensions (12 dimensions are considered herein (Table 1)), a stochastic generation methodology must be used to generate data sets that enable the training of surrogate models applicable to different feed water composition and water treatment processes.

The nonlinear nature of aqueous chemistry requires a nonuniform sampling of ion compositions to ensure

relevant data is generated from chemistry software. For example, Calcium is an ion of great interest in water treatment and will form Gypsum ( $\text{CaSO}_4$ ) in the presence of  $\text{SO}_4$ , which has solubility below 2.5 g/kgw and limits Ca concentrations. However, in the absence of  $\text{SO}_4$ , calcium concentrations can exceed 10 g/kgw, although such waters are uncommon, as  $\text{SO}_4$  is ubiquitous in real waters. Thus, uniformly sampling across ion concentrations will generally form ion compositions, scaling tendencies, and precipitation fractions that are not representative of real waters and conditions under which water treatment processes operate.

Skewing the sampling of ion concentrations to prioritize lower concentrations generates ion composition distributions that are more representative of real water compositions (Figure 2A). Increasing skewing magnitude from 10 to 10,000 results in a sampling of much lower ion concentrations on the median that approach USGS control set values but still sample high ion concentrations that occur in desalination processes.

Testing NNs against real ion composition and chemical additions that would be encountered by desalination and precipitation water treatment processes (USGS control set) demonstrates that accuracy increases with an increase in the skew magnitude (Figure 2B). At skew magnitude below 100, the median error for Calcite scaling tendency approached 10%. Increasing the skew magnitude above 200 reduced 75<sup>th</sup> percentile errors to below 5% for Calcite scaling tendency and reduced errors for precipitation fractions and pH.

Testing NNs using stochastically generated test data demonstrated a slight increase in prediction errors of less 0.5 % on the median with an increase in skew magnitude from 10 to 10,000. These test errors were significantly lower than those for USGS control set, with 75<sup>th</sup> percentile errors being less than <1% across all skew magnitudes. This clearly shows that NNs are learning chemistry in underlying training data, and skewing is shifting training to learning chemistry relevant to real water composition, as found in USGS control data set.

Increasing the data size used for training improves NN accuracy (Figure 2C). Training with less than 100,000 samples produces poor NN accuracy for Calcite scaling tendency, with median errors exceeding 10%. Increasing the sample size to above 200,000 samples reduced errors to below 5% in the 75<sup>th</sup> percentile for Calcite scaling tendency and the 95<sup>th</sup> percentile for pH and Calcite precipitation fraction. Increasing the sample size from 200,000 to 2,000,000 provided minimal improvements in overall network accuracy, suggesting that data sizes of around 500,000 could be sufficient for predicting scaling tendency. However, in the case of pH and precipitation fractions, data sizes of 50,000-100,000 samples could be adequate. The data size requirement is expected to increase with the number of inputs.

## Effect of architecture on accuracy

We explore the effect of NN architecture by training 24 different NN types using the 2M sample data set with a skew magnitude of 5,000. Each NN predicts a single output, resulting in five scaling NNs per architecture that predict scaling tendencies for the four scalants and pH, and six precipitation NNs per architecture that estimate precipitation fractions for the five solids and pH, as shown in Table 1 and Table 2. Calcite scaling tendencies and calcite precipitation fractions always have higher errors and are thus used as benchmarks for the performance of scaling tendencies and precipitation fraction prediction.

Prediction error decreased with an increase in the number of NN trainable parameters (Figure 3). The smallest NNs with 3 deep layers and 30 neurons (3x30) have 2,281 trainable parameters and exhibited median error in scaling tendencies on the order of ~5% and ~7% for sigmoid and tanh activation functions, respectively. In contrast, the largest 5x90 networks with 34,021 trainable parameters had errors <3% in 75<sup>th</sup> percentile (Figure 3A). Similar improvements were observed for precipitation networks (3B).

In general, we observed that deeper networks had slightly lower errors than shallow networks with a similar number of trainable parameters. This is demonstrated by the 3x90 network with 17,641 trainable parameters typically exhibiting a higher or similar level of error compared to their 5x60 network counterparts with 15,481 trainable parameters. In theory, the three-layer deep and five-layer deep network should be able to encode the same degree of nonlinear behavior, but it appears that higher depth improves encoding performance.

Activation function type and weight decay had a small effect on prediction error relative to network size impact. On average, NNs with sigmoid activation function had lower errors than NNs with tanh activation function, but the difference was less than 0.5% for median errors. The weight decay did impact the accuracy of NNs, but the impact depends on their architecture, implying that each architecture has an optimal weight decay value.

## Effect of Neural Network architecture on optimization

Surrogate models developed for equation-oriented programming models must faithfully capture underlying phenomena and provide smooth, continuous functions that ensure a successful optimization and avoid local solutions when solved with a local solver such as IPOPT. We test the NNs stability in optimization by integrating them in two black box desalination problems, the scaling tendency problem (Figure 1B), and the precipitation problem (Figure 1C).

The optimization results demonstrate that optimization problems that use NNs with tanh activation function



solve more often, are less likely to fall into a local minimum, and solve more frequently regardless of initial guess (Figure 4). The use of the tanh activation function resulted in >98% of problems successfully solving for the scaling tendency problem and >88% for the removal fraction problem (Figure 4A and 4C). In contrast, the use of the sigmoid activation function in small and shallow networks that use low weight decay ( $10^{-6}$ ) resulted in a large number of failed solves.

The optimization problems that used NNs with tanh activation function were not strongly impacted by initial guess relative to problems that used NNs with sigmoid activation function (Figure 4B and 4D). For the scaling tendency problem, using the tanh activation function resulted in less than 5% of solutions having different objective values caused by different initial guesses across all NNs. This contrasts the results for optimization problems that used NNs with the sigmoid activation function, even when it had a comparable number of solves to tanh, the sigmoid activation function produced a higher number of local solves (exceeding 10% in 4 of the NNs). This trend was not observed in the removal fraction problem, where a similar number of local minimums (~10% of samples) were observed for both activation functions. Similarly, for cases where the number of successful solves is comparable between tanh and sigmoid, the initial guess resulted in a similar number of failed solves.

We hypothesize that tanh function provides a more mathematically stable formulation than the sigmoid activation function, resulting in a higher likelihood that a problem is solved successfully, even as the use of the sigmoid function resulted in more accurate NNs. The tanh function provides larger gradients than the sigmoid activation function, improving the potential for IPOPT in avoiding local minimums. In addition, the sigmoid formulation utilizes an exponential in its denominator, which is prone to overflow errors during computation. The probability of overflow decreases as the network size increases and the weight decay parameter gets higher. Larger network size and weight decay regularize weights and reduce the potential for extreme values being passed into the activation function. This results in a lower likelihood of an exponential overflow occurring when changing NN input values, explaining why the sigmoid function performs better with the increase in trainable parameters and weight decay. With the tanh formulation, however, overflow can not occur, and optimization algorithms can freely vary the inputs.

NN architecture did not impact the number of iterations required to solve an optimization problem, but it did increase computational time. On average, the scaling problem required 20 iterations to solve, with 95<sup>th</sup> percentile of problems requiring 40 iterations, regardless of NNs used. The removal fraction problems required on average 40 iterations to solve, with 95<sup>th</sup> percentile requiring ~100 iterations, regardless of NNs used. The largest impact of

increasing network size was on solving time, with the use of the 5x90 NNs requiring upto ten times more time per iteration than for 3x30 NNs. The activation function type and weight decay did not impact the solve time.

The integration of NNs into optimization problems did not impact their accuracy. The results demonstrated similar error distributions to those observed in USGS control set, confirming the expectation that using NNs in an equation-oriented model does not affect their accuracy.

## CONCLUSION

The results of this work have demonstrated that NNs can be effective surrogate models for accurately predicting scaling tendencies, pH change, and removal fraction across a broad range of real water compositions and use in the equation-oriented programming models of water treatment processes. Our results have extracted a set of generalizable guidelines for data generation and design of NNs for chemistry prediction in water treatment equation-oriented process optimization:

- Sampling of ion compositions and chemical addition should be skewed to lower values to provide good accuracy in real water compositions.
- The number of required stochastically generated data samples depends on the underlying chemistry being modeled.
- Deeper networks provide slightly better performance over shallower networks with the same number of trainable parameters.
- Tanh activation function provides better stability than sigmoid activation functions in equation-oriented models, and should be preferred over sigmoid even at the cost of accuracy. This is a fundamental mathematical limitation and applies to all NNs, regardless of the data they are predicting.
- Weight decay should be optimized to a specific network architecture to extract the highest accuracy.
- Large NNs do not increase the number of iterations but do increase the solving time required per iteration.

The methods presented herein, for the first time, enable optimization studies of water treatment processes to consider a broad range of real water compositions, capturing their performance while faithfully representing real and complex aqueous chemistry.

## ACKNOWLEDGEMENTS

This material is based upon work supported by the National Alliance for Water Innovation (NAWI), funded by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy (EERE), Advanced Manufacturing Office, under Funding Opportunity Announcement Number DE-FOA-0001905. The version of IPOPT implemented in this work was compiled using HSL, a collection of Fortran codes for large-scale scientific



computation. See <http://www.hsl.rl.ac.uk>.

and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



## REFERENCES

1. Description of Input and Examples for PHREEQC Version 3—A Computer Program for Speciation, Batch-Reaction, One-Dimensional Transport, and Inverse Geochemical Calculations. (2013)
2. Pirdashti, M., Curteanu, S., Kamangar, M.H., Hassim, M.H., Khatami, M.A.: Artificial neural networks: applications in chemical engineering. *Reviews in Chemical Engineering*. 29, 205–239 (2013).
3. Gabelich, C.J., Rahardianto, A., Northrup, C.R., Yun, T.I., Cohen, Y.: Process evaluation of intermediate chemical demineralization for water recovery enhancement in production-scale brackish water desalting. *Desalination*. 272, 36–45 (2011).
4. Hydranautics: Chemical Pretreatment For RO and NF. Nitto Denko, (2008)
5. Hoover, C.P.: Review of Lime-Soda Water Softening. *Journal - American Water Works Association*. 29, 1687–1696 (1937).
6. Qi, S.L., Harris, A.C.: Geochemical Database for the National Brackish Groundwater Assessment of the United States, (2017)
7. Mease, D., Bingham, D.: Latin Hyperrectangle Sampling for Computer Experiments. *Technometrics*. 48, 467–477 (2006).
8. Ahdab, Y.D., Thiel, G.P., Böhlke, J.K., Stanton, J., Lienhard, J.H.: Minimum energy requirements for desalination of brackish groundwater in the United States with comparison to international datasets. *Water Research*. 141, 387–404 (2018).
9. Sheikholeslami, R., Bright, J.: Silica and metals removal by pretreatment to prevent fouling of reverse osmosis membranes. *Desalination*. 143, 255–267 (2002).
10. Stevens, E., Antiga, L., Viehmann, T.: Deep learning with PyTorch. Manning Publications Co, Shelter Island, NY (2020)
11. Loshchilov, I., Hutter, F.: Decoupled Weight Decay Regularization, <http://arxiv.org/abs/1711.05101>, (2019)
12. Beattie, Keith, Gunter, Daniel, Knueven, Ben, Lee, Andrew, Ladshaw, Austin, Drouven, Markus, Bartholomew, Tim, Bi, Xiangyu, Bianchi, Ludovico, Arnold, Travis, Atia, Adam, Wang, Chenyu, Miara, Ariel, Kurban, Sitterley, Allu, Srikanth, Dudchenko, Alexander, Amusat, Oluwamayowa, Kinshuk, Panda, Young, Ethan: WaterTAP v1.0.0, (2021)
13. Ceccon, F., Jalving, J., Haddad, J., Thebelt, A., Tsay, C., Laird, C.D., Misener, R.: OMLT: Optimization & Machine Learning Toolkit.
14. Wächter, A., Biegler, L.T.: On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Program.* 106, 25–57 (2006).

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator

# Improving Mechanistic Model Accuracy with Machine Learning Informed Physics

William Farlessyost<sup>a,\*</sup>, Shweta Singh<sup>a,b</sup>

<sup>a</sup> Purdue University, Agricultural & Biological Eng., West Lafayette, Indiana, USA

<sup>b</sup> Purdue University, Ecological & Environmental Engineering, West Lafayette, Indiana, USA

\* Corresponding Author: [wfarless@purdue.edu](mailto:wfarless@purdue.edu).

---

## ABSTRACT

Machine learning presents opportunities to improve the scale-specific accuracy of mechanistic models in a data-driven manner. Here we demonstrate the use of a machine learning technique called Sparse Identification of Nonlinear Dynamics (SINDy) to improve a simple mechanistic model of algal growth. Time-series measurements of the microalga *Chlorella Vulgaris* were generated under controlled photobioreactor conditions at the University of Technology Sydney. A simple mechanistic growth model based on intensity of light and temperature was integrated over time and compared to the time-series data. While the mechanistic model broadly captured the overall growth trend, discrepancies remained between the model and data due to the model's simplicity and non-ideal behavior of real-world measurement. SINDy was applied to model the residual error by identifying an error derivative correction term. Addition of this SINDy-informed error dynamics term shows improvement to model accuracy while maintaining interpretability of the underlying mechanistic framework. This work demonstrates the potential for machine learning techniques like SINDy to aid simple mechanistic models in scale-specific predictive accuracy.

---

**Keywords:** Dynamic Modelling, Machine Learning, System Identification, Surrogate Model, Batch Process

## INTRODUCTION

Within process design and system analysis, mechanistic models have long stood as pillars for understanding and predicting the behavior of complex systems. Rooted in fundamental physical and chemical principles, these models offer a structured approach to dissecting system dynamics, providing insights that are crucial for industrial applications [1]. However, as we delve deeper into the nuances of these systems, particularly at varying scales, the limitations of mechanistic models become apparent. A key challenge lies in their generalization; while these models are adept at capturing overarching trends and basic interactions, they often fall short in accurately representing the intricate dynamics of specific systems under varied conditions.

This is especially true in biological and ecological modeling, where the complexity of interactions and the sensitivity to environmental variables pose significant hurdles. The modeling of microalgal growth, such as that of *Chlorella vulgaris*, exemplifies this challenge. Here, the

transition from laboratory-scale observations to commercial-scale operations is fraught with complexities, as the models struggle to account for the myriad of factors influencing algal growth dynamics at different scales [3,4].

Yet the uptick of machine learning methods bolstering traditional system identification offer a beacon of hope in this landscape. Among the various machine learning techniques, Sparse Identification of Nonlinear Dynamics (SINDy) emerges as a particularly promising tool. SINDy is a novel algorithm that seeks to identify the governing equations of a dynamical system in a sparse, interpretable form [2-4]. This approach is particularly relevant in process modeling, where understanding the fundamental dynamics is as crucial as achieving predictive accuracy [5].

The objective of this study is to showcase the application of SINDy in refining a mechanistic model of algal growth. Focusing on the microalga *Chlorella vulgaris*, we utilize time-series data collected under controlled photobioreactor conditions at the University of Technology

Sydney (Unpublished Data).

The following paper delves into the methods, where we describe the initial mechanistic model based on light intensity and temperature, followed by the integration of the SINDy algorithm derived error term. The results section presents a comparative analysis between the original and the enhanced models, highlighting the improvements in accuracy and interpretability. In the discussion, we explore the implications of these findings for process design, particularly in the scaling of microalgal growth systems. Finally, the paper concludes with a summary of our key findings and their significance in the broader context of industrial applications.

## BACKGROUND

### Process Design

Historically, process design has been deeply rooted in the principles of chemical engineering, with a strong emphasis on understanding and manipulating material and energy balances [6-8]. Mechanistic models, characterized by their basis in fundamental scientific principles, have been extensively applied across various domains of process engineering [1, 9-13]. These models excel in their ability to predict system behavior under a range of conditions, providing a reliable foundation for process design and optimization. However, their application is not without limitations. One significant challenge is their generalization, which can lead to inaccuracies when applied to specific systems or at different scales [14]. Biological systems are inherently complex, and their dynamics can vary significantly with changing environmental conditions and scales, posing a substantial challenge to traditional mechanistic modeling approaches.

### Machine Learning

In recent years, the emergence of machine learning as a complementary tool in scientific research has opened new avenues in process design. Machine learning's data-driven nature allows it to uncover complex, non-linear relationships within high-dimensional data sets, offering a level of insight and prediction that is often beyond the reach of traditional models [15, 16]. This has been particularly advantageous in process modeling, where machine learning algorithms have been successfully integrated with mechanistic models to enhance their accuracy and adaptability [14].

Among the various machine learning techniques, SINDy has emerged as a promising tool for process design. Developed as an algorithm to identify governing equations of dynamical systems, SINDy distinguishes itself by its ability to distill complex data sets into sparse, interpretable models [2, 3]. This capability makes SINDy particularly suitable for process design applications where understanding the fundamental dynamics is as

crucial as achieving predictive accuracy.

## Microalgae

The modeling of microalgal growth, specifically the growth dynamics of *Chlorella vulgaris*, provides a compelling case study for the application of SINDy in process design [17, 18]. Microalgae, recognized for their potential in industrial applications, present significant challenges in scaling up from laboratory to commercial operations [18]. Traditional mechanistic models, while providing a basic understanding of algal growth, struggle to capture the full complexity of these biological systems at different scales [19, 20]

## METHODOLOGY

### Algae Mechanistic Model

We attempt to improve upon a simplified model that based on two key environmental factors: light intensity and temperature. This model combines the principles of Haldane kinetics and the Cardinal Temperature Model with Inflection (CTMI), offering a comprehensive understanding of algal growth dynamics. The model's first component, based on Haldane kinetics, addresses the effect of light intensity on algal growth. It recognizes that growth rate increases with light intensity up to an optimal level, after which it starts to decline due to photoinhibition. This relationship is mathematically represented by the equation:

$$\mu(I) = \mu_{opt} * I \left( I + \left( \frac{\mu_{opt}}{\alpha} \right) * \left( \left( \frac{I}{I_{opt}} \right) - 1 \right)^2 \right) \quad (1)$$

Here,  $\mu(I)$  represents the growth rate as a function of light intensity  $I$ ,  $\mu_{opt}$  is the maximum growth rate,  $I_{opt}$  is the optimal light intensity, and  $\alpha$  is a parameter that moderates the effect of light intensity on the growth rate.

The second component of the model, the CTMI, quantifies the influence of temperature on algae growth. This model defines the growth rate within a range of temperatures, specifying minimum ( $T_{min}$ ), optimal ( $T_{opt}$ ), and maximum ( $T_{max}$ ) temperatures for growth. The CTMI model is encapsulated in the following equation:

$$\mu(T) = k * \frac{(T - T_{min}) * (T_{opt} - T)}{((T_{opt} - T_{min}) * (T_{max} - T))} \quad (2)$$

In this formulation,  $\mu(T)$  denotes the growth rate as a function of temperature  $T$ , with  $T_{min}$ ,  $T_{opt}$ , and  $T_{max}$  being the minimum, optimal, and maximum temperatures for algal growth, respectively.  $k$  is a scaling constant.

The final model, encapsulating both the temperature and light components, is then the product of  $\mu(T)$  and  $\mu(I)$ .

### Data Source and Preparation

Our study utilizes data collected from the PBR 1250L photobioreactor system from Industrial Plankton as the primary cultivation equipment. This system provides a substantial working volume of 1250 liters, enabling a continuous harvesting capacity of up to 500 liters daily. Data collection was centered around minute measurements of pH level, irradiation, temperature, and relative density of the growth media within the photobioreactor. The pH control was managed by adjusting its setpoint and stabilized through automatic CO<sub>2</sub> injection. Irradiation was provided by LED light columns. Temperature control was maintained externally and the relative density of the algae was monitored through an optical density sensor.

The study involved a batch culture maintained over eleven days with alternating irradiation and pH setpoints, keeping the temperature constant and assuming a sufficient initial nutrient supply. The operational sequence was adjusted bi-daily, and the bioreactor was prepared with specific concentrations of culture media, followed by a regulated air bubbling process to maintain optimal growth conditions. This resulted in 22,301 data points for use in training and testing models.

### SINDy-Derived Error Model

We employ the SINDy algorithm [21] to identify a differential equation to model the rate of change of error between the actual and predicted *Chlorella vulgaris* densities. Here, the error given by Equation 4 as

$$e = \text{Density\_actual} - \text{Density\_model} \quad (4)$$

varies over time. The state data, represented by  $X_{sindy}$ , includes the error, capturing the essential aspects of the model's performance over time. Additionally, control data, denoted as  $U_{sindy}$ , incorporates key environmental factors, namely temperature  $T$ , light intensity  $I$ , and  $pH$  that significantly influence algal growth.

$$\dot{X}_{sindy} = e \quad (5)$$

$$\dot{X}_{sindy} = \dot{e} \quad (6)$$

$$U_{sindy} = [T \ I \ pH] \quad (7)$$

We develop a control-only error model using the SINDy framework with the introduction of a dummy state variable. This dummy variable, a zero-valued array, is designed to match the dimensions of  $X_{sindy}$ , the derivative of the error. Its primary function is not to represent a 'real' state variable in the traditional sense; instead, it serves as a tool for facilitating the SINDy algorithm process of system dynamics identification, which typically expects state variables to influence the system dynamics. The SINDy algorithm is thus attempting to fit sparsity matrix  $\Phi$  such that:

$$\dot{X}_{\{sindy\}} = U_{sindy}\Phi \quad (9)$$

To capture the complexity of biological growth

processes in our SINDy model, we included a variety of custom functions in the feature library. Each function represents a specific biological growth dynamic, such as logistic growth, exponential growth, and other well-known biological models shown in Table 1.

**Table 1:** Functional basis for SINDy library.

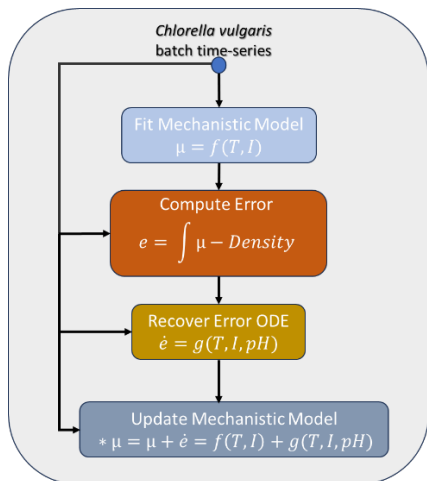
<b>Logistic growth</b>	$f(x) = \frac{1}{\{1 + e^{-x}\}} \quad (10)$
<b>Exponential growth</b>	$f(x) = e^{\{x\}} \quad (11)$
<b>Gompertz function</b>	$f(x, c_1, c_2) = e^{\{c_1 e^{-c_2 x}\}} \quad (12)$
<b>Allee effect</b>	$f(x, A) = \frac{x}{\{A - x + \epsilon\}} \quad (13)$
<b>Michaelis-Menten kinetics</b>	$f(x, K_m) = \frac{x}{\{K_m + x + \epsilon\}} \quad (14)$
<b>Holling's Type II function</b>	$f(x, a) = \frac{ax}{\{1 + ax + \epsilon\}} \quad (15)$

Here, the logistic growth function represents a growth process that accelerates rapidly at first and then slows down as it approaches a maximum limit, commonly used to model population growth. The exponential growth function describes a process where the rate of growth is proportional to the current amount, leading to a rapid increase over time. The Gompertz function models growth that initially accelerates and then decelerates, commonly used in biology to describe the growth of tumors or certain populations [22]. The parameters  $c_1$  and  $c_2$  reflect the maximum value the function will reach and the rate at which that point will be reached respectively. The Allee effect function represents a scenario where population growth is positively correlated with the population size, highlighting the difficulties in growth at very low densities, with  $A$  representing this population threshold value [23]. The parameter  $\epsilon$  is used here as well as the following functions to as a small value that prevents division by zero. The Michaelis-Menten kinetics function is typically used to describe the rate of enzymatic reactions, showing how reaction rate varies with substrate concentration [24]. The parameter  $K_m$  is the substrate concentration at which the growth rate is half of its maximum. Holling's Type II function, models the rate of predation in an ecosystem, increasing linearly at low prey density and saturating at higher densities [25]. The growth rate  $a$  here represents the efficiency of resource consumption by the growing population.

### Mechanistic & Error Model Coupling

After the ODE model for the error derivative is recovered, this sum of functions representing the

derivative is added directly to the original mechanistic CTMI-light model, as shown in Figure 1. This updated, or “corrected,” model can then be integrated overtime to return the specific density from the estimated growth rate.



**Figure 1:** Methodological flow resulting in an updated mechanistic model for *Chlorella vulgaris* rate of change of density.

## Model Evaluation

In the evaluation of the error ODE model, key metrics of RMSE, MAPE, and  $R^2$  are calculated for each fold of the time series cross-validation to assess model accuracy. RMSE is determined by taking the square root of the average of squared differences between predicted and observed values. MAPE is calculated as the average of the absolute differences between predicted and observed values, divided by the observed values, and expressed as a percentage.  $R^2$  is computed as the proportion of variance in the observed data that is predictable from the model inputs.

Visual comparisons of both the corrected and non-corrected models' predictions against the observed cumulative growth data are integrated and tested for each validation fold through a process of data preparation, plotting, visual analysis, and comparison. These plots allow for the assessment of how closely the predicted growth curves follow the observed data and the extent to which the corrected predictions improve alignment with the observed data compared to the non-corrected predictions.

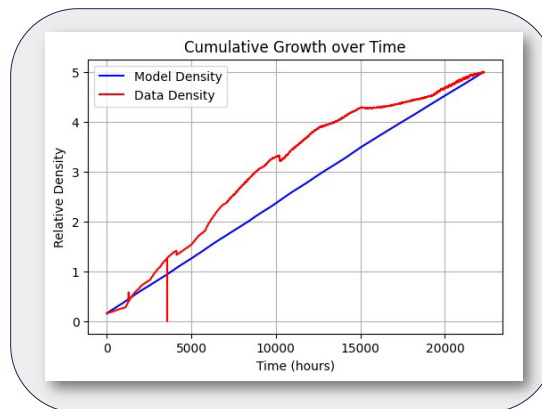
## RESULTS

### Algae Mechanistic Model

We fit the combined CTMI-light model for density using the full 22,301 measurement points available for temperature, light intensity, and density. Since the mechanistic model is built for non-standardized relationships

between components, we choose to pass in the raw data without standardization or normalization. We employed non-linear curve fitting techniques, utilizing the 'curve\_fit' function from the Python 'scipy.optimize' library, to estimate the parameters of the model with  $T_{opt} = 27.329\text{ C}$ .

To check the performance of this model, we integrate the fitted model over time and compare the result to the measured specific density. As seen in Figure 2, the combined model predicts a linear increase in algae density overtime.



**Figure 2:** Combined CTMI-light model for growth rate integrated over time versus measured model specific density over all 22,301 data points. Combined model RMSE is 0.009.

## Data Preparation

To make the SINDy optimization more robust against outliers in feature weights, the inputs, temperature, light intensity, and pH are each normalized by subtracting their time-series mean and dividing by standard deviation. This approach was chosen since the inputs to the batch were periodic, meaning their distribution was close to Gaussian.

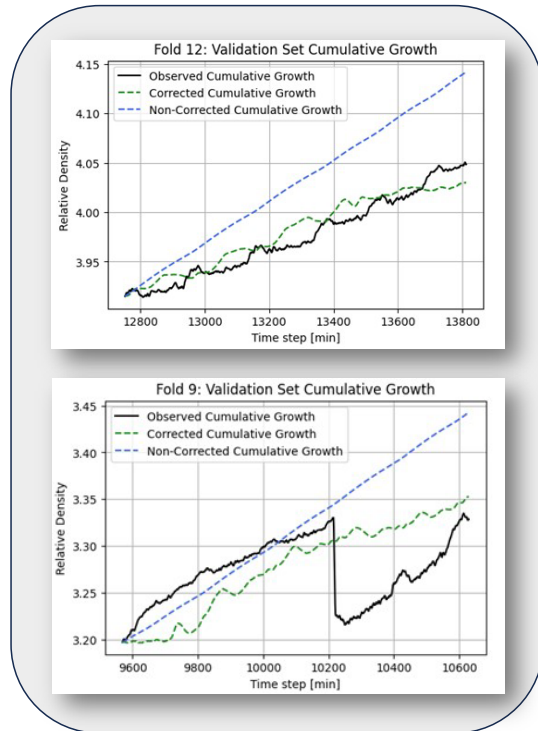
The full timeseries data, with standardized control inputs, is then broken into twenty folds of sequential data. Cross-validation is employed such that in each iteration, the model is trained on an expanding window of past data and validated on a subsequent, non-overlapping window of future data. This strategy allows the model to learn from a progressively larger dataset, capturing more of the underlying temporal patterns and dynamics as it moves through the folds.

## SINDy Error ODE Model Recovery

We leverage the sequentially thresholded least squares (STLSQ) algorithm as the optimization engine for Sparse Identification of Nonlinear Dynamics (SINDy). In this implementation, the STLSQ sparsity threshold is set at 0.0001, allowing the model to focus on the most impactful predictors and enhance its transparency. Additionally, the maximum iteration limit is set to 2000,



balancing computational efficiency with sufficient convergence time for the optimizer. These carefully chosen parameters ensure that the STLSQ algorithm effectively extracts the underlying dynamics of the system while maintaining interpretability and computational efficiency.



**Figure 3:** Updated growth rate model (integrated model versus actual specific density) showing better performance than original CTMI-light model.

We use a combination of custom, Polynomial, and Fourier libraries in SINDy. The custom library included logistic growth, exponential growth, Gompertz, Allee effect, Michaelis-Menten kinetics, and Holling's Type II functions, with all parameters uniformly set to one. This simplified configuration allowed for a consistent and straightforward interpretation of the model's outcomes while retaining the flexibility to model complex biological phenomena. The Polynomial Library, configured to include terms up to the second degree, enables the model to identify both linear and quadratic relationships. Meanwhile, the Fourier Library plays a key role in capturing remaining periodicity.

An ensemble method is utilized to enhance the model's robustness and accuracy [26]. This approach involves fitting the model multiple times using subsets of the training data, allowing for a comprehensive capture of potential dynamics and mitigating the impact of outliers or anomalous data points. The ensemble method generates 1000 models, each trained on a random subset comprising half of the training data.

## Model Performance

A SINDy model is trained for each of the 20 validation folds and the error metric scores for RMSE, MAPE, and  $R^2$  are tabulated in **Table 2**. Scores across the early folds reflect completely unbounded and divergent behavior in the corrected model when integrated over time, due to the limited training data up to that point. Scores continue to improve (RMSE and MAPE decreasing and  $R^2$  increasing) until Fold 6, where a measurement fluke in specific density temporarily biases the training data. These high RMSE and MAPE and highly negative  $R^2$  are seen again in Folds 14 and 15, reflecting completely divergent behavior likely the result of similar poor data quality around that point in time.

**Table 2:** Error metric scores across full 20 validation folds.

Fold #	Uncorrected			Corrected		
	RMSE	MAPE	$R^2$	RMSE	MAPE	$R^2$
1	1.15e+8	2.31e+10%	-7.52e+17	1.055e+12	1.50e+14%	-6.29e+25
2	0.075	5.353%	0.585	1.80e+10	1.25e+12%	-2.38e+22
3	0.0925	inf%	0.160	1.64e+11	inf%	-2.64e+24
4	0.0300	1.34%	0.876	0.399	18.7%	-20.9
5	0.117	4.62%	0.328	0.867	32.5%	-36.1
6	0.0969	3.77%	0.0973	3.92	125%	-1475
7	0.0884	2.75%	0.361	0.0817	2.71%	0.455
8	0.0364	1.06%	0.805	0.0909	1.95%	-0.215
9	0.0827	1.80%	-4.90	0.0479	1.28%	-0.981
10	0.055	1.40%	0.639	0.0684	1.47%	0.443
11	0.0235	0.541%	0.913	0.0480	1.16%	0.635
12	0.0595	1.35%	-1.17	0.0132	0.278%	0.893
13	0.0138	0.242%	0.948	0.0263	0.549%	0.811
14	0.120	2.37%	-310	2.53e+5	3.76e+6%	-1.38e+15
15	0.0949	1.90%	-29.2	12.1	177%	-492000
16	0.0720	1.39%	-7.01	0.223	3.97%	-76.1
17	0.0844	1.65%	-10.8	0.216	4.14%	-76.5
18	0.0144	0.260%	0.942	0.0642	1.19%	-0.144
19	0.0192	0.326%	0.873	0.0484	0.806%	0.197
20	0.0586	0.964%	-2.34	0.109	1.81%	-10.6

The fold models not exhibiting this fully divergent behavior either outperform the original combined CTMI-light model both visually quantitatively (for example models for Folds 9 and 12, see Table 2) or capture some pattern or periodicity not found in the original mechanistic models, while simultaneously slowly diverging from the actual truth specific density (for example models for Folds 8 and 18, see **Table 2**). In Fold 9, we see that the corrected model outperforms the original mechanistic model even with experimental error due to the drop in

specific density.

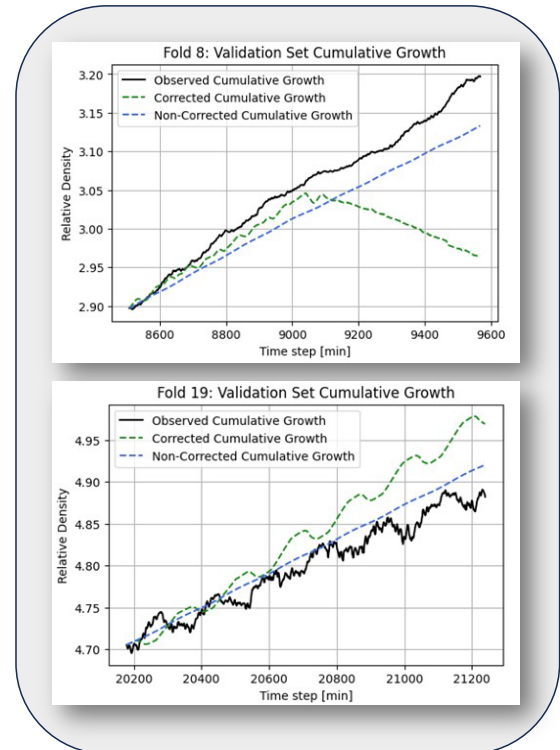
## Model Trends

The SINDy error-based models are provided as Supplementary Material. The natural functions are given only by name and variable they operate on (see Table 1). For example,  $13.512 \text{ gompertz}(Temperature, Light)$  refers to the Gompertz function applied to both temperature and light variables (separately) but with a coefficient of 13.512 for each.

Early fold equation coefficients (Folds 1-7) show noticeable fluctuation in the impact of temperature. Initially, the effect is strongly negative (e.g., -104689.0 in Fold 1) but shows variability in subsequent folds, indicating changing sensitivity to temperature (see Supplementary Material). Light maintains a predominantly negative coefficient, with large magnitudes (e.g., -43231100.0 in Fold 1). pH coefficients show significant changes, starting positive and large (8179580.0 in Fold 1) but becoming negative by Fold 4, indicating a shifting role of pH in these systems. Squared terms like  $Temperature^2$  and  $Light^2$  indicate evolving non-linear effects, with coefficients varying and reflecting changing dynamics. Interaction terms start to appear and vary in significance, suggesting the evolving interplay between factors. Logistic and exponential growth models are applied with varying coefficients, indicating differing growth dynamics under changing conditions.

Folds (Folds 8-14) see the coefficient of Temperature stabilizing somewhat but continuing to show variation, possibly reflecting a nonlinear role not captured by the supplied function basis for SINDy. Light consistently shows a negative impact but with varying magnitudes, indicating a sustained but dynamically changing influence. The role of pH continues to evolve with coefficient changes. Cyclical terms become more pronounced, with sin and cos functions indicating more significant seasonal or cyclic effects.

Late Folds (Folds 15-20) show temperature coefficients become more moderate, suggesting a refined understanding or a stabilizing effect. The influence of Light remains significant with large negative coefficients, indicating a consistently crucial role. pH effects continue to show variability but within a narrower range, suggesting a more consistent but still variable role. The complexity in terms of diverse interactions and non-linear effects shows signs of convergence, with key relationships and factors becoming more defined. There is a tendency towards more consistent use of growth models, indicating a stabilization in how growth or change processes are being represented.



**Figure 4:** Updated growth rate model (integrated model versus specific density) showing divergent behavior.

## CONCLUSION

This study demonstrates the potential of sparse machine learning techniques like SINDy to enhance the accuracy of mechanistic models while retaining interpretability. However, it also highlights the limitations and challenges that can arise in applying data-driven methods to improve mechanistic insights, especially given constrained data availability.

By focusing solely on a single experimental dataset for the microalga *Chlorella vulgaris* cultivated under tightly controlled photobioreactor conditions, the model's flexibility was restricted. While providing high temporal resolution measurements, spanning just one batch limited the diversity of growth dynamics captured. Periodic disturbances in the data, stemming from operational issues, further hampered efforts to train robust models.

Across 20 validation folds, adding the data-driven error dynamics equation resulted in improved error metrics. However, early folds suffered due to insufficient training data leading to a diverging trajectory with exploding error. Later folds demonstrated closer alignment with observed growth trajectory or responded to oscillation over time neglected by the linear mechanistic model. Inaccuracy in the phase of the oscillatory trajectory predicted may be due to sensitivity to initial conditions and warrants further exploration. The variation in types of biological growth functions applied over folds potentially indicated the model homing in on appropriate

representations for different mechanisms or growth phases as training data and timespan was increased.

These outcomes highlight the power of refining first-principles models with nuanced insights from real system observations. However, the instability across folds underscores the need for more comprehensive datasets before operational deployment. Future efforts should focus on generating measurements encapsulating a wider array of cultivation conditions and durations. Increased scale and variance in datasets may enable machine learning to extract more robust patterns. Beyond SINDy, exploring alternative algorithms could further improve the mechanistic correction.

## DIGITAL SUPPLEMENTARY MATERIAL

The full set of error ODEs generated can be found here. Link: [[Error\\_ODEs](#)].

## ACKNOWLEDGEMENTS

The author William Farlessyost acknowledges the National Science Foundation for support under the Graduate Research Fellowship Program (GRFP) undergrant number DGE-1842166. Thanks to Sebastian Oberst (UTS) for his help and support on this project.

## REFERENCES

1. L. Mears, S. M. Stocks, M. O. Albaek, G. Sin, and K. V. Gernaey, "Mechanistic Fermentation Models for Process Design, Monitoring, and Control," *Trends Biotechnol*, vol. 35, no. 10, pp. 914–924, Oct. 2017, doi: 10.1016/J.TIBTECH.2017.07.002.
2. M. Quade, M. Abel, J. Nathan Kutz, and S. L. Brunton, "Sparse identification of nonlinear dynamics for rapid model recovery," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 28, no. 6, p. 63116, 2018.
3. S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *Proceedings of the national academy of sciences*, vol. 113, no. 15, pp. 3932–3937, 2016.
4. E. Kaiser, J. N. Kutz, and S. L. Brunton, "Sparse identification of nonlinear dynamics for model predictive control in the low-data limit," *Proceedings of the Royal Society A*, vol. 474, no. 2219, p. 20180335, 2018.
5. F. Harirchi, "On sparse identification of complex dynamical systems: A study on discovering influential reactions in chemical reaction networks," *Fuel*, vol. 279, p. 118204, 2020.
6. G. Towler and R. Sinnott, *Chemical Engineering Design: Principles, Practice and Economics of Plant and Process Design*. Elsevier, 2021. doi: 10.1016/B978-0-12-821179-3.01001-3.
7. M. Sharifzadeh, "Integration of process design and control: A review," *Chemical Engineering Research and Design*, vol. 91, no. 12, pp. 2515–2549, Dec. 2013, doi: 10.1016/J.CHERD.2013.05.007.
8. B. W. Bequette, "Prentice Hall PTR: Process Control: Modeling, Design, and Simulation," 2002.
9. K. V. Gernaey, A. E. Lantz, P. Tufvesson, J. M. Woodley, and G. Sin, "Application of mechanistic models to fermentation and biocatalysis for next-generation processes," *Trends Biotechnol*, vol. 28, no. 7, pp. 346–354, Jul. 2010, doi: 10.1016/J.TIBTECH.2010.03.006.
10. S. Esplugas, S. Contreras, and D. F. Ollis, "Engineering Aspects of the Integration of Chemical and Biological Oxidation: Simple Mechanistic Models for the Oxidation Treatment," *Journal of Environmental Engineering*, vol. 130, no. 9, pp. 967–974, Sep. 2004, doi: 10.1061/(ASCE)0733-9372(2004)130:9(967).
11. R. T. Kapoor, M. Danish, R. S. Singh, M. Rafatullah, and A. K. Abdul, "Exploiting microbial biomass in treating azo dyes contaminated wastewater: Mechanism of degradation and factors affecting microbial efficiency," *Journal of Water Process Engineering*, vol. 43, p. 102255, Oct. 2021, doi: 10.1016/J.JWPE.2021.102255.
12. A. K. Datta, "Toward computer-aided food engineering: Mechanistic frameworks for evolution of product, quality and safety during processing," *J Food Eng*, vol. 176, pp. 9–27, May 2016, doi: 10.1016/J.JFOODENG.2015.10.010.
13. Q. Lu and F. Jiao, "Electrochemical CO<sub>2</sub> reduction: Electrocatalyst, reaction mechanism, and process engineering," *Nano Energy*, vol. 29, pp. 439–456, Nov. 2016, doi: 10.1016/J.NANOEN.2016.04.009.
14. M. Mozaffar *et al.*, "Mechanistic artificial intelligence (mechanistic-AI) for modeling, design, and control of advanced manufacturing processes: Current state and perspectives," *J Mater Process Technol*, vol. 302, p. 117485, Apr. 2022, doi: 10.1016/J.JMATPROTEC.2021.11748.
15. J. Huang, Y. Chen, S. Jiang, and C. Yuan, "Machine learning-enabled intelligent modeling and optimization of chemical processes," *Chem Eng Sci*, vol. 203, pp. 290–310, 2019.
16. H.-J. Song, D.-Y. Jeong, and J.-H. Ryu, "Prediction of chemical reaction yield using nonparametric regression and machine learning methods," *Ind Eng Chem Res*, vol. 54, no. 44, pp. 10868–10876, 2015.
17. C. Safi, B. Zebib, O. Merah, P. Y. Pontalier, and C. Vaca-Garcia, "Morphology, composition, production, processing and applications of *Chlorella vulgaris*: A review," *Renewable and*

- Sustainable Energy Reviews*, vol. 35, pp. 265–278, Jul. 2014, doi: 10.1016/J.RSER.2014.04.007.
18. M. T. Ahmad, M. Shariff, F. Md. Yusoff, Y. M. Goh, and S. Banerjee, “Applications of microalga *Chlorella vulgaris* in aquaculture,” *Rev Aquac*, vol. 12, no. 1, pp. 328–346, Feb. 2020, doi: 10.1111/RAQ.12320.
  19. M. F. Blair, B. Kokabian, and V. G. Gude, “Light and growth medium effect on *Chlorella vulgaris* biomass production,” *J Environ Chem Eng*, vol. 2, no. 1, pp. 665–674, Mar. 2014, doi: 10.1016/J.JECE.2013.11.005.
  20. K. Rezaei, A. Javanshir, R. Barghbani, K. Rezaei, and A. Javanshir, “Investigating the Effects of Several Parameters on the Growth of *Chlorella vulgaris* Using Taguchi’s Experimental Approach,” *Article in International Journal of Biotechnology for Wellness Industries*, vol. 1, pp. 128–133, 2012, doi: 10.6000/1927-3037/2012.01.02.04.
  21. S. L. Brunton, J. L. Proctor, and J. N. Kutz, “Sparse identification of nonlinear dynamics with control (SINDYc),” *IFAC-PapersOnLine*, vol. 49, no. 18, pp. 710–715, 2016.
  22. P. Waliszewski and J. Konarski, “A Mystery of the Gompertz Function,” *Fractals in Biology and Medicine*, pp. 277–286, Jan. 2005, doi: 10.1007/3-7643-7412-8\_27.
  23. J. C. Gascoigne and R. N. Lipcius, “Allee effects driven by predation,” *Journal of Applied Ecology*, vol. 41, no. 5, pp. 801–810, Oct. 2004, doi: 10.1111/J.0021-8901.2004.00944.X.
  24. B. Srinivasan, “A guide to the Michaelis–Menten equation: steady state and beyond,” *FEBS J*, vol. 289, no. 20, pp. 6086–6098, Oct. 2022, doi: 10.1111/FEBS.16124.
  25. H. N. Agiza, E. M. ELabbasy, H. EL-Metwally, and A. A. Elsadany, “Chaotic dynamics of a discrete prey–predator model with Holling type II,” *Nonlinear Anal Real World Appl*, vol. 10, no. 1, pp. 116–129, Feb. 2009, doi: 10.1016/J.NONRWA.2007.08.029.
  26. U. Fasel, J. N. Kutz, B. W. Brunton, and S. L. Brunton, “Ensemble-SINDy: Robust sparse model discovery in the low-data, high-noise limit, with active learning and control,” *Proceedings of the Royal Society A*, vol. 478, no. 2260, 2022, doi: 10.1098/RSPA.2021.0904.

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Optimal Design of Antibody Extraction Systems using Protein A Resin with Multicycling

Fred Ghanem<sup>a</sup>, Purnima M. Kodate<sup>b</sup>, Gerard M. Capellades<sup>a</sup>, and Kirti M. Yenkie<sup>a\*</sup>

<sup>a</sup> Rowan University, Department of Chemical Engineering, Glassboro, NJ, USA

<sup>b</sup> Department of Pathology, Government Medical College, Nagpur, India

\* Corresponding Author: yenkie@rowan.edu.

## ABSTRACT

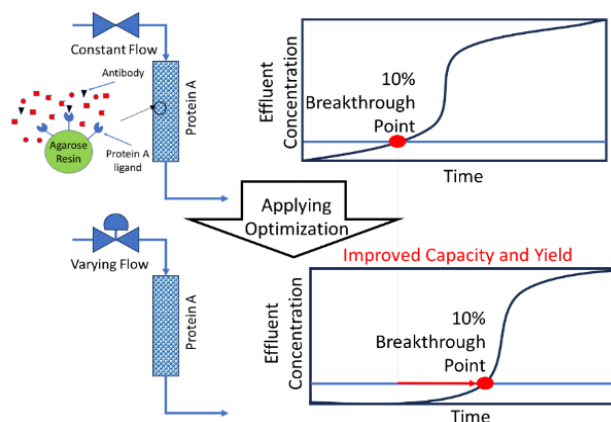
Antibody therapies are important in treating life-threatening ailments such as cancer and autoimmune diseases. Purity of the antibody is essential for successful applications and Protein A selective resin extraction is the standard step for antibody recovery. Unfortunately, such resins can cost up to 30% of the total cost of antibody production. Hence, the optimal design of this purification step becomes a critical factor in downstream processing to minimize the size of the column needed. An accurate predictive model, as a digital twin representing the purification process, is necessary where changes in the flow rates and the inlet concentrations are modeled via the Method of Moments. The system uncertainties are captured by including the stochastic Ito process model of Brownian motion with drift. Pontryagin's Maximum Principle under uncertainty is then applied to predict the flowrate control strategy for optimized resin use, column design, and efficient capturing of the antibodies. In this study, the flow rate is controlled to optimize the process efficiency via maximizing the theoretical plate number with time, the objective for efficient resin usage within a fixed-size column. This work successfully achieved optimality, which was also confirmed via experimentation, leading to higher antibody resin loading capacity. When the work was expanded to 200 cycles of Protein A usage, significant improvements in the downstream process productivity were achieved allowing for smaller footprint columns to be used.

**Keywords:** Optimization, Dynamic Modeling, Stochastic Optimization, Model Reduction, Antibody Extraction.

## INTRODUCTION

Antibodies are produced by many pharmaceutical companies to treat various autoimmune diseases [1]. Unfortunately, up to 30% of the production cost of the antibody [2] is based on a single purification step, the Protein A resin, that selectively captures the antibody from the fermentation broth with up to 99% purity. Therefore, the optimal usage of such an expensive resin is necessary. Since the cost of antibody production depends on the cost of this resin, an optimal control strategy is needed to maximize the efficient use of the resin as seen in Figure 1 giving the possibility to use smaller systems.

The effluent concentration of the expensive capture step will also follow a sigmoidal function [3] as seen in Figure 1. With many sigmoidal models used to evaluate packed bed systems [4], the model parameters are highly dependent on flow rates and inlet concentrations [5].

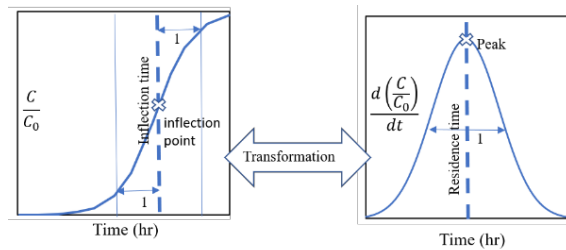


**Figure 1.** Optimal design and operating policy for the Protein A resin step in antibody extraction.



A generalized fixed bed model can be used where the parameters are defined by their relationships to the flow rate and inlet concentration [4]. Pontryagin's maximum principle [6] defines our purification system as requiring an optimization strategy via adjustments of controllable process variables such as flow rate.

To accomplish such optimality, the method of moments is used to simplify the system as a univariate function of time [7]. However since the moments are typically used to represent chromatographic steps [8] with a population distribution curve, the Protein A capture step cumulative curve will need to be transformed via its derivative into a distribution curve (as shown in Figure 2). The Hamiltonian function [9] is then formulated using moments as state variables and introducing adjoint variables to define the path taken during the optimization via maximizing the theoretical plate number [10] for the most efficient use of the resin.



**Figure 2.** Sigmoidal function transformation from a cumulative to a distribution curve via its derivative.

Since the inlet concentrations and model parameters can vary during production, capturing such uncertainties during optimization is important. The stochastic Ito process [11] was picked to capture the variability of the system. The purpose of this work is to confirm that varying the flow rate with time maximizes the resin performance over using fixed flow rates, a standard practice in existing industrial chromatography column operations [12]. While some industries adopted continuous chromatography [13], such as the use of simulated moving beds [14], to maximize the efficient use of the resin, considerable investments in such complex systems are required [15]. The novelty of our approach lies in its simplicity in maximizing the system performance of a single-column system avoiding the high cost of a continuous system.

## METHODOLOGY

The procedure consists of feeding IgG4 antibody [16] through 1 ml of packed Minichrom column (supplied by Repligen) packed with Praesto Jetted A50 Protein A resin (supplied by Purolite). Three different concentrations of the antibody (2, 5, and 7 mg/ml) were used at 4 different residence times (3, 4.8, 6, and 10 min). A standard sigmoidal function [17], predicting the column output is seen below.

$$\frac{C}{C_0} = \frac{1}{1 + \exp(a - b t)} \quad (1)$$

where  $C_0$  - inlet concentration,  $C$  - effluent concentration varying with time,  $a$  and  $b$  - empirical parameters of the sigmoidal function, and  $t$  - time. The model can be further modified into a modified Yoon-Nelson equation [18] where  $t_{\frac{1}{2}}$  - time needed to get to  $C/C_0 = 1/2$ .

$$\frac{C}{C_0} = \frac{1}{1 + \exp\left(b\left(\frac{a}{b} - t\right)\right)} = \frac{1}{1 + \exp\left(b\left(t_{\frac{1}{2}} - t\right)\right)} \quad (2)$$

The parameters in the model can be calculated by minimizing the Sum of Squares of the Error (SSE) between model predictions and experimental measurements [19]. Since the sigmoidal function is represented by a cumulative curve, it is necessary to transform Eq. 2 into a distribution curve, closely depicting chromatographic separations [20]. Such a transformation is achieved by taking the derivative of the sigmoidal model,  $d(C/C_0)/dt$ .

## The Method of Moments

The general form of the truncated temporal moments [21] used in chromatographic separations is defined by the following equation [22]:

$$m_i = \int_0^t t^i \frac{d\left(\frac{C}{C_0}\right)}{dt} dt \quad (3)$$

Eq. 3 is the basic form to present the 4 moments describing our system [23]. In chromatographic operations, the truncated temporal moment equations are represented as follows [24]:

$$\begin{aligned} \text{Zeroth moment} &= m_0(t) = \mu_0(t) \\ &= \int_0^t \frac{d\left(\frac{C}{C_0}\right)}{dt} dt = \frac{C}{C_0}(t) = y_1(t) \end{aligned} \quad (4)$$

$$\begin{aligned} \text{First normalized moment} &= \mu_1(t) \\ &= \frac{\int_0^t t \frac{d\left(\frac{C}{C_0}\right)}{dt} dt}{\int_0^t \frac{d\left(\frac{C}{C_0}\right)}{dt} dt} = \frac{\int_0^t t \frac{d\left(\frac{C}{C_0}\right)}{dt} dt}{\mu_0(t)} = t_m(t) = y_2(t) \end{aligned} \quad (5)$$

$$\begin{aligned} \text{Second central normalized moment} &= \mu_{2c}(t) \\ &= \frac{\int_0^t (t - \mu_1)^2 \frac{d\left(\frac{C}{C_0}\right)}{dt} dt}{\mu_0(t)} = \sigma^2(t) = y_3(t) \end{aligned} \quad (6)$$

$$\begin{aligned} \text{Third central normalized moment} &= \mu_{3c}(t) \\ &= \frac{1}{\mu_{2c}^{1.5}} \frac{\int_0^t (t - \mu_1)^3 \frac{d\left(\frac{C}{C_0}\right)}{dt} dt}{\mu_0(t)} = s^3(t) = y_4(t) \end{aligned} \quad (7)$$

where  $t_m$  - column residence time or the time needed to reach the maximum peak,  $\sigma^2$  - variance or square of the standard deviation, and  $s^3$  - skewness cubed.

## Capturing Uncertainties

As uncertainties can affect the protein A resin performance leading to inefficient removal of the antibodies, it is necessary to account for these variations via the adoption of stochastic process equations and models [25]. One of the simplest stochastic process models is the Brownian motion / Wiener process [26], where one begins at a known initial value and takes equiprobable

steps in any direction [27]. An example of the Brownian motion method is the Ito process with drift [28] defined below as a modified Euler step.

$$y_t = y_{t-\Delta t} + \alpha \Delta t + \kappa \varepsilon (\Delta t)^{1/2} \quad (8)$$

where  $y_t$  – the moment at time  $t$ ,  $y_{t-\Delta t}$  – the moment at time  $t - \Delta t$ ,  $\alpha$  – Ito process drift parameter,  $\kappa$  – Ito process deviation parameter,  $\varepsilon$  – random number with a normal distribution of zero as the mean and a standard deviation of 2, and  $\Delta t$  – time step. The Ito process has been applied to other modes of purification such as crystallization [29] and distillation [30].

### Objective function and boundary conditions

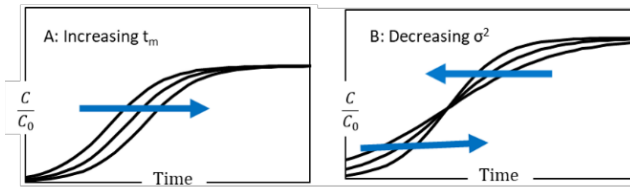
Figure 3 demonstrates that an increase in the normalized first moment,  $t_m$ , results in more antibody extraction, while a decrease in the second central moment,  $\sigma^2$ , results in lower leakage. Therefore, the objective would be to maximize  $t_m$  and minimize the variance  $\sigma^2$ . In chromatography, better purification is achieved when the theoretical plate number (TPN) is maximized [31].

$$TPN = \frac{t_m^2}{\sigma^2} \quad (9)$$

The objective,  $J_{max}$ , is to maximize the theoretical plate number.

$$J_{max(Q)} = \max\left(\frac{t_m^2}{\sigma^2}\right) = \max\left(\frac{\mu_1^2}{\mu_{2c}}\right) \quad (10)$$

The maximum value is achieved when the derivative of the objective function is equated to zero.



**Figure 3.** Effect of increasing the residence time (A) compared to decreasing the variance (B).

### Optimal Control

To simplify the optimization approach, we can linearize the four stochastic moments into a single Hamiltonian function as follows[25]:

$$H = \sum_{i=1}^4 [z_i F_i + \omega_i \frac{g_{y_i}^2}{2}] \quad (11)$$

where  $H$  – Hamiltonian function,  $z_i$  – adjoint variable,  $F_i$  – differential changes of the state variable with time =  $\frac{dy_i}{dt}$ ,

$\omega_i$  – stochastic adjoint variable, and  $g_{y_i} = \sqrt{\frac{var(\Delta y_i)}{\Delta t}}$  (12)

The adjoint variables limit the pathways to achieve the objective and are defined by their derivatives:

$$\frac{dz_i}{dt} = - \sum_{k=1}^4 \left[ \frac{\partial F_k}{\partial y_i} z_k + \frac{1}{2} \frac{\partial g_k^2}{\partial y_i} \omega_k \right] \quad (13)$$

$$\frac{d\omega_i}{dt} = - \sum_{k=1}^4 \left[ 2\omega_k \frac{\partial F_k}{\partial y_i} + \frac{\partial^2 F_k}{\partial y_i^2} z_k + \frac{1}{2} \frac{\partial^2 g_k^2}{\partial y_i^2} \omega_k \right] \quad (14)$$

Pontryagin's maximum principle (PMP) is used to apply the optimal control strategy [32] to our antibody extraction system. PMP utilizes the Hamiltonian by imposing the optimality conditions for the objective function [33]. Another optimization method that could be used is the orthogonal collocation method [34]. The collocation numerical method would simplify our optimization strategy for our deterministic approach [35] but is challenging when dealing with uncertainties captured in the Hamiltonian function used in the PMP method.

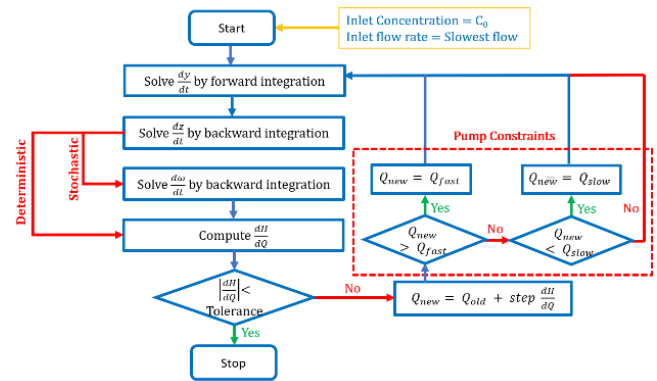
Optimal control is applied using a manageable system parameter such as the flow rate. Derivatization of the Hamiltonian function to the flow rate ( $Q$ ) can be represented by its coordinates (the state and the adjoint variables) using the following expression[29] of a derivative chain rule:

$$\frac{dH}{dQ} = \sum_{i=1}^4 \frac{dH}{dy_i} \frac{dy_i}{dQ} + \sum_{i=1}^4 \frac{dH}{dz_i} \frac{dz_i}{dQ} + \sum_{i=1}^4 \frac{dH}{d\omega_i} \frac{d\omega_i}{dQ} \quad (15)$$

Eq. 15 defines the Hamiltonian gradient where the gradient is nullified when flow rate changes are no longer necessary. The integrations were done using the Euler steps[29]. The initial values of the state variables are obtained from the starting point, and the final values of the adjoint variables ( $z$  and  $\omega$ ) are obtained from the optimal objective for the purification system.

### Optimal Control Solution Strategy

The pump flow rate is bound between 2 flow rates corresponding to residence times of 3 and 10 min in our experimental work. Considering that the inlet antibody concentration enters at 5 mg/ml, the flow rate, initially set at the slow residence time of 10 min, is expected to change to optimize the extraction of the antibody.



**Figure 4.** Flowchart for stochastic optimal control.

Figure 4 summarizes the algorithm using Matlab as the optimization tool for the deterministic and stochastic approach[36]. The main difference between the two

approaches is the bypass loop to calculate the stochastic adjoint variable  $\omega$ . Figure 4 demonstrates how our state and adjoint variables are calculated before calculating the Hamiltonian gradient which is compared to the tolerance picked for the system. Depending on the Hamiltonian gradient value, the optimization would be achieved (below the tolerance) or the flow rate would be adjusted (above the tolerance) between the two flow rate boundaries for another iteration.

## RESULTS

The experimental data for the concentration ratio of antibody igG4 with time from a 1 ml Minichrom column packed with Praesto Jetted A50 resin is seen in Table 1.

**Table 1:** Experimental data.

<i>Inlet Concentration <math>C_0 = 2</math> mg/ml</i>			
Residence Time (min)	Time to $C/C_0 = 0.05$ (min)	Time to $C/C_0 = 0.10$ (min)	Time to $C/C_0 = 0.50$ (min)
3	82.7	90.6	135.2
4.8	157.6	170.3	216.9
6	198.3	216.8	271.3
10	334.8	364.6	456.2
<i>Inlet Concentration <math>C_0 = 5</math> mg/ml</i>			
Residence Time (min)	Time to $C/C_0 = 0.05$ (min)	Time to $C/C_0 = 0.10$ (min)	Time to $C/C_0 = 0.50$ (min)
3	31.5	35.3	57.9
4.8	61.4	67.8	97.8
6	87.4	94.2	123.6
10	162.8	173.2	209.2
<i>Inlet Concentration <math>C_0 = 7</math> mg/ml</i>			
Residence Time (min)	Time to $C/C_0 = 0.05$ (min)	Time to $C/C_0 = 0.10$ (min)	Time to $C/C_0 = 0.50$ (min)
3	23.7	24.9	30.2
4.8	40.9	42.6	49.4
6	52.0	54.1	62.0
10	87.7	92.5	104.9

We extrapolated the sigmoidal model parameters via the minimization of the SSE (sum of square errors).

These parameters are summarized in Table 2 as they vary with changes in inlet concentration and flow rate.

**Table 2:** The sigmoidal function parameters.

Residence Time (min)	Concentration (mg/ml)	Empirical parameter $t_{1/2}$ (min)	Empirical parameter $b$ ( $\text{min}^{-1}$ )
3	2	135.1	0.0519
4.8	2	217.1	0.0480
6	2	271.4	0.0403
10	2	455.8	0.0241
3	5	57.8	0.1031
4.8	5	97.8	0.0762
6	5	123.5	0.0772
10	5	209.2	0.0619
3	7	30.2	0.4317
4.8	7	49.4	0.3344
6	7	61.9	0.2849
10	7	105.0	0.1748

From Table 2, we extrapolate the relationships between the model parameters and system characteristics.

$$t_{\frac{1}{2}} = 82.81 RT \exp(-0.295 C_0) \quad (16)$$

$$b = \frac{1}{-0.57 RT C_0 + 4.49 RT - 1.12 C_0 + 9.08} \quad (17)$$

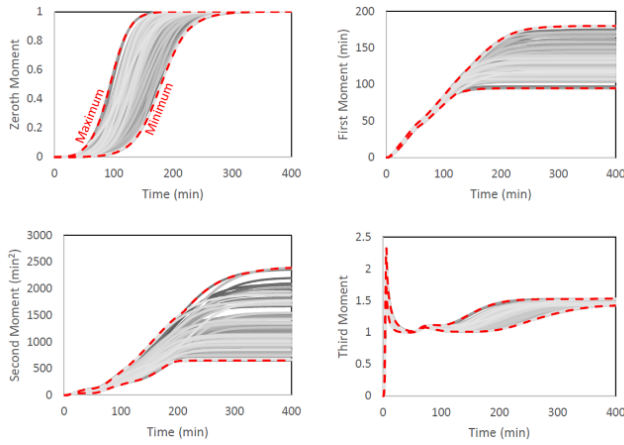
where  $RT$  – residence time =  $RV/Q$ , and  $RV$  – resin volume in the column. The results comparing the empirical parameters with the calculated data from Eq. 16 and Eq. 17 are summarized in Table 3. It displays that the fluctuations in the parameters  $t_{\frac{1}{2}}$  and  $b$  are within  $\pm 10\%$  and  $\pm 20\%$  respectively.

## Capture Uncertainty via Stochastic Processes

When considering an inlet antibody concentration that fluctuates 10% around 5 mg/ml and capturing the parameter fluctuations described by Table 3, we can display the effect of these uncertainties on the 4 moments.

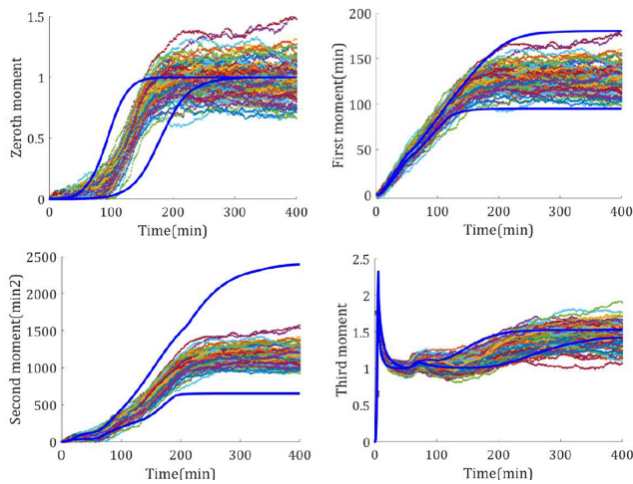
**Table 3:** Results of the sigmoidal function parameters for calculated versus empirical.

Residence Time (min)	Concentration (mg/ml)	Empirical parameter $t_{1/2}$ (min)	Empirical parameter $b$ ( $\text{min}^{-1}$ )	Calculated parameter $t_{1/2}$ (min)	Calculated parameter $b$ ( $\text{min}^{-1}$ )	Percent difference on $t_{1/2}$	Percent difference on $b$
3	2	135.1	0.0519	137.7	0.0592	-1.9	-14.1
4.8	2	217.1	0.0480	220.3	0.0436	-1.5	9.1
6	2	271.4	0.0403	275.4	0.0371	-1.5	7.9
10	2	455.8	0.0241	459.0	0.0248	-0.7	-2.9
3	5	57.8	0.1031	56.8	0.1190	1.7	-15.5
4.8	5	97.8	0.0762	90.9	0.0881	7.0	-15.6
6	5	123.5	0.0772	113.7	0.0751	8.0	2.8
10	5	209.2	0.0619	189.5	0.0503	9.4	18.7
3	7	30.2	0.4317	31.5	0.3650	-4.4	15.5
4.8	7	49.4	0.3344	50.4	0.2747	-2.1	17.8



**Figure 5.** Static uncertainty (100 scenarios via sampling) manifested as Dynamic uncertainty.

Figure 5 shows the effect of static uncertainties in the initial condition or process parameters manifesting as dynamic uncertainties in all the state variables (i.e. the 4 moments with respect to time). Figure 6 demonstrates that the Ito process of Brownian motion with drift can capture the dynamic uncertainties of the 4 moments.



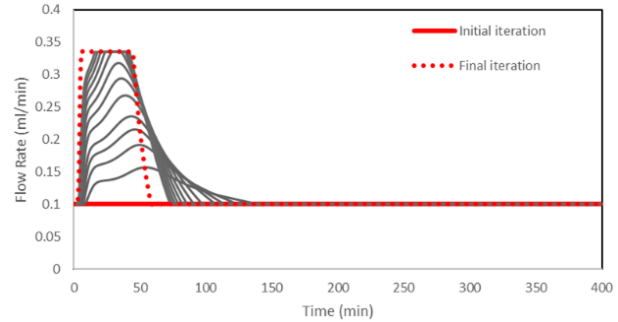
**Figure 6.** Ito process capturing the uncertainties for the four moments.

### Results from Stochastic Optimal Control

We run the Matlab flowchart program seen in Figure 4, for over 40,000 iterations to achieve the optimal flow rate path. While the state variables,  $y_i$ , are integrated in the forward direction from the initial conditions,  $[y_1(0) \ 0 \ 0 \ 0]$ , the adjoint variables,  $z_i$  and  $\omega_i$ , are integrated in a backward direction from the final conditions,  $[0 \ (2\mu_1(t_{final}))/(\mu_{2c}(t_{final})) - (\mu_1^2(t_{final}))/(\mu_{2c}^2(t_{final})) \ 0]$  and  $[0 \ 0 \ 0 \ 0]$ , respectively based on the final optimum of Eq. 11. We use Euler's method for integration with a very small-time step. At each time step, the values of the Hamiltonian and its derivative are

calculated.

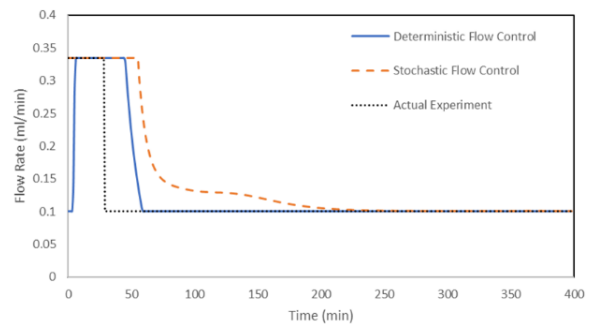
Two approaches were evaluated: a deterministic approach, ignoring the uncertainties, and a stochastic approach. During each of the iterations, the flow rate is adjusted until the optimal flow is achieved. Figure 7 displays the changes in the flow rates during the deterministic optimal control approach. It shows that deterministic optimal flow control will start by increasing the flow towards the maximum flow to get as much product processed before it reduces the flow to avoid early antibody breakthrough.



**Figure 7.** Iterations of the flow rate to achieve deterministic optimal control.

### Experimental Validation of Optimal Control

From Figure 8, Both the stochastic optimal flow and the deterministic optimal flow will adjust to the maximum flow rate initially before reducing it to minimum flow rate around 70 minutes. The deterministic optimal flow decreases slightly earlier to the slow flow rate of around 60 minutes. Since running the process at the fast flow rate initially followed by a slowdown improved the process performance, we ran an experiment to confirm the improvement by manually changing the flow from the fastest flow rate to the slowest flow rate after  $C/C_0 = 0.05$  is reached.



**Figure 8.** Optimal flow control for both deterministic and stochastic approaches with manual flow rate adjustment.

Table 4 demonstrates the effectiveness of using an optimal flow rate to maximize the antibody extraction over using a constant flow rate. Both stochastic and

deterministic approaches succeeded in improving the system performance by extracting more antibodies in less time than a fixed slow flow rate. The actual experiment also demonstrated an improvement in antibody extraction over running at a constant flow rate. While the experimental manually controlled work did not duplicate the data obtained from the deterministic or the stochastic approach, the results confirmed that designing targeted operating policies such as the fast flow followed by decreased flow later in the process, will achieve more antibody extraction.

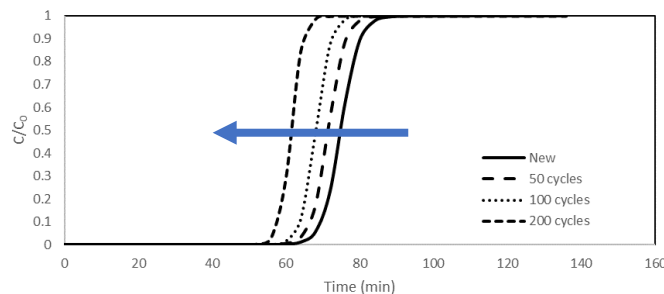
**Table 4:** Optimal control results versus constant flow.

Conditions	Max Flow	Min Flow	Deterministic	Stochastic	Actual test
Antibody Extracted to $C/Co = 0.1$ mg (mg antibody/ml resin)	65.0	73.6	97.1	105.7	105.3
Antibody Extracted to $C/Co = 0.5$ mg (mg antibody/ml resin)	86.0	92.9	112.8	122.6	113.4

When checking the amount of antibody extracted until 10% breakthrough, Table 4 demonstrates that the flow rate adjustment used in the experimental validation resulted in 43% more antibody removed than using the slowest flow rate setting. This validated improvement in the resin capacity leads to efficient utilization of the expensive Protein A resin. As the system residence time is influenced by the size of the column used, the optimization allows for the efficient utilization of any size resin column within the limitations of the system flow rates.

### Multicycle Optimization

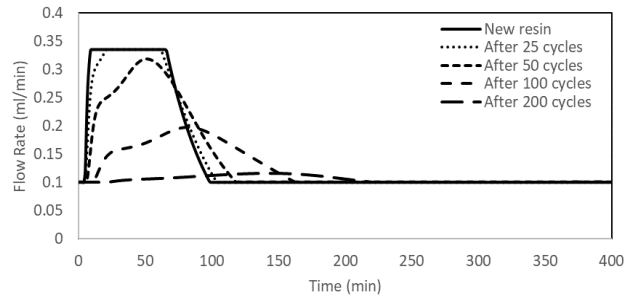
Since the Protein A resin is the most expensive step in the downstream processing of antibodies, cycling it for multiple uses is necessary to get the most value out of its resin life. But with every cycle, the Protein A resin will degrade due to caustic cleaning or due to fouling from the various molecules in the fermentation broth.



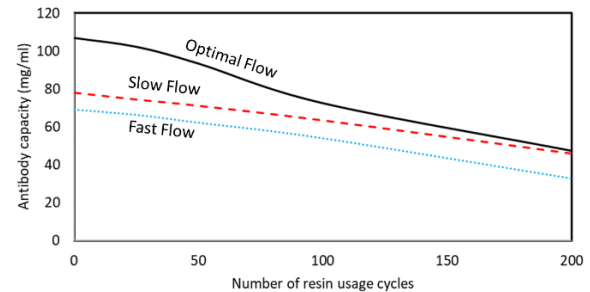
**Figure 9.** Protein A resin degradation during multicycle operations at 6 min residence time.

Figure 9 shows the actual effect of Protein A degradation during multiple cycles as the breakthrough tends

to occur sooner in late cycles (also called life time study). For example, the time to reach 50% breakthrough dropped from 75 min to 68 minutes after 100 cycles. While the model parameters change with contact time and inlet concentration (as seen in equations 20 and 21), further modification of these parameters can be made to depict the degradation with each cycle. Therefore, it is expected that optimal flow rate pathways might be different as the resin degrades.



**Figure 10.** Optimal flow rate during multicycle use



**Figure 11.** Effect of resin degradation on the optimal flow

Figure 10 confirms that hypothesis showing that the system will not be running at the high flow rates for as long as when using a newer resin. At the end of the 200 cycles, the optimal flow rate is to keep the flow constant at its lowest setting. Figure 11 confirms that the optimal flow rate will get closer to the slowest flow rate as the number of cycles increases. Once optimal conditions cannot be improved beyond running at the slowest flow rates, resin replacement becomes the next logical step.

Table 5 summarizes the effect of using the optimal flow rate over 200 cycles of Protein A resin usage. Table 5 demonstrates that over 20% increase in cumulative binding capacity, over 30% savings on cycle time, and over 80% in productivity can be achieved during the 200-cycle life of the Protein A resin. The impact of such improvement allows the use of any size column while avoiding system overdesign usually necessary for similar results. An 84% improvement in productivity translates to using 46% less resin to get the same performance as the slowest flow rate and therefore 46% less footprint. In addition, reducing the loading time by 34% allows the extracted antibody to have less contact time with the host



cell proteins containing the damaging enzymes that can reduce the process yield [37].

**Table 5:** Performance Improvement in the Resin Usage Capacity over 200 Cycles.

Dynamic binding capacity (total grams/ml-resin)	C/C <sub>0</sub> = 10%	C/C <sub>0</sub> = 50%
Optimal control for 200 cycles	15.2	19.6
Constant low flow at residence time = 10 min for 200 cycles	12.5	16.8
% Improvement over 200 cycles	21.6	16.7
Total time of loading production needed (hours)		
Optimal control for 200 cycles	351	565
Constant low flow at residence time = 10 min for 200 cycles	531	772
% Improvement (200 cycles)	33.9	26.8
Productivity (total grams/liter-resin-hour)		
Optimal control for 200 cycles	43.3	34.7
Constant low flow at residence time = 10 min for 200 cycles	23.5	21.8
% Improvement over 200 cycles	84.2	59.2

## CONCLUSIONS

Protein A resins are the most expensive step in the manufacturing of critical antibody therapies. Successful control strategies become necessary to maximize the efficiency of resin usage in downstream processing. When applying optimal control strategies via flow rate adjustment, we demonstrated through prediction at first and then confirmed through experimentation that optimal design is achieved by varying the flow rates rather than using a fixed flow rate. Optimality is achieved by initially running the resin column at high flow rates to get as much feed processed before reducing the flow to avoid an early breakthrough. During the design of multicycle operation of the Protein A resin, the optimal results show a decreasing trend until it reaches similar results as running at the slowest flow rate. However, optimization during the multicycle operation was able to achieve over 80% productivity in the use of Protein A resin, leading to better column design.

With the successful implementation of such an optimal strategy, we demonstrate that smaller designs can be used more efficiently while expecting savings in resins, buffers/water, and utilities that accompany productivity improvements.

## ACKNOWLEDGEMENTS

The authors thank the Department of Chemical Engineering at Rowan University for their assistance in acquiring computational licenses and resources, and Puro-lite for providing the necessary data and testing the results predicted in this study.

## REFERENCES

1. A. Puthenpurail, H. Rathi, S. M. Nauli, and A. Ally, "A Brief Synopsis of Monoclonal Antibody for the Treatment of Various Groups of Diseases.," 2022.
2. D. Stanton, "Protein A's mAb capture efficacy offsets cost, says GE launching new offering." *www.biopharma-reporter.com*, Sep. 26, 2017. [Online]. Available: [https://www.biopharma-reporter.com/Article/2017/09/26/Protein-A-s-efficacy-offsets-cost-says-GE-as-it-launches-new-offering?utm\\_source=copyright&utm\\_medium=OnSite&utm\\_campaign=copyright](https://www.biopharma-reporter.com/Article/2017/09/26/Protein-A-s-efficacy-offsets-cost-says-GE-as-it-launches-new-offering?utm_source=copyright&utm_medium=OnSite&utm_campaign=copyright)
3. H. Motulsky and A. Christopoulos, *Fitting Models to Biological Data using Linear and Nonlinear Regression*. 2003.
4. F. Ghanem, S. S. Jerpoth, and K. M. Yenkie, "Improved Models for Chromate Removal Using Ion Exchangers in Drinking Water Applications," *J. Environ. Eng.*, vol. 148, no. 5, May 2022, doi: 10.1061/(ASCE)EE.1943-7870.0001997.
5. M. Korkmaz, C. Özmetin, E. Özmetin, and Y. Süzen, "Modelling of boron removal from solutions by ion exchange for column reactor design in boron mine wastewater treatment," *DESALINATION WATER Treat.*, vol. 179, pp. 63–74, 2020, doi: 10.5004/dwt.2020.25041.
6. I. M. Ross, *A primer on Pontryagin's principle in optimal control*, 2nd Edition. Collegiate Publishers, 2009. [Online]. Available: <https://search.library.wisc.edu/catalog/9910894537702121>
7. S. Misra, M. F. Wahab, D. C. Patel, and D. W. Armstrong, "The utility of statistical moments in chromatography using trapezoidal and Simpson's rules of peak integration," *J. Sep. Sci.*, vol. 42, no. 8, pp. 1644–1657, Apr. 2019, doi: 10.1002/jssc.201801131.
8. G. Carta, A. R. Ubiera, and T. M. Pabst, "Protein Mass Transfer Kinetics in Ion Exchange Media: Measurements and Interpretations," *Chem. Eng. Technol.*, vol. 28, no. 11, pp. 1252–1264, Nov. 2005, doi: 10.1002/ceat.200500122.
9. P. Hamill, *Lagrangians and Hamiltonians*, 4th printing. Cambridge University Press, 2018.
10. A. Staby, I. H. Jensen, and I. Møllerup, "Comparison of chromatographic ion-exchange resins I. Strong anion-exchange resins," *J Chromatogr A*, p. 13, 2000.
11. V. Rico-Ramirez and U. M. Diwekar, "Stochastic maximum principle for optimal control under uncertainty," *Comput. Chem. Eng.*, p. 5, 2004.
12. M. Hedhammar, A. E. Karlström, and S. Hober, *Chromatographic methods for protein purification*. AlbaNova University Center, Dept. of

- Biotechnology, SE-106 91 Stockholm, Sweden: Royal Institute of Technology, 2006.
13. B. A. Patel *et al.*, "On-Line Ion Exchange Liquid Chromatography as a Process Analytical Technology for Monoclonal Antibody Characterization in Continuous Bioprocessing," *Anal. Chem.*, vol. 89, no. 21, pp. 11357–11365, Nov. 2017, doi: 10.1021/acs.analchem.7b02228.
  14. J. Dawson, "A Review of Chromatography Methods for the Large Scale Downstream Protein Purification of Monoclonal Antibodies," 2019.
  15. S. Engell and A. Toumi, "Optimisation and control of chromatography," *Comput. Chem. Eng.*, vol. 29, no. 6, pp. 1243–1252, May 2005, doi: 10.1016/j.compchemeng.2005.02.034.
  16. S. Arora, B. V. Ayyar, and R. O'Kennedy, "Affinity Chromatography for Antibody Purification," in *Protein Downstream Processing*, vol. 1129, N. E. Labrou, Ed., in *Methods in Molecular Biology*, vol. 1129, Totowa, NJ: Humana Press, 2014, pp. 497–516. doi: 10.1007/978-1-62703-977-2\_35.
  17. N. Kyurkchiev and S. Markov, "Sigmoidal Functions: Some Computational and Modelling Aspects," *Biomath Commun.*, vol. 1, no. 2, Mar. 2015, doi: 10.11145/j.bmc.2015.03.081.
  18. E. R. Arenas-Flores *et al.*, "Treatment of a mining water effluent for the sorption of heavy metal ions using polymer composite beads in a continuous column," *Water Environ. J.*, p. wej.12710, Apr. 2021, doi: 10.1111/wej.12710.
  19. S. Biswas and U. Mishra, "Continuous Fixed-Bed Column Study and Adsorption Modeling: Removal of Lead Ion from Aqueous Solution by Charcoal Originated from Chemical Carbonization of Rubber Wood Sawdust," *J. Chem.*, vol. 2015, pp. 1–9, 2015, doi: 10.1155/2015/907379.
  20. A. Staby *et al.*, "Comparison of chromatographic ion-exchange resins," *J. Chromatogr. A*, vol. 1164, no. 1–2, pp. 82–94, Sep. 2007, doi: 10.1016/j.chroma.2007.06.048.
  21. J. Luo, O. A. Cirpka, and P. K. Kitanidis, "Temporal-moment matching for truncated breakthrough curves for step or step-pulse injection," *Adv. Water Resour.*, vol. 29, no. 9, pp. 1306–1313, Sep. 2006, doi: 10.1016/j.advwatres.2005.10.005.
  22. M. N. Goltz and P. V. Roberts, "Using the method of moments to analyze three-dimensional diffusion-limited solute transport from temporal and spatial perspectives," *Water Resour. Res.*, vol. 23, no. 8, pp. 1575–1585, Aug. 1987, doi: 10.1029/WR023i008p01575.
  23. H. S. Fogler, "Residence Time Distributions of Chemical Reactors," in *Elements of chemical reaction engineering*, Fifth edition., Boston: Prentice Hall, 2016, pp. 767–806.
  24. J. W. Jawitz, "Moments of truncated continuous univariate distributions," *Adv. Water Resour.*, vol. 27, no. 3, pp. 269–281, Mar. 2004, doi: 10.1016/j.advwatres.2003.12.002.
  25. S. Abbasi and U. M. Diwekar, "Characterization and stochastic modeling of uncertainties in the biodiesel production," *Clean Technol. Environ. Policy*, vol. 16, pp. 79–94, Mar. 2013.
  26. T. Szabados, "An elementary introduction to the Wiener process and stochastic integrals," *ArXiv10081510 Math*, Aug. 2010, Accessed: Feb. 21, 2021. [Online]. Available: <http://arxiv.org/abs/1008.1510>
  27. D. T. Brereton, in *Stochastic Simulation of Processes, Fields and Structures*, Ulm University: Institute of Stochastics, 2014, pp. 108–121.
  28. U. Diwekar, "Optimal Control and Dynamic Optimization," in *Introduction to Applied Optimization*, vol. 22, in *Springer Optimization and Its Applications*, vol. 22, Boston, MA: Springer US, 2008, pp. 215–277. doi: 10.1007/978-0-387-76635-5.
  29. K. M. Yenkie and U. Diwekar, "Stochastic Optimal Control of Seeded Batch Crystallizer Applying the Ito Process," *Ind. Eng. Chem. Res.*, pp. 108–122, Jun. 2012.
  30. Z. Lei, C. Li, and B. Chen, "Extractive Distillation: A Review," *Sep. Purif. Rev.*, vol. 32, no. 2, pp. 121–213, Jan. 2003, doi: 10.1081/SPM-120026627.
  31. P. Ravisankar, S. Anusha, K. Supriya, and U. A. Kumar, "Fundamental Chromatographic Parameters," *Int J Pharm Sci Rev Res*, vol. 55, no. 09, pp. 46–50, Apr. 2019.
  32. J. Harmand, C. Lobry, A. Rapaport, and T. Sari, *Optimal Control in Bioprocesses: Pontryagin's Maximum Principle in Practice*, First Edition., vol. 3. Wiley, 2019.
  33. Z. Artstein, "Pontryagin Maximum Principle Revisited with Feedbacks," *Eur. J. Control*, vol. 1, pp. 46–54, 2011.
  34. L. C. Young, "Orthogonal collocation revisited," *Comput. Methods Appl. Mech. Eng.*, vol. 345, pp. 1033–1076, 2019, doi: <https://doi.org/10.1016/j.cma.2018.10.019>.
  35. Santosh K Gupta, "Orthogonal Collocation (OC)," in *Numerical Methods for Engineers*, 3rd ed., New Age International (P) Limited, Publishers, 1995, pp. 229–246.
  36. F. Ghanem and K. M. Yenkie, "Optimal Control of Chromate Removal via Enhanced Modeling using the Method of Moments," 2023. [Online]. Available: <https://api.semanticscholar.org/CorpusID:259316597>
  37. S. Y. E, Y. Hu, R. Molden, H. Qiu, and N. Li, "Identification and Quantification of a Problematic

Host Cell Protein to Support Therapeutic Protein Development," *J. Pharm. Sci.*, vol. 112, no. 3, pp. 673–679, Mar. 2023, doi: 10.1016/j.xphs.2022.10.008.

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Exploring Quantum Optimization for Computer-aided Molecular and Process Design

Ashfaq Iftakher<sup>a</sup> and M. M. Faruque Hasan<sup>a,b,\*</sup>

<sup>a</sup> Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX 77843-3122, USA.

<sup>b</sup> Texas A&M Energy Institute, Texas A&M University, College Station, TX 77843-3122, USA.

\* Corresponding Author: [hasan@tamu.edu](mailto:hasan@tamu.edu).

## ABSTRACT

Computer-aided Molecular and Process Design (CAMPD) is an equation-oriented multi-scale decision making framework for designing both materials (molecules) and processes for separation, reaction, and reactive separation whenever material choice significantly impacts process performance. The inherent nonlinearity and nonconvexity in CAMPD optimization models, introduced through the property and process models, pose challenges to state-of-the-art solvers. Recently, quantum computing (QC) has shown promise for solving complex optimization problems, especially those involving discrete decisions. This motivates us to explore the potential usage of quantum optimization techniques for solving CAMPD problems. We have developed a technique for directly solving a class of mixed integer nonlinear programs using QC. Our approach represents both continuous and integer design decisions by a set of binary variables through encoding schemes. This transformation allows to reformulate certain types of CAMPD problems into Quadratic Unconstrained Binary Optimization (QUBO) models that can be directly solved using quantum annealing techniques. We illustrate this technique for the selection of optimal ionic liquids (IL) and the configuration of a reactor-separator process network. We also discuss several challenges that are associated with quantum optimization when solving large scale CAMPD problems.

**Keywords:** Optimization, Multiscale Modeling, Process Design, Quantum Optimization, CAMPD

## INTRODUCTION

Computer-Aided Molecular and Process Design (CAMPD) allows systematic identification of optimal materials and processes, thereby facilitating the transition towards more sustainable and cost-effective chemical processes. CAMPD typically involves solving mixed-integer nonlinear program (MINLP) of the following form:

$$\begin{aligned} \min \quad & C(x_d^{mat}, x_p^{mat}, x_d^{process}, x_p^{process}) \\ \text{s.t.} \quad & g_1(x_d^{mat}, x_p^{mat}) \leq 0 \\ & g_2(x_d^{mat}, x_p^{mat}) = 0 \\ & g_3(x_d^{process}, x_p^{process}) \leq 0 \\ & g_4(x_d^{process}, x_p^{process}) = 0 \\ & g_5(x_p^{mat}, x_p^{process}) \leq 0 \\ & g_6(x_p^{mat}, x_p^{process}) = 0 \end{aligned} \quad (P1)$$

$$x^L \leq x_d^{mat}, x_p^{mat}, x_d^{process}, x_p^{process} \leq x^U,$$

where the decision variables include both discrete and continuous variables. Molecular design is achieved through the optimal choices for a set of design variables ( $x_d^{mat}$ ) that correspond to the molecular fingerprints or groups that constitute the material, as well as performance variables ( $x_p^{mat}$ ) that dictate the material properties. Process design and configuration is achieved through a set of variables ( $x_d^{process}$ ) that represent process design specifications and typically involve both integer variables (such as feed location, number of column stages, etc.) and continuous variables (such as solvent flow rate, reflux ratio, operating pressure, and temperature). The process performance variables ( $x_p^{process}$ ) dictate the separation efficiency, product purity, etc. The objective function  $C$  aims to minimize the overall cost of the process and is subject to a set of constraints. The constraints  $g_1(x_d^{mat}, x_p^{mat}) \leq 0$  and  $g_2(x_d^{mat}, x_p^{mat}) = 0$  include the thermodynamic and transport property

relations and molecular structure-property models. In many cases, empirical structure-property relationships, such as group contribution models [1-2] or trained neural networks [3] are considered. They also contain a set of constraints that describe molecular generation rules to form structurally feasible molecules from a collection of groups or descriptors [4]. For solvent design, for example, these constraints typically involve the octet rule to ensure that the molecule satisfies atom balance. Often, bounds are imposed on the number of groups of a particular type to ensure the generation of a finite number of molecules. The constraints  $g_3(x_d^{process}, x_p^{process}) \leq 0$  are process inequality constraints (e.g., bounds on temperature, pressure, and other operational design decisions), while the constraint  $g_4(x_d^{process}, x_p^{process}) = 0$  pertains to the mass and energy conservation laws governing the process, and thermodynamic models such as the Gamma-Phi model [5-6] or equation of states that dictate the vapor-liquid equilibria (VLE) inside a separation unit, or nonlinear kinetic models integrated with VLE models for reactive-distillation processes [7-8]. Finally,  $g_5(x_p^{mat}, x_p^{process}) \leq 0$  and  $g_6(x_p^{mat}, x_p^{process}) = 0$  are linking constraints that describe the interaction between the material and process level performance through material property, thermodynamic property, and process models.

Several solution techniques and CAMPD approaches have been reported in the past, especially in systematically identifying organic solvents and solvent-based separation processes [9-12]. These approaches can be bottom-up, top-down, or simultaneous [13]. The bottom-up approach decomposes the CAMPD problem by first considering the molecular decisions, and sequentially reducing the feasible space by adding constraints, thereby deriving optimal product and process pairs [10-11]. The top-down approach also decomposes the CAMPD, and first optimizes the process performance by considering an ideal hypothetical molecule, and then optimizes over the molecular design space to match the molecular property targets [9,14]. The simultaneous approach often considers a relatively smaller design space for tractability and simultaneously solves CAMPD using modified outer approximation [15] or generalized benders decomposition algorithm [16]. CAMPD problems are multiscale in nature and often involve models with highly nonlinear, nonconvex terms, or many mixed variables and constraints. Despite significant progress in global optimization [17-18], large-scale discrete-continuous problems, as commonly occurred in CAMPD, can still be daunting, requiring extensive computational time to achieve feasible solutions [19].

Recent advances in quantum computing (QC) show promise to unlock new avenues for solving complex optimization problems [20-21]. Quantum optimization techniques are primarily used to solve Quadratic Unconstrained Binary Optimization (QUBO) problems [21].

Linear and quadratic integer programs can be converted to QUBO, with constraints reformulated as quadratic penalties to the objective function [21-22]. Once formulated, these QUBO models are mapped onto quantum devices, leveraging the principles of quantum annealing to solve these models as energy minimization problems [22]. However, CAMPD problems are characterized by their constraints and continuous decision variables, which do not have direct QUBO reformulations.

To bridge this gap, we have formulated a technique that represents continuous and integer variables by a set of binary variables through efficient encoding schemes. This allows us to reformulate a class of MINLP to QUBO and subsequently solve them using QC platforms, thereby facilitating the solution to certain CAMPD problems. We illustrate our technique for the simultaneous selection of ionic liquids (IL) and the configuration of a reactor separator process network. We convert the CAMPD model into QUBO and then solve them using D-Wave's quantum annealer, with results compared against deterministic solvers. Our results suggest that QC is able to obtain good feasible solutions in terms of process configuration and IL selection, thereby establishing the efficacy of QC for CAMPD.

## COMPUTER-AIDED MOLECULAR AND PROCESS DESIGN WITH IONIC LIQUIDS

Designing efficient separation systems using ILs as solvent is gaining increasing research attention. ILs possess many desirable properties (e.g., high thermal and chemical stability, low vapor pressure, etc.) that make them a suitable choice for solvent, especially for separating gas mixtures [5-6]. ILs are salts that exist in the liquid state at ambient conditions. These are largely constructed by a pair of cation core, anion, and organic functional groups that are attached to the cation core. For example, 1-butyl-3-methylimidazolium chloride, a common IL known as [BMIM][Cl], is formed by an imidazolium cation core, a chloride ion, and a butyl group attached to the cation core. Based on the type and number of groups and ions that are present in an IL, the corresponding material performance can vary. As such, ILs can be molecularly designed for the desired performance that makes them suitable for CAMPD application. However, the design space for ILs is very large, owing to the combinatorial possibilities in the assignment of their anion and cation components as well as side groups. In addition to the optimal choice of IL, process optimization needs to be performed which typically involves identifying the optimal process design variables. A process network synthesis with the consideration of IL selection involves mixed-integer variables across multiple scales, making the optimization highly non-trivial. To illustrate, consider a mixture separation task in an extractive distillation column



using ILs. For this case, the molecular structural constraints for ILs,  $\mathbf{g}_1$  and  $\mathbf{g}_2$  can take the following form [23]:

$$\begin{aligned} \sum_{j \in Ca} c_j &= 1, \sum_{j \in An} c_j = 1 \\ \sum_{j \in Ca \cup Sub} (2 - v_j) \cdot c_j - 2 &= 0 \\ LB &\leq \sum_{j \in Sub} c_j \leq UB \end{aligned} \quad (1)$$

where  $Ca$ ,  $An$ ,  $Sub$  refer to the set of groups or molecular descriptors for the cation core, anion, and side groups respectively.  $v_j$  is the number of bonds for group  $j$ ,  $LB$  and  $UB$  are the minimum and maximum number for each group.  $\mathbf{g}_1$  and  $\mathbf{g}_2$  can also include property models of the following form:

$$\begin{aligned} T_m(K) &= c_1 + \sum_{j \in Ca \cup An \cup Side} n_j G_j \leq T_{max} \\ \ln \mu(cp) &= c_2 + \sum_{j \in Ca \cup An \cup Side} n_j a_j \\ &+ \sum_{j \in Ca \cup An \cup Side} \frac{n_j b_j}{T} \leq \ln \mu_{max} \end{aligned} \quad (2)$$

$$(3)$$

where the set of properties can be a function of the groups, as well as of design variables such as temperature or pressure. Process constraints  $\mathbf{g}_3$  and  $\mathbf{g}_4$  typically enforce bounds on the flowrates, operating temperature, and pressure. These also include mass, energy, and momentum balances of the following form:

$$\sum_{i \in Inlet} x_i = \alpha_k \sum_{i \in Outlet} x_i \quad \forall k \quad (4)$$

where for a separation unit  $k$ ,  $\alpha_k$  is the separation efficiency that is dictated by the choice of ILs. The linking constraints  $\mathbf{g}_5$  and  $\mathbf{g}_6$  relate the process level performance with the material selection, and can take the form as follows:

$$\beta_i = \frac{1}{\gamma_i^\infty} \times \frac{MW_i}{MW_{IL}} \geq \beta_{min} \quad (5)$$

$$S_{ij} = \frac{\beta_j^\infty}{\beta_i^\infty} \geq S_{min} \quad (6)$$

The distribution coefficient ( $\beta_i$ ) and selectivity ( $S_{ij}$ ) can be obtained through a VLE model and are linked with the IL design through the molecular weight,  $MW_{IL}$ . Finally, the objective function can be a nonlinear function of the process streams to minimize the overall cost. Therefore, the overall CAMPD model can be posed as a mixed-integer nonlinear program (MINLP) involving nonlinearities and nonconvexities across multiple scales making them challenging to solve.

## SOLVING MIXED-INTEGER PROGRAMS USING QUANTUM COMPUTING

We will first assume a quadratic approximation of all nonlinear models, making the overall CAMPD a mixed-integer quadratically constrained quadratic program (MIQCQP). Then, we will demonstrate the reformulation of MIQCQP to QUBO. Finally, we will extend it to a class of MINLP by introducing nonlinear encoding.

Assume that the vectors of all material and process level variables ( $\mathbf{x}_d^{mat}$ ,  $\mathbf{x}_p^{mat}$ ,  $\mathbf{x}_d^{process}$ ,  $\mathbf{x}_p^{process}$ ) are aggregated in a vector  $\mathbf{x}$ . We can then express the CAMPD (under quadratic approximation, MIQCQP) in a compact form as follows:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{x}^T \mathbf{H} \mathbf{x} + \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{x}^T \mathbf{C}_k \mathbf{x} + \mathbf{a}_k^T \mathbf{x} \leq b_k, \quad k = 1, \dots, u \\ & \mathbf{x} \in \mathbb{R}^n \times \mathbb{Z}^p, \end{aligned} \quad (P2)$$

where  $\mathbf{H} \in \mathbb{R}^{(n+p) \times (n+p)}$  is a symmetric matrix of real constants,  $\mathbf{c} \in \mathbb{R}^{(n+p)}$ ,  $\mathbf{C}_k \in \mathbb{R}^{(n+p) \times (n+p)}$ ,  $\mathbf{a}_k \in \mathbb{R}^{(n+p)}$ , and  $b_k \in \mathbb{R}$ .

### Encoding

To solve the model (P2) using QC, it first needs to be reformulated to a QUBO. For that, we utilize several encoding schemes for an approximate representation of both the continuous and integer variables in terms of binary variables. To illustrate this, let  $x_m \in \mathbb{R}$ , with  $x_m^L \leq x_m \leq x_m^U$ . It can then be scaled to  $\tilde{x}_m \in \mathbb{R}$  with  $0 \leq \tilde{x}_m \leq 1$  as follows:

$$x_m = x_m^L + (x_m^U - x_m^L) \tilde{x}_m \quad (7)$$

Then  $\tilde{x}_m$  can be approximately represented by the binary variables using unary encoding as follows [24]:

$$\tilde{x}_m \approx \sum_{j=1}^J \sum_{i=1}^9 10^{-j} z_{i,j,m} + 10^{-J} z_m, \quad (8)$$

where  $z_{i,j,m} \in \{0,1\}$ , and  $J$  is the number of decimal places. To illustrate this, consider  $\tilde{x}_m = 0.342$ . For  $J = 3$ , we can write  $\tilde{x}_m = 10^{-1}(z_{1,1,m} + \dots + z_{9,1,m}) + 10^{-2}(z_{1,2,m} + \dots + z_{9,2,m}) + 10^{-3}(z_{1,3,m} + \dots + z_{9,3,m})$ . The number of binary variables to represent  $\tilde{x}_m$  can be reduced by using a binary encoding as follows:

$$\tilde{x}_m \approx \sum_{j=1}^J 10^{-j} (z_{1,j,m} + 2z_{2,j,m} + 3z_{3,j,m} + 3z_{4,j,m}) + 10^{-J} z_m \quad (9)$$

For an integer variable, that is,  $x_m \in \mathbb{Z}$  let  $I$  be the least positive integer such that  $2^I \geq |x_m|$ . Then, using binary encoding,  $|x_m| = \sum_{i=1}^I 2^i z_{i,m}$ , where  $z_{i,m} \in \{0,1\}$ .

### Reformulating Mixed-Integer Quadratic Program with Linear Constraints to QUBO

Let us first consider a mixed-integer quadratic

program with linear constraints (MIQP) by setting  $C_k = [0]^{(n+p) \times (n+p)}$ . Then (P2) can be written as:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{x}^T \mathbf{H} \mathbf{x} + \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A} \mathbf{x} \leq \mathbf{b}, \\ & \mathbf{x} \in \mathbb{R}^n \times \mathbb{Z}^p, \end{aligned} \quad (\text{P3})$$

where  $\mathbf{A} \in \mathbb{R}^{u \times (n+p)}$  and  $\mathbf{b} \in \mathbb{R}^u$ . At first, we convert all inequality constraints into equalities by adding slack variables ( $\mathbf{s}$ ) to each constraint. Then we add a quadratic penalty to the objective function by introducing penalty parameters ( $\lambda$ ), as follows:

$$\min \mathbf{x}^T \mathbf{H} \mathbf{x} + \mathbf{c}^T \mathbf{x} + \lambda^T (\mathbf{A} \mathbf{x} - \mathbf{b} + \mathbf{s})^T (\mathbf{A} \mathbf{x} - \mathbf{b} + \mathbf{s}) \quad (10)$$

where  $\lambda \in \mathbb{R}^{u \times u}$ , and  $\mathbf{s} \in \mathbb{R}^{u \times u}$ . We scale the continuous variables  $\{x_m\}_{m=1}^n$  and  $\{s_v\}_{v=1}^u$  to  $\{\tilde{x}_m\}_{m=1}^n$  and  $\{\tilde{s}_v\}_{v=1}^u$ , respectively with  $0 \leq \tilde{x}_m, \tilde{s}_v \leq 1$ . Note that,  $s_v^L = 0, \forall v \in u$ , and  $s_v^U$  can be calculated as follows:

$$s_v^U = \begin{cases} b_v - x_m^L A_{vm}, & \text{if } A_{vm} \geq 0 \\ b_v - x_m^U A_{vm}, & \text{otherwise} \end{cases} \quad (11)$$

After scaling, Equation (10) can be expressed in terms of scaled variables as follows:

$$\min \tilde{\mathbf{x}}^T \tilde{\mathbf{H}} \tilde{\mathbf{x}} + \tilde{\mathbf{c}}^T \tilde{\mathbf{x}}, \quad (12)$$

where  $\tilde{\mathbf{H}} \in \mathbb{R}^{(n+p+u) \times (n+p+u)}$ ,  $\tilde{\mathbf{c}} \in \mathbb{R}^{n+p+u}$ , and  $\tilde{\mathbf{x}} \in \mathbb{R}^{n+p+u}$  with  $\tilde{\mathbf{x}} = [[\tilde{x}_1, \dots, \tilde{x}_n] \quad [\tilde{x}_{n+1}, \dots, \tilde{x}_{n+p}] \quad [\tilde{s}_1, \dots, \tilde{s}_u]]$ , where the first  $n$  decision variables are continuous, the next  $p$  variables are binary, and the remaining  $u$  variables are slack variables. We then represent all these scaled variables using unary or binary encoding, and finally reformulate as QUBO as follows:

$$\begin{aligned} \min_{\mathbf{z}} \quad & \mathbf{z}^T \mathbf{Q} \mathbf{z} \\ \text{s.t.} \quad & \mathbf{z} = \{0,1\}^\beta, \end{aligned} \quad (\text{P4})$$

where  $\mathbf{Q} \in \mathbb{R}^{\beta \times \beta}$  and  $\beta = (n+u) \cdot (9J+1) + p$  and  $(n+u) \cdot (4J+1) + p$  for unary and binary encoding, respectively. Thus, the original  $n+p$  dimensional MIQP (P3) is reformulated into a  $\beta$  dimensional QUBO.

## Reformulating Mixed-Integer Quadratically Constrained Quadratic Program to QUBO

Reformulating a mixed-integer quadratically constrained quadratic program (MIQCQP) to QUBO involves two steps. In the first step, we reformulate MIQCQP to an MIQP. In the second step, we follow the steps described in the previous section to convert the reformulated MIQP to QUBO. For Step 1 we introduce additional binary variables to substitute binary quadratic terms that appear in the constraints and for each of them, add an equivalent penalty function,  $f$  to the objective function as follows:

$$f = P(3w - 2z_1w - 2z_2w + z_1z_2) \quad (13)$$

where,  $w = z_1z_2$ , and  $w, z_1, z_2 \in \{0,1\}$ .  $z_1z_2$  is a binary

quadratic term that appears in the constraint, and  $P \in \mathbb{R}_+$  is a penalty parameter. As shown in Table 1, for a feasible configuration of  $z_1, z_2$ , and  $w$ ,  $f = 0$  whereas infeasible configuration implies  $f > 0$ .

**Table 1:** Penalty function for MIQCQP

$z_1$	$z_2$	$w$	$f$
1	1	1	0
1	1	0	$> 0$
1	0	0	0
1	0	1	$> 0$
0	1	0	0
0	1	1	$> 0$
0	0	0	0
0	0	1	$> 0$

## Illustrative Example

$$\begin{aligned} \min_{x_1, y_1, y_2} \quad & -2x_1y_1 + y_2^2 \\ \text{s.t.} \quad & x_1 + y_1y_2 \leq 1, \\ & 0 \leq x_1 \leq 1; y_1, y_2 \in \{0,1\} \end{aligned} \quad (\text{P5})$$

We substitute  $y_1y_2$  in the constraint by defining  $w = y_1y_2$ , and add the penalty to the objective as follows:

$$\begin{aligned} \min_{x_1, y_1, y_2, w} \quad & -2x_1y_1 + y_2^2 + P(3w - 2y_1w - 2y_2w + y_1y_2) \\ \text{s.t.} \quad & x_1 + w \leq 1, \\ & 0 \leq x_1 \leq 1; w, y_1, y_2 \in \{0,1\} \end{aligned} \quad (\text{P6})$$

Here, Problem P6 is an MIQP with linear constraints. We then encode the continuous variables using binary encoding:

$$\begin{aligned} \min \quad & -2\{10^{-1}(z_{111} + 2z_{211} + 3z_{311} + 3z_{411}) + z_1\}y_1 + y_2 \\ & + P(3w - 2y_1w - 2y_2w + y_1y_2) \\ \text{s.t.} \quad & 10^{-1}(z_{111} + 2z_{211} + 3z_{311} + 3z_{411}) + z_1 + w \leq 1, \\ & z_{111}, z_{211}, z_{311}, z_{411}, z_1, w, y_1, y_2 \in \{0,1\} \end{aligned} \quad (\text{P7})$$

Here, Problem (P7) is a binary quadratic program with linear constraints whose equivalent QUBO model is constructed with the  $\mathbf{Q}$  matrix given in Table 2. For  $J = 1$ ,  $\mathbf{Q} \in \mathbb{R}^{13 \times 13}$  with the last 5 binary variables used to encode the slack variable that corresponds to the constraint. We set  $P = 10$ , and solve the resulting QUBO in D-Wave's quantum annealer [25], and obtain the optimal solution:  $x_1^* = 1, y_1^* = 1, y_2^* = 0$ .

## Encoding Nonlinear Terms

Assume that the CAMPD model consists of terms such as  $\tilde{x}_m^a$ , where  $a \in \mathbb{R}$ . We can then encode this nonlinear term using the same binary variables corresponding to  $\tilde{x}_m$  as follows:

$$\tilde{x}_m^a \approx \sum_{j=1}^J \sum_{i=1}^9 P_{i,j,m} z_{i,j,m} + Q_m z_m, \quad (14)$$

$$z_{i+1,j,m} \leq z_{i,j,m} \quad \forall i = 1, 2, \dots, |I|, j = 1, 2, \dots, |J| \quad (15)$$

Note that the constraints corresponding to Equation (15) have an exact penalty function that can be added to the objective as follows:  $\sum_{j=1}^J \sum_{i=1}^8 P(z_{i+1,j,m} - z_{i,j,m} \cdot z_{i,j,m})$ . Also note that, the parameters  $P_{i,j,m}$  and  $Q_m$  can be precomputed for all instances of  $\tilde{x}_m^a$  since  $a$  is known. Specifically, consider  $J = 1$ . This implies  $\tilde{x}_m = \sum_{i=1}^9 0.1 z_{i,1,m} + 0.1 z_m$ . Then,  $\tilde{x}_m^a = \sum_{i=1}^9 P_{i,1,m} z_{i,1,m} + Q_m z_m$ , where  $P_{i,1,m} = \gamma_{i,1,m} - \gamma_{i-1,1,m} \quad \forall i \in I$ .  $\gamma_{0,1,m} = 0$ , and  $Q_m = 1 - \gamma_{9,1,m}$ , where  $\gamma_{i,1,m} = (0.1i)^a$ . To illustrate, assume that  $\tilde{x}_m$  takes the value 0.3. Then Equation (14) and (15) together ensures that the only valid solution is  $z_{1,1,m} = z_{2,1,m} = z_{3,1,m} = 1$  and the others are 0. This then enforces  $\tilde{x}_m^a = \gamma_{3,1,m} = (0.1 \cdot 3)^a = (0.3)^a$ , as desired.

We can further encode other nonlinear terms that can be approximately expressed in terms of  $\tilde{x}_m^a$ , e.g.,  $\exp(x)$ ,  $\log(x)$ ,  $\sin(x)$ ,  $\cos(x)$ ,  $\log(1+x)$  and  $\log(x)$  using McLaurin expansion. Finally, we can also encode general nonlinear product terms such as  $\prod_m \tilde{x}_m^{a_m}$  using the exp-log-sum technique by encoding  $\exp(\sum_m a_m \log \tilde{x}_m)$ , thereby allowing to reformulate MINLP with signomial terms to QUBO.

## INTEGRATED IL SELECTION AND CONFIGURATION OF REACTOR SEPARATOR NETWORK

We formulate and solve a reactor-separator network synthesis problem with integrated IL selection as shown in Figure 1. The separator performance depends on the choice of ILs, thus making the overall problem a CAMPD. In this illustrative case study, we assume that the selection of ILs is made by choosing one anion and one cation from a pool of anions and cations. Formally, consider  $Cat = \{c_1, c_2, \dots, c_C\}$  and  $An = \{a_1, a_2, \dots, a_A\}$ . Then all possible feasible ILs are constructed by  $IL = \{c \times a \mid c \in Cat, a \in An\}$ . With this construction, we formulate the following optimization model that integrates the selection of ILs with the synthesis of the reactor-separator network.

$$\begin{aligned} \min \quad & \sum_{k \in K} c_k^f y_k + \sum_{i \in I_{K_r}^{in}} c_i^f x_i^{0.6} \\ & + \sum_{k \in K_s} c_k^l \left( \sum_{i \in I_{K_s}^{in}} x_i \right)^2 \\ & + \sum_{k \in K_s} c_k^e \left( \sum_{i \in I_{K_s}^{in}} x_i - \sum_{i \in I_{K_s}^{out}} x_i \right) \\ \text{s.t.} \quad & f_k^l y_k \leq \sum_{i \in I_{K_s}^{in}} x_i \leq f_k^u y_k, \quad \forall k \in K \end{aligned} \quad (P8)$$

$$\sum_{i \in I_{K_r}^{out}} x_i = \alpha_k \sum_{i \in I_{K_r}^{in}} x_i, \quad \forall k \in K_r$$

$$\sum_{c \in Cat} z_c = 1$$

$$\sum_{a \in An} z_a = 1$$

$$x_i \geq \beta_{k,c,a} \sum_{i \in I_{K_s}^{in}} x_i - M(2 - z_c - z_a), \\ \forall i \in I_{K_s}^{out}, k \in K_s, c \in Cat, a \in An$$

$$x_i \leq \beta_{k,c,a} \sum_{i \in I_{K_s}^{in}} x_i + M(2 - z_c - z_a), \\ \forall i \in I_{K_s}^{out}, k \in K_s, c \in Cat, a \in An$$

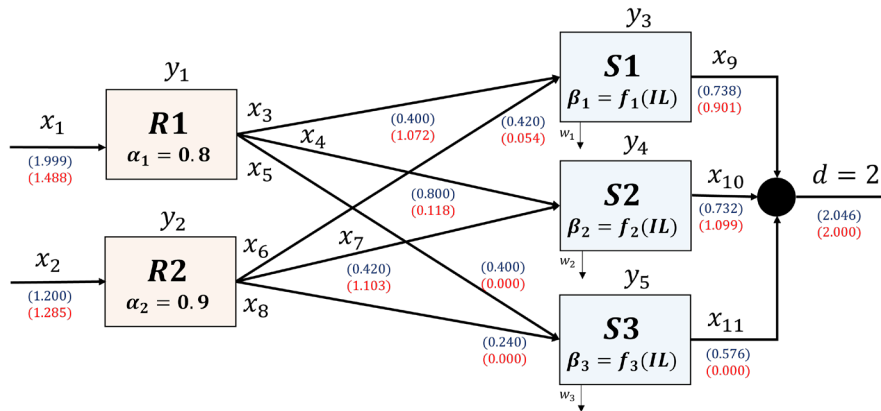
$$\sum_{i \in I_{K_s}^{out}} x_i \geq d,$$

where  $K$  is the set of all process units,  $K_r, K_s \subset K$  are the sets of all reactors and separators, respectively.  $\alpha_k$  is the stoichiometric reactor conversion factor for a reactor  $k$ ,  $\beta_{k,c,a}$  is the separation factor for a separator  $k$  with an IL constructed by cation  $c$  and anion  $a$ .  $I_{K_s}^{in}$  and  $I_{K_s}^{out}$  are the set of all inlet and outlet streams, respectively, and  $d$  is the demand. The decision variables involve all the flowrate streams  $x_i$  and the selection of cation and anion,  $z_c$  and  $z_a$ , respectively. The objective is to minimize the overall cost that includes the fixed cost, investment cost, operating cost, and cost associated with emissions or waste.  $c_k^f$  are coefficients for fixed cost;  $c_i^f$  and  $c_k^l$  are coefficients for operating cost for reactors and separators, respectively; and  $c_k^e$  are coefficients for the cost associated with waste or emission. The first constraint is related to the unit selection and flowrate bounds. The second constraint is for the overall mass balance around the reactors. The third and fourth constraints ensure that a feasible IL is constructed by choosing only one cation and one anion. The fifth and sixth constraints relate the output flows from the separators to the IL selection. In this way, the effect of the material selection on the process performance is captured. The final constraint ensures that the sum of all output flows must satisfy the demand. The overall problem involves both binary and continuous variables. We solve the optimization problem using a quantum annealer and compare the solution against a deterministic solver that uses classical computing. All the quantum computation is performed on D-Wave's LeapHybrid solver, and the results are compared to that of a deterministic solver (BARON v44.2) run on a Dell Windows system with Intel(R) Core (TM) i7-8750 2.20 GHz CPU and 16 GB RAM.

We select an initial pool of two cations and two anions. The goal is to select the best IL from the cation and anion pool, as well as to select the best reactor-separator network from two reactors and three separators. For that,

**Table 2:** Q matrix for illustrative example.

	$z_{111}$	$z_{211}$	$z_{311}$	$z_{411}$	$z_1$	$y_1$	$y_2$	$w$	$z_{S_{111}}$	$z_{S_{211}}$	$z_{S_{311}}$	$z_{S_{411}}$	$z_S$
$z_{111}$	0	0	0	0	0	$-10^{-1}$	0	0	0	0	0	0	0
$z_{211}$	0	0	0	0	0	$-2 \cdot 10^{-1}$	0	0	0	0	0	0	0
$z_{311}$	0	0	0	0	0	$-3 \cdot 10^{-1}$	0	0	0	0	0	0	0
$z_{411}$	0	0	0	0	0	$-3 \cdot 10^{-1}$	0	0	0	0	0	0	0
$z_1$	0	0	0	0	0	$-10^{-1}$	0	0	0	0	0	0	0
$y_1$	$-10^{-1}$	$-2 \cdot 10^{-1}$	$-3 \cdot 10^{-1}$	$-3 \cdot 10^{-1}$	$-10^{-1}$	0	$P/2$	$-P$	0	0	0	0	0
$y_2$	0	0	0	0	0	$P/2$	0	$-P$	0	0	0	0	0
$w$	0	0	0	0	0	$-P$	$-P$	0	0	0	0	0	0
$z_{S_{111}}$	0	0	0	0	0	0	0	0	$10^{-2}$	$2 \cdot 10^{-2}$	$3 \cdot 10^{-2}$	$3 \cdot 10^{-2}$	$10^{-2}$
$z_{S_{211}}$	0	0	0	0	0	0	0	0	$2 \cdot 10^{-2}$	$4 \cdot 10^{-2}$	$6 \cdot 10^{-2}$	$6 \cdot 10^{-2}$	$2 \cdot 10^{-2}$
$z_{S_{311}}$	0	0	0	0	0	0	0	0	$3 \cdot 10^{-2}$	$6 \cdot 10^{-2}$	$9 \cdot 10^{-2}$	$9 \cdot 10^{-2}$	$3 \cdot 10^{-2}$
$z_{S_{411}}$	0	0	0	0	0	0	0	0	$3 \cdot 10^{-2}$	$6 \cdot 10^{-2}$	$9 \cdot 10^{-2}$	$9 \cdot 10^{-2}$	$3 \cdot 10^{-2}$
$z_S$	0	0	0	0	0	0	0	0	$10^{-2}$	$2 \cdot 10^{-2}$	$3 \cdot 10^{-2}$	$3 \cdot 10^{-2}$	$10^{-2}$



**Figure 1:** Schematic of an integrated IL selection and reactor-separator network

the parameters used are as follows:  $\beta_{1,1,1} = 0.7, \beta_{1,1,2} = 0.9, \beta_{1,2,1} = 0.8, \beta_{1,2,2} = 0.9, \beta_{2,1,1} = 0.8, \beta_{2,1,2} = 0.7, \beta_{2,2,1} = 0.9, \beta_{2,2,2} = 0.6, \beta_{3,1,1} = 0.9, \beta_{3,1,2} = 0.6, \beta_{3,2,1} = 0.7, \beta_{3,2,2} = 0.9$ . The conversion factors for the two reactors are:  $\alpha_1 = 0.8, \alpha_2 = 0.9$ . Cost parameters in the objective function are as follows:  $c_1^f = 10, c_2^f = 15, c_3^f = 8, c_4^f = 9, c_5^f = 10$ . Operating cost parameters for the reactors are  $[0.5, 0.7]$ . Operating cost parameters for the separators are  $[0.2, 0.7, 0.5]$ . Emission cost parameters for the separators are  $[1.2, 1.5, 0.9]$ . The upper bound on the flowrate for all streams (see Figure 1) is 2. This results in a MINLP with 39 constraints. Using these model parameters, we first

solve the optimization model in GAMS using BARON v44.2. The optimal objective and the network selections are as follows:  $f_{obj} = 46.009, y_1 = 1, y_2 = 1, y_3 = 1, y_4 = 1, y_5 = 0, z_1^{cation} = 0, z_2^{cation} = 1, z_1^{anion} = 1, z_2^{anion} = 0$ . The optimal IL is constructed by combining the 2<sup>nd</sup> cation with the 1<sup>st</sup> anion. Also, both reactors are selected, and only the first and the second separators are selected.

We then solve the same optimization model in D-Wave's LeapHybrid solver. For that, we first eliminate the constraints for anion and cation selection, that is,  $\sum_{c \in Cat} z_c = 1$  and  $\sum_{a \in An} z_a = 1$ , by observing that, for  $|Cat| = |An| = 2$ , an exact penalty function  $P(1 - z_1 - z_2 +$

$2z_1z_2$ ) can be achieved, where  $P$  is a penalty parameter whose value is set to 2. After removing the two constraints and adding the equivalent penalties to the objective, we now have a reduced MINLP with 37 constraints. We then reformulate it to QUBO by adding a quadratic penalty to the objective function and representing continuous variables by unary encoding ( $J = 5$ ). We obtain a feasible solution as follows:  $f_{obj} = 58.369, y_1 = 1, y_2 = 1, y_3 = 1, y_4 = 1, y_5 = 1, z_1^{cation} = 0, z_2^{cation} = 1, z_1^{anion} = 0, z_2^{anion} = 1$ . From the results, we can observe that all the process units are selected resulting in a higher value for the objective. However, it confirms that it is possible to obtain good feasible solutions using QC.

## CONCLUSIONS

We have developed a technique for solving mixed-integer programs using Quantum Computing (QC) that allows us to solve Computer-aided Molecular and Process Design (CAMPD) problems. Our methodology utilizes efficient encoding schemes and transforms a class of Mixed-Integer Nonlinear Programs into Quadratic Unconstrained Binary Optimization (QUBO) models, thus making them possible to solve using QC platforms. We demonstrate the framework for the selection of optimal ionic liquid and configuration of a reactor-separator process with promising results suggesting QC's potential to solve CAMPD problems. However, scalability and applicability to large-scale industrial problems remain a challenge for the quantum optimization-based approach. Also, the solution quality and feasibility of the QC solution heavily depend upon the penalty parameter. Ongoing works include the decomposition of CAMPD models as well as the QUBO matrix by exploiting their structure, thus enabling the solution to larger CAMPD problems.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge support from the NSF CAREER award CBET-1943479. Part of the research was conducted with the computing resources provided by Texas A&M High Performance Research Computing.

## REFERENCES

- Constantinou L, Gani R. New group contribution method for estimating properties of pure compounds. *AIChE J* 40(10): 1697-1710 (1994)
- Marrero J, Gani R. Group-contribution based estimation of pure component properties. *Fluid phase equilib* 183: 183-208 (2001)
- Grimstad B, Andersson H. ReLU networks as surrogate models in mixed-integer linear programs. *Comput Chem Eng* 131: 106580 (2019)
- Churi N, Achenie LE. Novel mathematical programming model for computer aided molecular design. *Ind Eng Chem Res* 35(10): 3788-3794 (1996)
- Shiflett MB, Yokozeki A. Solubility and diffusivity of hydrofluorocarbons in room-temperature ionic liquids. *AIChE J* 52(3), 1205-1219 (2006)
- Monjur MS, Iftakher A, Hasan MMF. Separation process synthesis for high-gwp refrigerant mixtures: Extractive distillation using ionic liquids. *Ind Eng Chem Res* 61(12): 4390-4406 (2022)
- Iftakher A, Mansouri SS, Nahid A, Tula AK, Choudhury MS, Lee JH, Gani R. Integrated design and control of reactive distillation processes using the driving force approach. *AIChE J* 67(6): e17227 (2021)
- Taylor R, Krishna R. Modelling reactive distillation. *Chem Eng Science*, 55(22): pp.5183-5229 2000.
- Bardow A, Steur K, Gross J. Continuous-Molecular Targeting for Integrated Solvent and Process Design. *Ind Eng Chem Res* 49(6):2834-2840 (2010)
- Burger J, Papaioannou V, Gopinath S, Jackson G, Galindo A, Adjiman CS. A hierarchical method to integrated solvent and process design of physical CO<sub>2</sub> absorption using the SAFT- $\gamma$  Mie approach. *AIChE J* 61(10): 3249-3269 (2015)
- Hostrup M, Harper PM, Gani R. Design of environmentally benign processes: integration of solvent design and separation process synthesis. *Comput & Chem Eng* 23 (10): 1395-1414 (1999)
- Liu Q, Zhang L, Liu L, Du J, Tula AK, Eden M, Gani R. OptCAMD: An optimization-based framework and tool for molecular and mixture product design, *Comput & Chem Eng* 124: 285-301 (2018)
- Iftakher A, Monjur MS, Hasan MMF. An Overview of Computer-aided Molecular and Process Design. *Chem Ing Tech* 95:315-333 (2023)
- Eljack FT, Eden MR, Kazantzi V, Qin X, El-Halwagi MM. Simultaneous process and molecular design—A property based approach. *AIChE J* 53(5): 1232-1239 (2007)
- Gopinath S, Jackson G, Galindo A, Adjiman CS. Outer approximation algorithm with physical domain reduction for computer-aided molecular and separation process design. *AIChE J*, 62(9), 3484-3504 (2016)
- Buxton A, Livingston AG, Pistikopoulos EN. Optimal design of solvent blends for environmental impact minimization. *AIChE J*, 45(4): 817-843 (1999)
- Floudas CA. Deterministic global optimization: theory, methods and applications (Vol. 37). Springer Science & Business Media (2013)
- Misener R, Floudas CA. Antigone: algorithms for continuous/integer global optimization of nonlinear equations. *J Glob Optim* 59(2-3):503-526 (2014)



19. Ajagekar A, Humble T, You F. Quantum computing based hybrid solution strategies for large-scale discrete-continuous optimization problems. *Comput Chem Eng* 132: 106630 (2020)
20. Grant EK, Humble TS. Adiabatic Quantum Computing and Quantum Annealing. Oxford University Press (2020)
21. Glover F, Kochenberger G, Hennig R, Du Y. Quantum bridge analytics i: a tutorial on formulating and using qubo models. *Ann Oper Res* 314:141-183 (2022)
22. Anthony M, Boros E, Crama Y, Gruber A. Quadratic reformulations of nonlinear binary optimization problems. *Math Programming* 162(1): 115–144 (2017)
23. Chao H, Song Z, Cheng H, Chen L, Qi Z. Computer-aided design and process evaluation of ionic liquids for n-hexane-methylcyclopentane extractive distillation. *Sep. Purif. Technol* 196: 157-165 (2018)
24. Iftakher A, Kazi MK, Hasan MMF. Mixed-integer Quadratic Optimization using Quantum Computing for Process Applications. *In Proceeding of the Foundations of Computer Aided Process Operations / Chemical Process Control*, 1-6 (2023)
25. Ushijima-Mwesigwa H, Negre CF, Mniszewski SM. Graph partitioning using quantum annealing on the d-wave system. *In Proceedings of the Second International Workshop on Post Moores Era Supercomputing*, 22-29 (2017)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Learning Hybrid Extraction and Distillation using Phenomena-based String Representation

Jianping Li\*

<sup>a</sup> Energy Systems and Infrastructure Analysis Division, Argonne National Laboratory, Lemont, IL 60439 USA

\* Corresponding Author: [jianping.li@anl.gov](mailto:jianping.li@anl.gov).

---

## ABSTRACT

We present a string representation for hybrid extraction and distillation using symbols representing phenomena building blocks. Unlike the conventional equipment-based string representation, the proposed representation captures the design details of liquid-liquid extraction and distillation. We generate a set of samples through the procedure of input parameter sampling and superstructure optimization that minimizes separation cost. We convert these generated samples into a set of string representations based on pre-defined rules. We use these string representations as descriptors and connect them with conditional variational encoder. The trained conditional variational encoder shows good prediction accuracy. We further use the trained conditional variational encoder to screen designs of hybrid extraction and distillation with desired cost investment.

---

**Keywords:** Extraction, Distillation, Machine Learning, Process Intensification

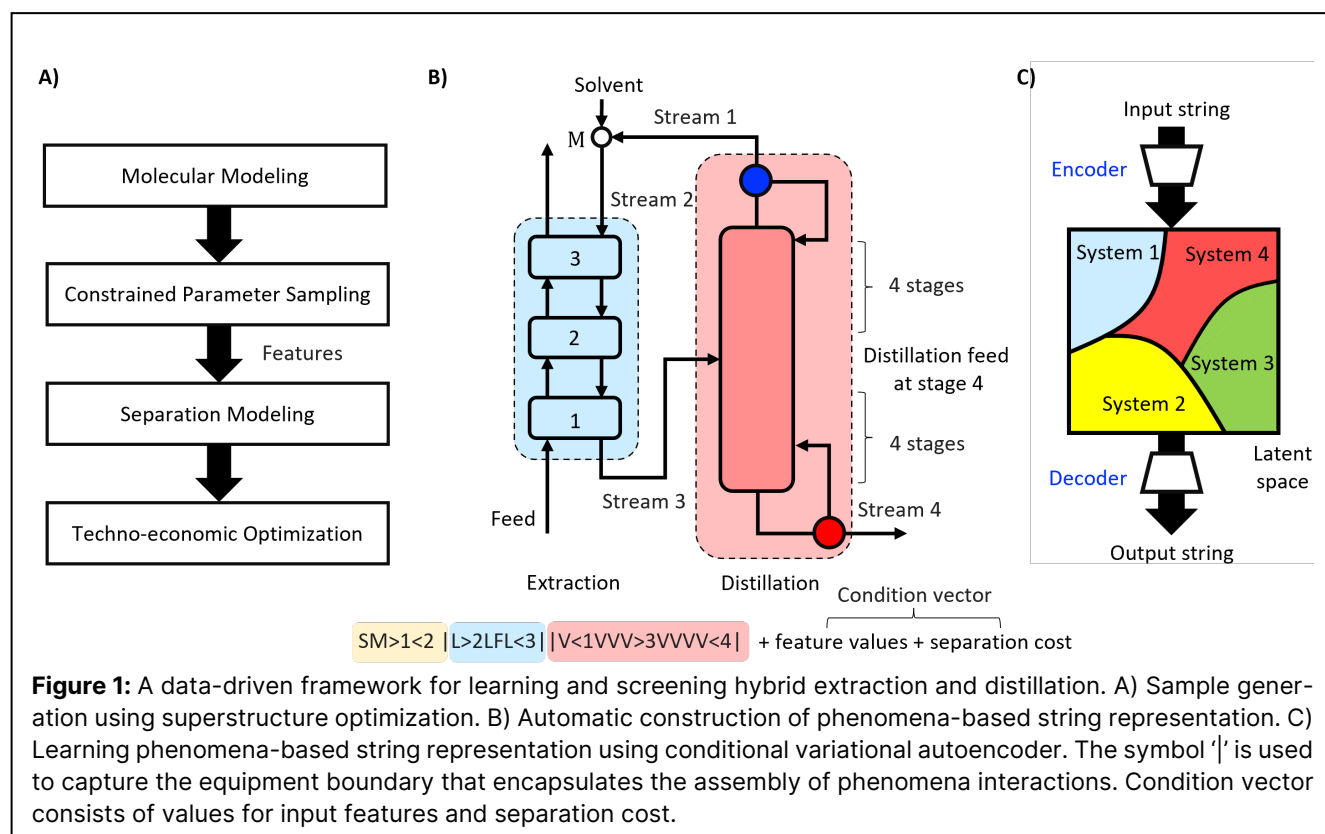
## INTRODUCTION

Separation design is critical for plastic upcycling, biomass valorization, carbon capture and utilization [1-3]. Liquid-liquid extraction and distillation are two commonly used separation methods. Hybrid extraction and distillation further reduces material and energy consumption by leveraging driving forces of solubility difference and volatility difference, which leads to process intensification [4-6]. Global optimization of hybrid extraction and distillation has been reported [7]. To reduce the computational efforts, machine learning shows promise to identify optimal hybrid extraction and distillation design with desired separation cost using machine-readable representations [8]. These motivate the research into constructing machine-readable data representation for learning flowsheets with design details and screening flowsheets that meet design objectives.

Construction of machine-readable representation is an active area in the field of machine learning. Development of such data representation with physical interpretation requires domain knowledge. In the field of molecular modeling, starting from SMILES (Simplified Molecular Input Line Entry System) [9], diverse and inclusive string representations have been reported. These include INCHI strings for standardized searching across databases

and the internet, DEEPSMILES for automated inverse design using deep generative models, CurlySMILES for macromolecules, and BigSMILES for polymers among others [9-13]. These string representations for molecules are constructed based on the molecular structure information such as atoms and their connections (SMILES) [9], information on tautomer, isotope, stereochemistry, and electronic charge (INCHI strings) [10], information on molecular details and extra-molecular features such as non-covalent interactions, surface attachment, solutions, and crystal structures (CurlySMILES) [12], and stochastic object that represents a molecule fragment that is stochastic in the molecular structure (BigSMILES) [13]. In the field of process modeling, d'Anterrosches and Gani proposed SFILES (simplified flowsheet-input line-entry system) using equipment-based flowsheet building blocks [14], which enables the flowsheet similarity comparison in the work of Zhang et al [15]. SFLIES has been recently extended to SFILES2.0 that describes top and bottom products, and the control structure [16] and to eSFILES that integrates process engineering domain knowledge and machine learning [17]. Despite the success in constructing string representation based on equipment, limited representations exist that capture design details and enable the identification of novel design opportunities.

Once these string representations are constructed



as descriptors and the corresponding training samples are generated, these samples are often related with design objectives through machine learning models. In the field of molecular modeling, commonly used machine learning models include graph neural network (GNN) [18], variational autoencoder (VAE) [19], and conditional variational autoencoder (CVAE) [20]. GNNs take molecular graphs as inputs and perform convolutions based on the graph topology by aggregating the features of a node and its connected neighbors. The node features are embedded into a fixed-dimension space where similar nodes are close to each other. VAEs generalize autoencoders, add stochasticity to the encoder which is combined with a penalty term, and encourage all areas of the latent space to correspond to a valid decoding. To identify the optimal molecule structures, VAEs are often jointly trained with a property prediction model [19]. CVAEs distinguish from the VAEs by imposing certain conditions in the encoding and decoding processes and do not require further optimization during the search of new molecular design, which is required for the framework with VAE and an additional neural network jointly trained [21].

In this work, with a novel phenomena-based string representation, we develop a data-driven framework for learning and screening the design details of hybrid extraction and distillation. Firstly, to generate data for automatic construction of phenomena-based string representation, we use superstructure optimization by

modeling hybrid extraction and distillation based on state-by-stage model for extraction column and shortcut model for distillation column. Then, these models are solved to global optimality with sampled input feature vectors to obtain minimum separation cost. These input feature vectors are sampled from the thermodynamic space approximated using Conductor-like Screening Model for Real Solvents (COSMO-RS) and molecular dynamics simulation [3]. We define rules for constructing the string representation with information on the phenomena nodes, equipment boundary, and equipment connections, which is extracted from the solution of superstructure optimization. These phenomena-based string representations are learned using conditional variational autoencoder, where new designs that meet desired cost investment are sampled from the latent space.

## PROBLEM STATEMENT

We consider the hybrid extraction and distillation systems with the following processing conditions: (1) the feed stream only contains water and one product while other components are neglected, and (2) the liquid-liquid extraction and distillation enrich product into a product stream from a ternary system of product (P), water (W), and solvent (S).

Given are (1) specification on the feed stream (i.e., product mole fraction in feed), (2) specifications on the product stream (i.e., total product flowrate, product

recovery from the feed, and product purity), (3) physical properties (liquid-liquid equilibrium constants, relative volatilities), (4) cost coefficients used for determining capital and operating costs of liquid-liquid extraction and distillation (solvent price is an operating cost coefficient for extraction), and (5) the maximum number of liquid-liquid extraction stages. The assumptions in this work include that: (1) we consider ternary systems where the product mole fraction is low such that liquid-liquid equilibrium constants are composition-independent and are identical at all stages of liquid-liquid extraction column, and (2) we consider distillation of near-ideal mixtures, allowing us to assume constant relative volatility.

The goal of this work is to construct a string representation with information on physicochemical phenomena, stream connections, and equipment boundary. With the proposed string representation, we use conditional variational autoencoder to map the discrete separation design space to a continuous latent space. This enables fast sampling within the latent space to screen designs of hybrid extraction and distillation that meets desired cost investment.

## COMPUTATIONAL FRAMEWORK

In this Section, we describe the computational framework for learning and screening hybrid extraction and distillation (Figure 1). This framework consists of (1) sample generation using superstructure optimization, (2) automatic construction of phenomena-based string representation, and (3) learning phenomena-based string representation using conditional variational autoencoder. With the trained conditional variational encoder, we sample within the latent space to screen designs of hybrid extraction and distillation.

### Sample Generation using Superstructure Optimization

To generate the flowsheet samples with design details, we adapted the workflow of integrated molecular modeling and process modeling from Li et al. to the reduced framework (shown in Figure 1A) [22]. Here, we used superstructure model as the first-principle model to generate flowsheet samples with detailed design information. Specifically, we used this framework to (1) predict a physically reasonable space of thermodynamic properties, (2) generate feature vectors by sampling within the space of thermodynamic properties and ranges of feature values, and (3) separation modeling and optimization to obtain the costs and design details of hybrid extraction and distillation for the sampled feature vectors.

The representation, modeling, and optimization of hybrid extraction and distillation were achieved using classic superstructure optimization [23]. The

representation of hybrid extraction and distillation involves stage-by-stage representation of liquid-liquid extraction column, equipment-based representation of distillation column, which were modeled using a shortcut model (detailed design information is extracted following the procedure described in the next subsection), and their stream connections [22]. This representation was modeled as a mixed-integer nonlinear programming (MINLP) problem. Discrete decisions are made on the activation of liquid-liquid extraction stages and the use of the second potential distillation column. We used both stage-by-stage model (for liquid-liquid extraction column) and shortcut Underwood-Fenske model (for distillation column) to balance the accuracy of separation physics and tractability of the proposed mathematical model [22]. The specifications for all hybrid extraction and distillation systems are (1) total product flowrate of 1 kmol/s, (2) 95% product purity, and (3) 90% product recovery. Each MINLP model were solved using BARON for each input feature vector with 1% optimality gap and within 5 CPU hours, using parallel computing resources [24].

In this work, we considered the input features as liquid-liquid equilibrium constants of solvent ( $\kappa_S$ ) and product ( $\kappa_P$ ), relative volatility between water and product ( $\alpha_{W,P}$ ), relative volatility between solvent and product ( $\alpha_{S,P}$ ), and solvent price ( $\pi_S$ ). Here, W, S, P are water, solvent, and product respectively. The bounds for these features are [1, 1000], [1, 50], [1.1, 2], [50, 6000] (\$ per kmol solvent) respectively to focus on a property region with difficult distillation. Liquid-liquid equilibrium constant of water was fixed at 0.01. We considered the product mole fraction as 0.1. Through this adapted framework, we prepared and generated 10,000 samples with values of input features, separation cost, and their corresponding *.st* files containing design details. 9816 samples remain after removing infeasible samples.

### Automatic Construction of Phenomena-based String Representation

Phenomena building blocks enable the identification of novel design opportunities and capture detailed design information compared with equipment-based process design and synthesis [25-27]. We leveraged phenomena building blocks to construct a string representation for learning and screening flowsheets.

We followed the rule of constructing SMILES and considered the criterion of leveraging information that uniquely and discretely represents the solutions of hybrid extraction and distillation models, which are obtained in the previous subsection. Following this criterion, we focus on the identity of physicochemical phenomena nodes and their connections. Specifically, we considered physicochemical phenomena as mixing (M), liquid-liquid phase equilibrium (L), and vapor-liquid phase equilibrium

(V) and constructed separation flowsheet strings with nodes as phenomena for each sample (Figure 1B). We did not consider other common factors for all samples such as system specifications (total product flowrate, product purity, and product recovery) and generic postulated superstructure representation with all equipment options (technology availability) in the proposed string representations since all optimal solutions are obtained based on these specifications and from the same search space represented by the superstructure. Moreover, since we solved each MINLP model to global optimality, we did not consider the initial conditions of solutions in the proposed string representation.

Note that the distillation column is modeled using equipment-based shortcut model and hence only equipment-wide design information can be obtained, which does not include the information on the position of feed stage. To obtain this information, based on the optimal solution for the MINLP model, we used Kirkbride equation [28] to determine the ratio of the number of stages above and below the feed stage. Based on this ratio and the total number of stage number in the distillation column, we located the position of feed stage and part of the python script is shown as follows:

```
def kirkbride(B, D, xFHK, xFLK, xBLK, xDHK):
    lnDB = 0.206*np.log(B/D*xFHK/xFLK*(xBLK/
    xDHK)**2)
    return np.exp(lnDB)
```

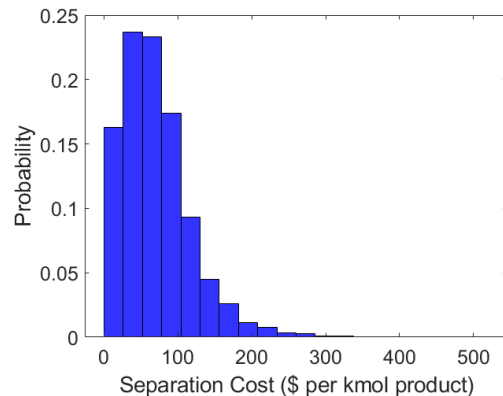
Here, the input B, D, xFHK, xFLK, xBLK, and xDHK for the above function are values of bottom flow rate, flow rate of distillate stream, composition of heavy key component in the feed stream, composition of light key component in the feed stream, composition of light key component in the bottom stream, and composition of heavy key component in the distillate stream of the distillation column. The output  $\text{np.exp}(\ln\text{DB})$  is the ratio of the number of stages above and below the feed stage.

The other key components of the proposed string representation are (1) integer number to label different connecting streams, (2) symbol '>' to represent inlet streams for equipment, (3) symbol '<' to represent outlet streams for equipment, and (4) symbol '|' to represent the first or the last stage of equipment to represent equipment boundary, which encapsulates all involved physico-chemical phenomena within certain equipment. Symbols '>' and '<' are always placed next to the right of phenomena nodes where stream connections exist. Multiple stream connections on the same phenomena node are ranked based on the index of stream connection. An example is illustrated in Figure 1B. Firstly, the solution in Figure 1B shows a mixer connected with an inlet solvent (S) stream, inlet stream 1, and outlet stream 2. This gives a substring as SM>1<2. Moreover, the liquid-liquid extraction column contains 3 stages with feed stream supplied

on stage 1. The liquid-liquid extraction column connects with inlet stream 2 on stage 3 and connects with outlet stream 3 on stage 1. This gives a substring as |L>2LFL<3|. Furthermore, the distillation column involves 8 stages with the 4th stage as a feed stage, top stream connected with outlet stream 1, and bottom stream connected with outlet stream 4. Inlet stream 3 enters feed stage. This gives a substring as |V<1VVV>3VVVV<4|. Following the ascending order of stream index, we combine these three substrings, which gives the full string representation for the hybrid extraction and distillation system in Figure 1B. The construction procedure is automatically enabled through a python script with part of codes shown as follows:

```
a = 0
Next = []
open('file_name.lst', 'rt') as f:
    data = f.readlines()
for line in data:
    if line.__contains__('---- VAR Next '):
        num=re.findall("(?<=[AZaz])?(?!\\d*)"[0-9.-+]"",data[a])[3]
        Next.append(num)
    a=a+1
if (Next == []):
    Next=['0']
if (float(Next[-1]) >= 0):
    string = '|L'+>2'+L'+*(int(float(Next[-1]))-2)+'FL<3|'
```

Here, file\_name.lst is the *.lst* file containing the optimal solution of hybrid extraction and distillation model for each sampled feature vectors. The symbol 'Next' is the variable for the stage number of liquid-liquid extraction column in the *.lst* file, which represents the multiplicity of 'L' nodes in the string representation. With the developed python script, we converted 9816 samples obtained in the previous section into a data set with 9816 hybrid extraction and distillation flowsheets represented by phenomena-based strings. Note that the substring of S(M)>1<2 exists for all samples in this work, and hence this substring is redundant. Computational comparison suggests that removing this substring improves the





performance of training CAVE. In the result interpretation, we appended this substring at the beginning of generated string representation.

**Figure 2:** Distribution of separation cost in the full dataset (9816 samples).

Figure 2 illustrates the distribution of separation cost in the full dataset. Most samples involve separation cost ranging from \$20 per kmol product to \$80 per kmol product, which indicates economical separation using hybrid extraction and distillation for a feed with product mole fraction of 0.1. Note that the constant relative volatility assumption does not extend the current work to separation systems with azeotropes due to challenges in sample generation. Future work to allow the sample generation and construction of string representations for systems with azeotropes involves: (1) replacing the features of relative volatility with the features of parameters or grouped parameters in activity coefficient model (e.g., Wilson model and UNIQUAC model) to account for non-ideality in the liquid phase [6], (2) enumeration of representative separation structures that allow azeotropic distillation, and (3) modeling the azeotropic distillation using stage-by-stage model.

### Learning phenomena-based string representation using CVAE

With the phenomena-based string representation constructed in the previous subsection, we canonicalize each representation by adding 'E' at the end of string representation. We combine all strings with 'E' into an input matrix. Each entry of the input matrix is converted to an entry with the size of 200 by adding 'E'. Then, each entry is combined with a condition vector, which results in the final string. We fill the sampled values for  $\kappa_S$ ,  $\kappa_P$ ,  $\alpha_{W,P}$ ,  $\alpha_{S,P}$ ,  $\pi_S$ , and the cost of hybrid extraction and distillation into the condition vector (shown in the bottom of Figure 1B). The condition vector represents the known thermodynamic properties, solvent price, and cost value for an optimized hybrid extraction and distillation system. We combine all final strings with condition vectors into the processed input matrix.

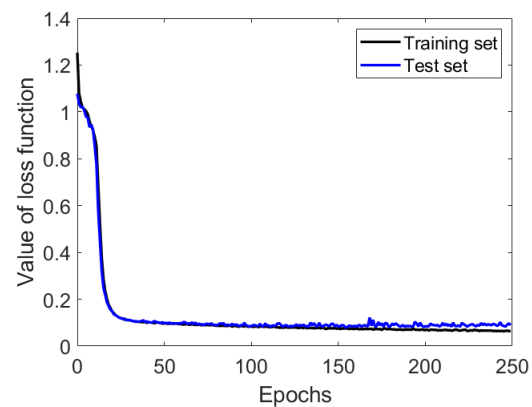
The processed input matrix is related with the encoder of the CVAE to generate a latent space (Figure 1C). We used the recurrent neural network (RNN) with the long short-term memory (LSTM) cell for both the encoder and decoder of the CVAE [29]. The architecture of the RNN consists of (1) 3-layer and (2) 500 hidden nodes on each layer. A layer with softmax function was used in each output of the decoder cell. The latent vector combined with the condition vector becomes an input of the decoder at each time step of the RNN cell. Finally, the output vector of each decoder cell was converted to a vector whose size is identical to that of the vector of the

input matrix. The softmax activation function was applied to each converted vector. The loss function of CVAE is given as follows [20]:

$$E[\log P(X|z, c)] - D_{KL}[Q(z|X, c) \parallel P(z|c)] \quad (1)$$

Here,  $E(f)$  represents the expectation value of function  $f$ .  $P$  and  $Q$  are probability distributions.  $D_{KL}$  is the kullback-leibler (KL) divergence.  $X$ ,  $z$ , and  $c$  indicate the data, latent space, and condition vector respectively. The first term is reconstruction error, and the second term is a regularization term that minimizes the KL-divergence between the variational and true encoders. We used cross entropy in the reconstruction error term of the cost function. To generate a flowsheet for hybrid extraction and distillation with given input feature values and desired separation cost imposed by the condition vector, the cell of the RNN decoder was unrolled for 200 times. The result was considered as invalid if the following conditions exist: (1) no 'L' exists within the 200 characters, (2) no 'V' exists within the 200 characters, (3) no 'E' exists within the 200 characters, (4) no '|' exists on the left side of the first 'L' symbol, and (5) no '|' exists on the right side of the last 'V' symbol. Each output vector of the decoder cell represents the probability distribution of the characters of the proposed phenomena-based representation and 'E'. Finally, the output vector was converted to the phenomena-based string representation. We implemented CVAE using TensorFlow 1.14.0 [30].

We set hyperparameters as follows: (1) 80 % of total samples are used for training and the remaining samples are used for test, and (2) the learning rate was set as 0.0001 and exponentially decayed at a rate of 0.97. The model was trained within 250 epochs. Figure 3 gives the



update of loss function values on training and test sets.

**Figure 3:** Update of loss function values of conditional variation autoencoder on training and test set.

### Results Discussion

In this Section, we use the proposed phenomena-based string representation as the descriptor and leverage the CVAE to generate hybrid extraction and

distillation systems with given thermodynamic property values and solvent price to meet desired cost investment.

The feature values are 4.59, 696.38, 1.42, 1.23, and \$5281.15 per kmol solvent for  $\kappa_P$ ,  $\kappa_S$ ,  $\alpha_{W,P}$ ,  $\alpha_{S,P}$ , and  $\pi_S$ , respectively. The separation cost for this feature vector is \$79.44 per kmol product according to the optimal solution for the hybrid extraction and distillation model. Based on this reference separation cost, we vary the separation cost value for the hybrid extraction and distillation system to screen the most economical design. We consider separation cost values as 10, 60, 80, 100, and 200 (with unit of \$ per kmol product). Five conditional vectors are constructed by filling into the feature values and separation cost value. We combine the condition vector with the latent vector, which are sampled by adding a Gaussian noise to the latent space of separation flowsheets selected randomly from the training set.

Based on the corresponding latent vectors and condition vectors, we generate the string representations and only keep the string representation with smallest length. These generated string representations with the smallest length for five separation cost values are given in Figure 4. The phenomena-based string representation (denoted as string 0) for the obtained optimal hybrid extraction and distillation system is as follows:

```
|L>2LLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLFL<3||V<1VVVVVVVVVVVVVVVVVVVVVVVVVVVV>3VVVVVVV<4|
```

This optimal separation system consists of one liquid-liquid extraction column and one distillation column.

As shown in Figure 4, the proposed data-driven framework has learned the representation of phenomena nodes (L and V), stream connections (< and >), and equipment boundary (|). For the separation cost of \$80 per kmol product, the predicted string representation is close to the one obtained from the optimal solution of MINLP model. This indicates good prediction performance. The differences are (1) the number of nodes (32 nodes) for liquid-liquid phase equilibrium (L) in the

\$10 per kmol product (string 1)	No valid string representations
\$60 per kmol product (string 2)	L>2LLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLFL<3  VV<1VVVVVVVVVVVVVV>3VVVVV<4  VV<1VVVVVVVVVVVVVV>4V
\$80 per kmol product (string 3)	L>2LLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLFL<3  V<1VVVVVVVVVVVVVVVVVVVVVV>3VVVVVVV<4
\$100 per kmol product (string 4)	L>2LLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLFL<3  V<1VVVVVV>3VVVVVVVVVVVVVVVVVVVVVVVVVVVVV<4
\$200 per kmol product (string 5)	L>2LLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLFL<3  V<1VVVVVVVVVVVVVV>3VVVVVVVVVVVVVVVVVVVVVVVVVVV<4

predicted string representation is more than those (31 nodes) in the string representation obtained from the solution of separation process optimization, and (2) the number of nodes (38 nodes) for vapor-liquid phase equilibrium (V) in the predicted string representation is more than those (37 nodes) in the string representation obtained from the solution of separation process optimization.

**Figure 4:** Generated phenomena-based string representations with varying separation costs.

At the separation cost of \$10 per kmol product, no valid string representations exist, which indicates that the desired separation cost of \$10 per kmol product can not be achieved. Moreover, we observe that comparing the string representations for \$60 per kmol product and \$80 per kmol product, the predicted string representation has larger length for \$60 per kmol product, which indicates more stages in liquid-liquid extraction column or distillation column. The increase of stage number in liquid-liquid extraction column or distillation column can decrease operating costs related with solvent or utility consumption. We also note that string 2 for \$60 per kmol product shows a separation system with two distillation columns. For the design of liquid-liquid extraction column, the separation system for string 2 has more extraction stages (49 stages) than the separation system for string 3 (32 stages). For the design of distillation column, the separation system for string 2 has more distillation stages (39 stages) than the separation system for string 3 (38 stages). This case indicates that other factors that influence separation costs are solvent and utility consumption, which are continuous variables and are not captured in the proposed string representation. When the desired separation cost is higher than \$80 per kmol product, we also observe string representations with larger length. This is because more stages in liquid-liquid extraction column or distillation column make capital costs for separation systems represented by string 4, 5 higher than the capital cost for separation system represented by string 3.

Moreover, the above analysis also indicates that more economical separation design can be achieved by adjusting solvent and utility consumption. Specifically, we observe an improved solution (string 2) for designing hybrid extraction and distillation system. To validate this, we fix the design of liquid-liquid extraction column, and activate two distillation columns in the MINLP model, and we obtain the optimal solution as \$75.5 per kmol product. Although this improved solution is still not identical to the desired separation cost of \$60 per kmol product, we expect to improve exploration performance of CVAE when using a training set with more samples.

## CONCLUSION

In this work, we propose a phenomena-based string presentation for learning and screening the designs of hybrid extraction and distillation. In the proposed representation, we leverage the phenomena building blocks of liquid-liquid phase equilibrium, vapor-liquid phase equilibrium, and mixing while considering stream connections and equipment boundary. We train conditional variational autoencoder with condition vector on the feature values of liquid-liquid equilibrium constants of solvent, and product, relative volatilities, and solvent price. The trained conditional variational autoencoder shows good prediction performance. We apply the proposed framework for learning and screening separation systems for product separation with product mole fraction in the feed as 0.1. With the trained conditional variational autoencoder, we fix the values of input features, vary the values of separation cost, and predict the corresponding optimal designs of hybrid extraction and distillation represented by phenomena-based strings. Our proposed approach shows promise as a bottom-up data representation for flowsheet learning, data standardization, and generation of flowsheet database. Future work involves the consideration of more phenomena node options, automatic conversion of phenomena-based string representations to equipment-based chemical flowsheets, and conversion among different representations of flowsheet data.

## REFERENCES

1. Korley, L.T. et. al. Toward polymer upcycling—adding value and tackling circularity. *Science*, 373(6550), 66-69 (2021).
2. Hasan, M.M.F., Zantye, M.S. and Kazi, M.K. Challenges and opportunities in carbon capture, utilization, and storage: A process systems engineering perspective. *Comput Chem Eng*, 107925 (2022).
3. Li, J., Maravelias, C.T. and Van Lehn, R.C. Adaptive Conformer Sampling for Property Prediction Using the Conductor-like Screening Model for Real Solvents. *Ind Eng Chem Res*, 61(25), 9025-9036 (2022).
4. Xu, S., Cremaschi, S., Eden, M.R. and Tula, A.K. An integrated framework for sustainable process design by hybrid and intensified equipment. *Comput Chem Eng*, 176, 108288 (2023).
5. Castillo-Landero, A., Dominguillo-Ramirez, D., Aburto, J., Sadhukhan, J. and Martinez-Hernandez, E. Improving the Economic, Environmental, and Safety Performance of Bio-Jet Fuel Production through Process Intensification and Integration Using a Modularity Approach. *Acs Sustain Chem Eng*, 11(2), 660-669 (2023).
6. Li, J. and Hasan, M.M.F. A parametric approach to identify synergistic domains of process intensification for reactive separation. *Chem Eng Sci*, 267, 118337 (2023).
7. Kraemer, K. et. al. Separation of butanol from acetone-butanol-ethanol fermentation by a hybrid extraction-distillation process. *Comput Chem Eng*, 35(5), 949-963 (2011).
8. Schweidtmann, A.M., Esche, E., Fischer, A., Kloft, M., Repke, J.U., Sager, S. and Mitsos, A. Machine learning in chemical engineering: A perspective. *Chem Ing Tech*, 93(12), 2029-2039 (2021).
9. Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J Chem Inf Comput Sci*, 28(1), 31-36 (1988).
10. Heller, S., McNaught, A., Stein, S., Tchekhovskoi, D. and Pletnev, I. InChI-the worldwide chemical structure identifier standard. *J Cheminf*, 5, 1-9 (2013).
11. Krenn, M., Ai, Q., Barthel, S., Carson, N., Frei, A., Frey, N.C., Friederich, P., Gaudin, T., Gayle, A.A., Jablonka, K.M. and Lameiro, R.F. SELFIES and the future of molecular string representations. *Pattern*, 3(10) (2022).
12. Drefahl, A. CurlySMILES: a chemical language to customize and annotate encodings of molecular and nanodevice structures. *J Cheminf*, 3, 1-7 (2011).
13. Lin, T.S., Coley, C.W., Mochigase, H., Beech, H.K., Wang, W., Wang, Z., Woods, E., Craig, S.L., Johnson, J.A., Kalow, J.A. and Jensen, K.F. BigSMILES: a structurally-based line notation for describing macromolecules. *ACS Cent Sci*, 5(9), 1523-1531 (2019).
14. d'Anterrosches, L. and Gani, R. Group contribution-based process flowsheet synthesis, design and modelling. *Fluid Ph Equilibria*, 228, 141-146 (2005).
15. Zhang, T., Sahinidis, N.V. and Siirola, J.J. Pattern recognition in chemical process flowsheets. *AIChE J*, 65(2), 592-603 (2019).
16. Vogel, G., Hirtreiter, E., Schulze Balhorn, L. and Schweidtmann, A.M. SFILES 2.0: an extended text-based flowsheet representation. *Optim Eng*, 1-23 (2023).
17. Mann, V., Sales-Cruz, M., Gani, R. and Venkatasubramanian, V. eSFILES: Intelligent process flowsheet synthesis using process knowledge, symbolic AI, and machine learning. *Comput Chem Eng*, 108505 (2023).
18. Medina, E.I.S., Linke, S., Stoll, M. and Sundmacher, K. Gibbs-Helmholtz graph neural network: capturing the temperature dependency of activity coefficients at infinite dilution. *Digit Discov*, 2(3), 781-798 (2023).
19. Gómez-Bombarelli, R. et.al. Automatic chemical

- design using a data-driven continuous representation of molecules. *ACS Cent Sci*, 4(2), 268-276 (2018).
20. Kingma, D.P., Mohamed, S., Jimenez Rezende, D. and Welling, M. Semi-supervised learning with deep generative models. *Adv Neural Inf Process Syst*, 27 (2014).
  21. Lim, J., Ryu, S., Kim, J.W. and Kim, W.Y. Molecular generative model based on conditional variational autoencoder for de novo molecular design. *J Cheminform*, 10(1), 1-9 (2018).
  22. Li, J., Van Lehn, R.C. and Maravelias, C.T. An Explainable Classification Framework for Determining and Understanding the Suitability of Solvent Extraction for Bioproduct Recovery. *Acs Sustain Chem Eng*. Accepted. doi.org/10.1021/acssuschemeng.3c07129. (2024)
  23. Li, J., Demirel, S.E. and Hasan, M.M.F. Process synthesis using block superstructure with automated flowsheet generation and optimization. *AIChE J*, 64(8), 3082-3100 (2018).
  24. Palach, J. Parallel programming with Python; Packt Publishing Ltd, 2014.
  25. Lutze, P., Babi, D.K., Woodley, J.M. and Gani, R. Phenomena based methodology for process synthesis incorporating process intensification. *Ind Eng Chem Res*, 52(22), 7127-7144 (2013).
  26. Skiborowski, M. Synthesis and design methods for energy-efficient distillation processes. *Curr Opin Chem Eng*, 42, 100985 (2023).
  27. Monjur, M.S., Demirel, S.E., Li, J. and Hasan, M.M.F. SPICE\_MARS: a process synthesis framework for membrane-assisted reactive separations. *Ind Eng Chem Res*, 60(20), 7635-7655 (2021).
  28. Chu, K.T., Cadoret, L., Yu, C.C. and Ward, J.D., 2011. A new shortcut design method and economic analysis of divided wall columns. *Ind Eng Chem Res*, 50(15), 9221-9235 (2011).
  29. Hochreiter, S. and Schmidhuber, J. Long short-term memory. *Neural Comput*, 9(8), 1735-1780 (1997).
  30. Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M. and Kudlur, M. TensorFlow: a system for Large-Scale machine learning. In 12th USENIX symposium on operating systems design and implementation (OSDI 16), 265-283 (2016).

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Optimizing Batch Crystallization with Model-based Design of Experiments

Hailey G. Lynch<sup>a</sup>, Aaron Bjarnason<sup>b</sup>, Daniel J. Laky<sup>a</sup>, Cameron J. Brown<sup>b</sup>, and Alexander W. Dowling<sup>a\*</sup>

<sup>a</sup> University of Notre Dame, Department of Chemical and Biomolecular Engineering, Notre Dame, Indiana 46556, USA

<sup>b</sup> University of Strathclyde, EPSRC Future Manufacturing Research Hub for Continuous Manufacturing and Advanced Crystallisation (CMAC), Glasgow G1 1RD, UK

\* Corresponding Author: [adowling@nd.edu](mailto:adowling@nd.edu).

## ABSTRACT

Adaptive and self-optimizing intelligent systems such as digital twins are increasingly important in science and engineering. Digital twins utilize mathematical models to provide added precision to decision-making. However, physics-informed models are challenging to build, calibrate, and validate with existing data science methods. Model-based design of experiments (MBD<sub>oE</sub>) is a popular framework for optimizing data collection to maximize parameter precision in mathematical models and digital twins. In this work, we apply MBD<sub>oE</sub>, facilitated by the open-source package Pyomo.DoE, to train and validate mathematical models for batch crystallization. We quantitatively examined the estimability of the model parameters for experiments with different cooling rates. This analysis provides a quantitative explanation for the heuristic of using multiple experiments at different cooling rates.

**Keywords:** Model-based Design, Batch Crystallization, Intelligent Systems, Pyomo, Digital Twins

## INTRODUCTION

### Digital Twins and Model-based Design of Experiments

Spearheading the data-driven revolution, the industrial implementation of inexpensive sensors and other data-capturing technologies has increased the availability of information [1]. However, the collected data has significant analysis limitations due to existing software and modeling capabilities [2]. At the confluence of these challenges, digital twins, i.e., virtual representations of physical systems, are increasing across industries ranging from agriculture to medicine [3,4]. At the heart of many digital twins are science-based mathematical models with physically meaningful parameters estimated from sensor data, often in real-time. These science-based models are created, maintained, and updated using data science tools, including model selection, parameter estimation, sensitivity and uncertainty analysis, and model-based design of experiments (MBD<sub>oE</sub>). MBD<sub>oE</sub> supports improved parameter precision or model discrimination for science-based models with physically meaningful parameters [5] through the following form:

$$y_i = m(x_i, \theta) + \epsilon_i, \epsilon_i \sim \mathcal{N}(0, \sigma_\epsilon^2) \quad (1)$$

where  $y$  represents the  $i$ -th measurement from an experiment with an observation error  $\epsilon_i$  following an assumed probability distribution, e.g., independent and identically normally distributed with mean 0 and variance  $\sigma_\epsilon^2$ . Here,  $m(\cdot, \cdot)$  denotes the mathematical model,  $x_i$  depicts the input and control variables for the  $i$ -th measurement, and  $\theta$  are the model parameters.

Training and validating a mathematical model using MBD<sub>oE</sub> is done sequentially. The model is updated and recalibrated before determining the next best set of experimental conditions to maximize information gain [6]. MBD<sub>oE</sub> aims to improve parameter precision in the estimated parameters by minimizing the associated uncertainty [5,7]. Encoding the mathematical model requires initial parameter values. Estimating the values for these parameters is done by using data collected from simulations of the experimental procedure. Model parameters are estimated from data using nonlinear regression:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \Psi := \frac{1}{2} \sum_i [y_i - m(x_i, \theta)]^2 \quad (2)$$

For more sophisticated observation error assumptions,



maximum likelihood estimation, variational inference, or Bayesian methods are used instead of Eq. (2).

Next, we consider the Hessian matrix  $\mathbf{H}$  to quantify the sensitivity of least-squares objective  $\Psi$  to the perturbations in  $\theta$ :

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 \Psi}{\partial \theta_1^2} & \dots & \frac{\partial^2 \Psi}{\partial \theta_n \partial \theta_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \Psi}{\partial \theta_1 \partial \theta_m} & \dots & \frac{\partial^2 \Psi}{\partial \theta_m^2} \end{bmatrix} \quad (3)$$

When the residuals  $y_i - m(x_i, \theta)$  are small,  $\mathbf{H} \approx \mathbf{Q}^T \mathbf{Q}$ , where  $\mathbf{Q}$  is the sensitivity of the model predictions with respect to the model parameters  $\theta$  [8]. The sensitivity matrix  $\mathbf{Q}(\theta)$  is as follows:

$$\mathbf{Q}(\theta) = \begin{bmatrix} \frac{\partial m(x_1, \theta)}{\partial \theta_1} & \dots & \frac{\partial m(x_1, \theta)}{\partial \theta_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial m(x_n, \theta)}{\partial \theta_1} & \dots & \frac{\partial m(x_n, \theta)}{\partial \theta_m} \end{bmatrix} \quad (4)$$

MBDoE uses the Fisher information matrix (FIM)  $M_{\hat{\theta}}$  to measure the information content about parameters  $\theta$  contained in proposed experiment(s)  $x_i$  for the model  $m(x_i, \theta)$ :

$$\mathbf{M}_{\hat{\theta}} \approx \Sigma_{\hat{\theta}}^{-1} \approx \frac{1}{\sigma_{\epsilon}^2} (\mathbf{Q}^T \mathbf{Q}) \quad (5)$$

where  $\sigma_{\epsilon}^2$  is the variance of the measurement uncertainty  $\epsilon$  and  $\Sigma_{\hat{\theta}}$  is the covariance of the estimated parameters. Finally, the MBDoE optimization formulation is given by:

$$\mathbf{x}^* = \underset{x}{\operatorname{argmax}} \Phi[\mathbf{M}(\hat{\theta}, x) + \mathbf{M}_0] \quad (6)$$

where  $\mathbf{x}^*$  are the optimal conditions for the next experiment considering prior information  $\mathbf{M}_0$ .  $\Phi(\cdot)$  is a matrix operation such as determinant (D-optimality), trace (A-optimality) or smallest eigenvalue (E-optimality).

Let  $\lambda_i$  and  $\mathbf{v}_i$  represent the eigenvalues and eigenvectors of  $\mathbf{M}_{\hat{\theta}}$ . Because  $\mathbf{M}_{\hat{\theta}} \approx \Sigma_{\hat{\theta}}^{-1}$ , the eigenvalues and eigenvectors of  $\Sigma_{\hat{\theta}}^{-1}$  are approximately  $\lambda_i^{-1}$  and  $\mathbf{v}_i$ . The covariance matrix  $\Sigma_{\hat{\theta}}$  is often represented as a confidence ellipsoid with axes  $\lambda_i^{-1} \mathbf{v}_i$  [5].

Model  $m(\cdot, \cdot)$  is identifiable if there exists a (possibly infinite) dataset for which it is possible to uniquely estimate the model parameters [9]. Structural issues, such as a parameter appearing only in a sum ( $\theta_1 + \theta_2$ ) or product ( $\theta_1 \times \theta_2$ ) will cause a model to be not identifiable; even with an infinite amount of data it is impossible to uniquely estimate both parameters ( $\theta_1$  and  $\theta_2$ ). In contrast, this short paper analyzes the practical identifiability, i.e., estimability, [10] of parameters in a batch crystallization model considering specific sets of experiments.

## Crystallization

Crystallization is a fundamental unit operation for the purification of substances necessary for objectives relating to separations for critical mineral recovery [11],

agriculture [12], food [13], fine chemicals [14], and pharmaceuticals [15]. Depending on the application (i.e., fertilizers, product quality and shelf life, fine chemicals, and active pharmaceutical ingredients), different critical quality attributes (i.e., crystal size, distribution) greatly impact the efficacy of the final product [16-19]. Tailored experimental procedures are required for each application to identify desired crystal properties. This remains difficult due to the highly nonlinear nature of crystallization dynamics and the need to control the operating conditions necessary for crystallization [20]. Due to the experimental design conditions of crystallization, open challenges include developing processes in confined volumes [21] as well as continuous manufacturing with time and material constraints [22].

The basic driving mechanisms for crystallization are nucleation and growth, which are thermodynamically favorable through a concentration gradient (i.e., supersaturation). Kinetic parameters for these processes are traditionally found through experimental procedures [23]. As such, determining the size and shape distribution of crystals is attainable through crystallization modeling [24]. Understanding crystallization processes with kinetic parameters is often achieved through population balance modeling [25]. However, estimating these parameters in crystallization models from experimental data is challenging due to the numerical complexities in a population balance equation (PBE) [26].

## Crystallization DoE Heuristics

For *in-silico* processes, we focus on the heuristics in crystallization for designing improved experimental campaigns. It is common practice to estimate primary nucleation ( $k_{p_1}$  and  $b_1$ ) through induction time measurements, i.e., unseeded with no cooling ( $\beta=0$ ), then later estimating secondary nucleation ( $k_{b_2}$  and  $b_2$ ) from seeded experiments. An aging process is necessary after the seeding process, and the decline of supersaturation can arise from growth only ( $k_g$  and  $g$ ) or secondary nucleation and growth, which is based on the kinetics [27]. For unseeded experiments, a suitable seed for secondary nucleation is obtained from primary nucleation [28]. The growth of the crystal nuclei is necessary for detecting primary nucleation and activating secondary nucleation [29]. The rates for nucleation as well as crystal growth are usually dependent on the power law relations of supersaturation [30].

## Paper Contribution

This paper uses practical identifiability analyses to provide mathematical insights into the above-mentioned crystallization modeling heuristics.

## CRYSTALLIZATION MODEL

We now describe a simple batch crystallization model. We then analyze the practical identifiability of crystallization kinetics parameters  $k_{b_1}, b_1, k_{b_2}, b_2, k_g,$  and  $g$  (described below) using one or more unseeded experiment performed at different cooling rates  $\beta$ .

## Population Balance Equation

A standard PBE represents the change in particle population for a specified volume. For the given problem, a one-dimensional PBE was used to describe the batch crystallization system with the form:

$$\frac{\partial n(L,t)}{\partial t} + \frac{\partial Gn(L,t)}{\partial L} = B\delta(L - L_0) \quad (7)$$

where the number of crystals per crystal length and the volume of slurry are depicted by the number density function  $n(L, t)$ ,  $G$  is the crystal growth rate,  $B$  is the nucleation rate,  $L$  is the length of the crystals,  $L_0$  is the nuclei size,  $\delta$  is the Dirac delta function,  $\sigma$  is relative supersaturation, and  $t$  represents the time. The kinetics for growth and nucleation of particles can be described with power laws:

$$\sigma = \frac{C - C_{sat}}{C_{sat}} \quad (8)$$

$$G = k_g \sigma^g \quad \text{if } \sigma > 0 \quad (9)$$

$$B_p = k_{b_1} \sigma^{b_1} \quad \text{if } \sigma > 0 \quad (10)$$

$$B_s = \varepsilon M_T k_{b_2} \sigma^{b_2} \quad \text{if } \sigma > 0 \quad (11)$$

$$B = B_p + B_s \quad (12)$$

Total nucleation rate ( $B$ ) is often described with both primary ( $B_p$ ) and secondary ( $B_s$ ) mechanisms. Here  $k_g, k_{b_1}, k_{b_2}$  and  $g, b_1, b_2$  are the rate constants and the power law exponents for growth, primary nucleation, and secondary nucleation, respectively. These six parameters are the focus of our analysis.  $M_T$  is the slurry density and  $\varepsilon$  is the power density.  $C_{sat}$  is the saturation concentration for the system temperature given in the equation below:

$$C_{sat} = A_1 \exp(A_2 T) \quad (13)$$

where  $A_1$  and  $A_2$  are parameters found through a solubility study. For an unseeded system (i.e.,  $n=0$ ), the initial and boundary conditions for this PBE are given by:

$$n(L, 0) = 0 \quad (14)$$

$$n(0, t) = \frac{B\delta(L-L_0)}{G} \quad \text{if } \sigma > 0 \quad (15)$$

where we assume that the initial birth rate is negligible and growth rate is independent of crystal length. As time progresses, a solute mass balance equation is used to define the change in concentration inside the crystallizer such that:

$$\frac{dC}{dt} = -3\rho_c K_v \int_0^\infty nGL^2 dL \quad (16)$$

with  $C$  as the solute concentration,  $\rho_c$  as the solid density

of a crystal, and  $K_v$  as the volumetric shape factor.

## Method of Moments

To reduce the complexity of the PBE, the method of moments converted the partial differential equations of the PBE to a set of ordinary differential equations [32] by solving the equations in the model using the first four moments calculated through the population density function:

$$\frac{d\mu_0}{dt} = B \quad (17)$$

$$\frac{d\mu_i}{dt} = iG\mu_{i-1} \quad \forall i \in \{1,2,3,4\} \quad (18)$$

$$\frac{dC}{dt} = -\frac{dM_T}{dt} = -3\rho_c K_v G \mu_2 \quad (19)$$

where  $\mu_i$  is the  $i$ -th moment of the crystal size distribution. The moments represent a physical significance per unit volume for total crystal number, total crystal length, total crystal surface area, and total crystal volume.

## Model Implementation in Pyomo

A typical experiment was defined for a batch crystallization system in the Pyomo modeling environment. We fixed the parameters for a solid density of a crystal  $\rho_c$ , volume shape factor  $K_v$ , and the solubility parameters  $C_1$  and  $C_2$ . For the linear cooling of the crystal suspension, an ordinary differential equation was used to describe temperature of the form:

$$\frac{dT}{dt} = -\beta \quad (20)$$

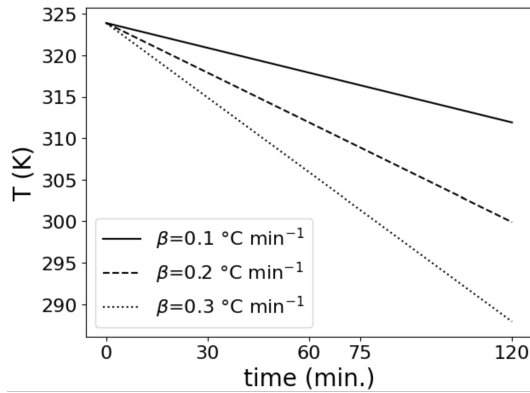
where  $\beta$  refers to the cooling rate. Empirical equations were used to define the growth and birth rate in the crystallizer [33,34]. The FIM  $\mathbf{M} \in \mathbb{R}^{6 \times 6}$  was calculated using Pyomo.DoE [31] for the following six parameters:  $k_{b_1}$  and  $b_1$  for primary nucleation,  $k_{b_2}$  and  $b_2$  for secondary nucleation, and  $k_g$  and  $g$  for growth.

## RESULTS

We simulated the batch crystallization experiment at four different cooling rates:  $\beta=0, \beta=0.1, \beta=0.2,$  and  $\beta=0.3$  in  $^\circ\text{C min}^{-1}$ . For each experiment, we calculated the FIM. Finally, we analyzed the practical identifiability of each experiment and several collections of multiple experiments. All experiments start as a 1% supersaturated solution.

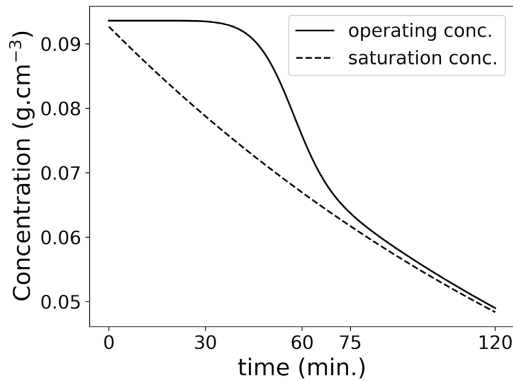
## Simulation Profiles

Figure 1 shows the temperature decreases linearly with time for  $\beta=0.1, 0.2,$  and  $0.3 \text{ }^\circ\text{C min}^{-1}$  as expected. As noted previously, when applying cooling to a vessel, solubility of the solute generally decreases. Therefore, when the temperature of a solution changes, supersaturation is generated thermodynamically, causing the solute to crystallize out of solution.

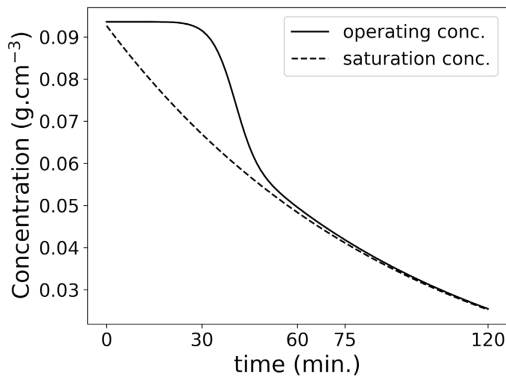


**Figure 1.** Change in temperature (K) over time for  $\beta=0.1$ , 0.2, and 0.3 ( $^\circ\text{C min}^{-1}$ ).

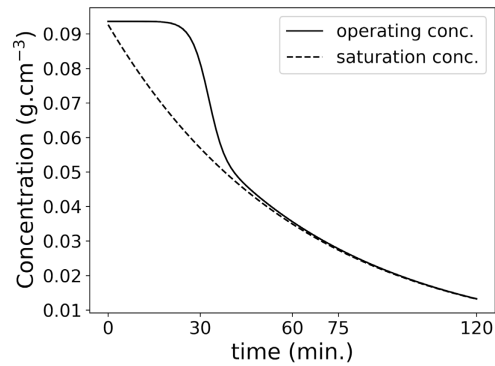
Subsequently, Figures 2 to 4 show the decrease in solute saturation concentration and operating concentration as a function of time for  $\beta=0.1$ , 0.2, and 0.3  $^\circ\text{C min}^{-1}$ . As expected, these concentrations decrease the fastest with the fastest cooling. This makes sense because the solubility will decrease as temperature decreases such that the solute separates from the solution as crystals form.



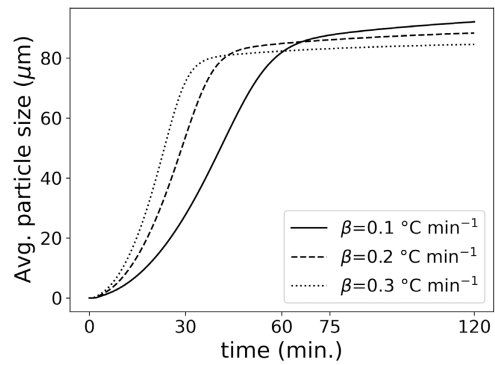
**Figure 2.** Operating and saturation concentration ( $\text{g cm}^{-3}$ ) over time for  $\beta=0.1$  ( $^\circ\text{C min}^{-1}$ ).



**Figure 3.** Operating and saturation concentration ( $\text{g cm}^{-3}$ ) over time for  $\beta=0.2$  ( $^\circ\text{C min}^{-1}$ ).



**Figure 4.** Operating and saturation concentration ( $\text{g cm}^{-3}$ ) over time for  $\beta=0.3$  ( $^\circ\text{C min}^{-1}$ ).



**Figure 5.** Change in average particle size ( $\mu\text{m}$ ) over time for  $\beta=0.1$ , 0.2, and 0.3 ( $^\circ\text{C min}^{-1}$ ).

Figure 5 investigates the change in average particle size over time for three different cooling rates to understand the growth distribution. As the cooling rate increases, the average particle size grows at a faster rate. This is likely due to incubation time (time to first nuclei forming) decreasing as the cooling rate increases, resulting in a longer period of low supersaturation where growth is the dominant mechanism.

### Eigendecomposition of the FIM

We now analyze the eigendecompositions of the FIMs from proposed experiments to explore the estimability of the model. Recall, the basis of eigenvectors is defined with respect to the model parameters. In Tables 1 to 6 (after references), each row (except the last) corresponds to a model parameter. Each column is an eigenvector. The last row reports the eigenvalues. Within each column, the largest absolute eigenvector element is highlighted in blue. For example, in Table 1, the fifth and sixth columns report  $\lambda_5 = 9.270 \times 10^{-10}$  and  $\lambda_6 = 2.622 \times 10^{-12}$ .

Both eigenvalues are small, indicating significant practical identifiability issues. Specifically, the largest components of  $\mathbf{v}_5$  and  $\mathbf{v}_6$  correspond to the secondary nucleation parameters  $k_{b_2}$  and  $b_2$ . In other words, these eigenvectors are predominantly in the direction of  $k_{b_2}$  and

**Table 1:** Eigendecomposition of the FIM for the  $\beta=0$  (no cooling) candidate experiment. The columns are eigenvectors, and the rows are model parameters. The last row are the eigenvalues. The largest absolute elements in each vector are highlighted in blue.

	$\mathbf{v}_1$	$\mathbf{v}_2$	$\mathbf{v}_3$	$\mathbf{v}_4$	$\mathbf{v}_5$	$\mathbf{v}_6$
$k_{b_1}$	$7.261 \times 10^{-2}$	$-6.303 \times 10^{-1}$	$4.392 \times 10^{-2}$	$-7.524 \times 10^{-1}$	$1.716 \times 10^{-1}$	$-2.966 \times 10^{-3}$
$b_1$	$-8.827 \times 10^{-2}$	$7.662 \times 10^{-1}$	$-2.861 \times 10^{-2}$	$-6.186 \times 10^{-1}$	$1.469 \times 10^{-1}$	$-2.069 \times 10^{-3}$
$k_{b_2}$	$4.403 \times 10^{-4}$	$-3.815 \times 10^{-3}$	$-4.539 \times 10^{-3}$	$1.758 \times 10^{-1}$	$7.684 \times 10^{-1}$	$6.154 \times 10^{-1}$
$b_2$	$-2.953 \times 10^{-4}$	$2.562 \times 10^{-3}$	$2.497 \times 10^{-3}$	$-1.417 \times 10^{-1}$	$-5.988 \times 10^{-1}$	$7.882 \times 10^{-1}$
$k_g$	$7.922 \times 10^{-1}$	$1.214 \times 10^{-1}$	$5.979 \times 10^{-1}$	$9.889 \times 10^{-3}$	$9.894 \times 10^{-4}$	$5.374 \times 10^{-4}$
$g$	$5.995 \times 10^{-1}$	$2.876 \times 10^{-2}$	$-7.998 \times 10^{-1}$	$-1.324 \times 10^{-2}$	$-1.322 \times 10^{-3}$	$-7.19 \times 10^{-4}$
$\lambda_i$	$2.192 \times 10^3$	$1.754 \times 10^1$	$2.118 \times 10^{-4}$	$1.052 \times 10^{-7}$	$9.270 \times 10^{-10}$	$2.622 \times 10^{-12}$

**Table 2:** Eigendecomposition of the FIM for the  $\beta=0.1$  ( $^{\circ}\text{C min}^{-1}$ ) candidate experiment.

	$\mathbf{v}_1$	$\mathbf{v}_2$	$\mathbf{v}_3$	$\mathbf{v}_4$	$\mathbf{v}_5$	$\mathbf{v}_6$
$k_{b_1}$	-0.031	0.029	-0.226	0.279	-0.355	0.862
$b_1$	0.061	-0.084	0.493	-0.768	-0.298	0.260
$k_{b_2}$	-0.513	0.818	0.067	-0.038	-0.222	-0.107
$b_2$	0.117	-0.198	-0.092	0.163	-0.858	-0.419
$k_g$	-0.808	-0.495	-0.260	-0.186	0.006	-0.019
$g$	-0.256	-0.194	0.791	0.520	0.005	0.039
$\lambda_i$	$5.753 \times 10^4$	$4.905 \times 10^3$	$2.357 \times 10^1$	$9.083 \times 10^0$	$1.441 \times 10^{-2}$	$2.127 \times 10^{-3}$

**Table 3:** Eigendecomposition of the FIM for the  $\beta=0.2$  ( $^{\circ}\text{C min}^{-1}$ ) candidate experiment.

	$\mathbf{v}_1$	$\mathbf{v}_2$	$\mathbf{v}_3$	$\mathbf{v}_4$	$\mathbf{v}_5$	$\mathbf{v}_6$
$k_{b_1}$	-0.025	0.009	0.318	-0.250	-0.095	-0.909
$b_1$	0.045	-0.032	-0.651	0.548	0.321	-0.413
$k_{b_2}$	-0.624	0.751	-0.020	-0.045	0.211	0.009
$b_2$	0.125	-0.165	0.206	-0.275	0.915	0.047
$k_g$	-0.743	-0.605	0.174	0.228	0.022	0.011
$g$	-0.203	-0.205	-0.634	-0.713	-0.080	-0.014
$\lambda_i$	$6.317 \times 10^4$	$3.206 \times 10^3$	$2.081 \times 10^1$	$1.897 \times 10^0$	$3.127 \times 10^{-2}$	$6.362 \times 10^{-6}$

$b_2$ . Because  $\lambda_5$  and  $\lambda_6$  are small, the proposed experiment contains little information about these two parameters. On the other hand, we can see for the largest eigenvalues,  $\lambda_1$  and  $\lambda_2$ , the predominant direction of  $\mathbf{v}_1$  is growth ( $k_g$ ), and  $\mathbf{v}_2$  is primary nucleation ( $b_1$ ), which contain the most information for this experiment. This makes sense since this experiment is not seeded, so we obtain information about primary nucleation  $b_1$ .

Table 2 shows a significant increase in the magnitude of the smallest eigenvalues and the number of eigenvalues above one for  $\beta=0.1$  ( $^{\circ}\text{C min}^{-1}$ ) compared to  $\beta=0$  ( $^{\circ}\text{C min}^{-1}$ ) (Table 1). This corresponds with an improved practical identifiability in the growth parameter ( $g$ ). This is likely because using a non-zero cooling rate disposes the system to the thermodynamic driving force:

solute supersaturation. As the cooling rate increases, the smallest eigenvalue decreases from order  $10^{-3}$  (Table 2) to  $10^{-6}$  (Table 3). In both tables,  $\mathbf{v}_6$  is predominantly in the direction of  $k_{b_1}$ .

As the cooling rate increases (Table 4), the magnitude and primary direction of the eigenvalues and eigenvectors change. The smallest eigenvalue increases from order  $10^{-6}$  (Table 3) to  $10^{-5}$  (Table 4). Primary nucleation  $k_{b_1}$  remains the parameter with the least information for Tables 2 to 4.

Often, in crystallization studies, multiple cooling rates are used to identify growth. Thus, we consider two possible experiment campaigns: (1) two experiments at  $\beta=0.2$  and  $\beta=0.3$  and (2) three experiments at  $\beta=0.1$ ,  $\beta=0.2$ , and  $\beta=0.3$ . For both scenarios, the FIMs are

summed before computing the eigendecomposition. Consistent with intuition, considering two experiments (Table 5) results in a larger smallest eigenvalue (order  $10^{-3}$ ) compared to each cooling rate separately (order  $10^{-6}$  to  $10^{-5}$ ). Similarly, three experiments (Table 6) further increase the smallest eigenvalue (order  $10^{-2}$ ).

## Optimality of the FIM

We now examine the A-, D-, and E-optimality measures of the FIMs from the proposed set of experiments to explore the information content of the cooling rates. Recall, the optimality criterion is defined by the trace (Eq. 21), determinant (Eq. 22), and smallest eigenvalue (Eq. 23) of the FIM:

$$\text{trace}(\mathbf{M}) = \sum_i \lambda_i \quad (21)$$

$$\det(\mathbf{M}) = \prod_i \lambda_i \quad (22)$$

$$\min(\text{eig}(\mathbf{M})) = \min(\lambda_1, \dots, \lambda_n) \quad (23)$$

where  $\lambda_i$  is the eigenvalue of the FIM.

As denoted in Figure 6, the optimality measure has an overall increase for the experiments with the multiple cooling rates as opposed to the experiments with individual cooling rates. We can see that for no cooling (i.e.,  $\beta=0$ ), there is minimal optimal measure. However, for the individual cooling rates, A-optimality is nearly constant, D-optimality decreases slightly, and the optimality measure is marginally better for E-optimality with  $\beta=0.1$  °C min<sup>-1</sup>. As expected, when we consider two measurements, we have more information for A- and D-optimality. Overall, the experiment set with the three cooling rates contains the most information.

## Computational Environment

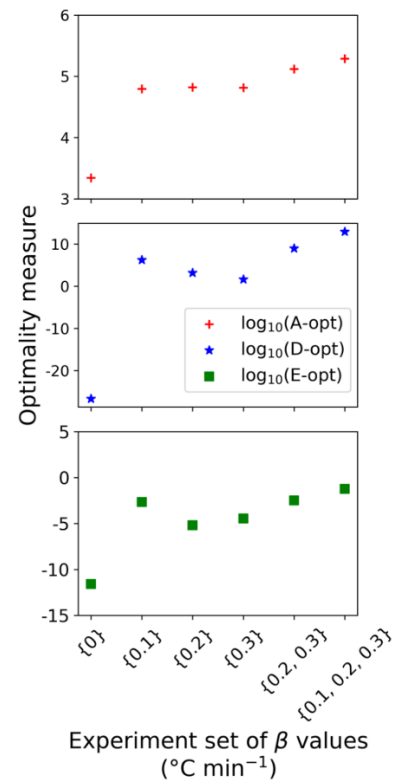
The computational results were conducted using the following software versions: Python 3.10.12, Ipopt 3.13.2, Pyomo 6.7.1, and CasADi 3.6.5. The computer specifications include Windows 11 Version 22H2, Intel(R) Core (TM) i7-10510U CPU @ 1.80 GHz 2.30 GHz, and 16.0 GB RAM. For each individual cooling rate ( $\beta=0$ ,  $\beta=0.1$ ,  $\beta=0.2$ , and  $\beta=0.3$ ), we found the average computational times for running the model, computing the FIM, and the eigendecomposition of the FIM which were approximately 0.078 s, 8.881 s, and 0.124 s respectively.

## CONCLUSIONS

Through our analysis of the practical identifiability of parameters for a batch crystallization system, we found that growth appears to be the most estimable parameter as it consistently corresponds to the largest eigenvalues. As expected, the secondary nucleation parameters  $k_{b_2}$  and  $b_2$  were not estimable in constant temperature experiments ( $\beta=0$ ). If only one experiment is possible, a modest cooling rate ( $\beta=0.1$  °C min<sup>-1</sup>) is

recommended to maximize estimability. However, performing experiments at multiple cooling rates is the most informative, as expected. Further analysis for the optimality measure of the FIM demonstrated an increase of information content when changing the cooling rate or adding multiple experiments.

In conclusion, MBD<sub>oE</sub> and practicable identifiability analysis provides a mathematical context for a common replicate crystallization heuristic. This analysis is limited to this model, and although the framework is general, the specific conclusions may not generalize. Analysis like that performed in this work demonstrates a mathematical impetus for fitting selective parameter sets for specific experiments. There is an opportunity to integrate these insights with existing crystallization parameter estimation workflows and tools [35].



**Figure 6.** A-, D-, and E-optimality measures for six sets of candidate experiments define by their  $\beta$  values (°C min<sup>-1</sup>).

## ACKNOWLEDGEMENTS

H.G.L. acknowledges support from the Remick Graduate Fellowship in Engineering administered by the University of Notre Dame. A.B. acknowledges support from the University of Strathclyde Research Studentship Scheme, the University of Strathclyde and CMAC's Industry Membership for funding and supporting this work.

This effort was funded by the U.S. Department of



Energy's Process Optimization and Modeling for Minerals Sustainability (PrOMMiS) Initiative, supported by the Office of Fossil Energy and Carbon Management's Office of Resource Sustainability.

Disclaimer: This effort was funded by the U.S. Department of Energy, National Energy Technology Laboratory an agency of the United States Government, through a support contract. Neither the United States Government nor any agency thereof, nor any of its employees, nor the support contractor, nor any of their employees, makes any warranty, expressor implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and

opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## REFERENCES

1. Tao F, Q Qi. Make more digital twins. *Nature* 490-491 (2019)
2. Kusiak A. Smart manufacturing must embrace big data. *Nature* 23-25 (2017)
3. Pyliandis C, S Osinga, IN Athanasiadis. Introducing digital twins to agriculture. *Comput Electron Agric* 105942 (2021)
4. Björnsson B, et al. Digital twins to personalize medicine. *Genome Med.* 1-4 (2020)
5. Franceschini G, S Macchietto. Model-based design of experiments for parameter precision. *Chem.* 4846-4872 (2008)
6. Agarwal AK, ML Brisk. Sequential experimental design for precise parameter estimation. 1. Use of

**Table 4:** Eigendecomposition of the FIM for the  $\beta=0.3$  ( $^{\circ}\text{C min}^{-1}$ ) candidate experiment.

	$\mathbf{v}_1$	$\mathbf{v}_2$	$\mathbf{v}_3$	$\mathbf{v}_4$	$\mathbf{v}_5$	$\mathbf{v}_6$
$k_{b_1}$	-0.027	-0.015	-0.418	0.140	-0.043	0.896
$b_1$	0.045	0.041	0.800	-0.389	0.101	0.441
$k_{b_2}$	-0.647	-0.735	-0.007	-0.137	0.147	-0.007
$b_2$	0.122	0.144	-0.353	-0.550	0.731	-0.038
$k_g$	-0.724	0.638	-0.053	-0.188	-0.173	-0.015
$g$	-0.198	0.171	0.239	0.687	0.634	0.031
$\lambda_i$	$6.124 \times 10^4$	$3.920 \times 10^3$	$9.993 \times 10^0$	$5.506 \times 10^{-2}$	$9.359 \times 10^{-3}$	$3.516 \times 10^{-5}$

**Table 5:** Eigendecomposition of the FIM for the  $\beta=0.2 + \beta=0.3$  ( $^{\circ}\text{C min}^{-1}$ ) candidate experiments.

	$\mathbf{v}_1$	$\mathbf{v}_2$	$\mathbf{v}_3$	$\mathbf{v}_4$	$\mathbf{v}_5$	$\mathbf{v}_6$
$k_{b_1}$	-0.026	0.013	0.355	0.215	-0.053	0.908
$b_1$	0.045	-0.037	-0.706	-0.341	0.481	0.387
$k_{b_2}$	-0.635	0.744	-0.005	0.075	0.190	-0.034
$b_2$	0.124	-0.153	0.262	0.457	0.812	-0.158
$k_g$	-0.734	-0.622	0.139	-0.223	0.074	-0.009
$g$	-0.200	-0.186	-0.536	0.757	-0.255	0.012
$\lambda_i$	$1.244 \times 10^5$	$7.148 \times 10^3$	$2.96 \times 10^1$	$7.506 \times 10^0$	$1.576 \times 10^0$	$3.185 \times 10^{-3}$

**Table 6:** Eigendecomposition of the FIM for the  $\beta=0.1 + \beta=0.2 + \beta=0.3$  ( $^{\circ}\text{C min}^{-1}$ ) candidate experiments.

	$\mathbf{v}_1$	$\mathbf{v}_2$	$\mathbf{v}_3$	$\mathbf{v}_4$	$\mathbf{v}_5$	$\mathbf{v}_6$
$k_{b_1}$	-0.028	0.015	0.143	0.332	0.191	0.912
$b_1$	0.051	-0.044	-0.462	-0.778	-0.157	0.391
$k_{b_2}$	-0.599	0.771	-0.087	-0.081	0.180	-0.026
$b_2$	0.122	-0.161	-0.363	-0.010	0.901	-0.122
$k_g$	-0.758	-0.579	-0.243	0.144	-0.098	-0.007
$g$	-0.218	-0.208	0.754	-0.508	0.291	0.003
$\lambda_i$	$1.810 \times 10^5$	$1.283 \times 10^4$	$8.377 \times 10^1$	$5.984 \times 10^1$	$1.246 \times 10^1$	$5.839 \times 10^{-2}$

- reparameterization. *Ind. Eng. Chem. Process. Des. Dev.* 203-207 (1985)
7. Bazil JN, GT Buzzard, AE Rundell. A global parallel model based design of experiments method to minimize model output uncertainty. *Bull. Math. Biol.* 688-716 (2012)
  8. Bard Y, Nonlinear Parameter Estimation. Academic Press (1974)
  9. Asprey SP, S Macchietto. Statistical tools for optimal dynamic model building. *Comput Chem Eng.* 1261-1267 (2000)
  10. McLean KAP, KB McAuley. Mathematical modelling of chemical processes—obtaining the best model predictions and parameter estimates using identifiability and estimability procedures. *Can J Chem Eng.* 351-366 (2012)
  11. Pramanik BK, et al. A critical review of membrane crystallization for the purification of water and recovery of minerals. *Rev. Environ. Sci. Biotechnol.* 411-439 (2016)
  12. Moerman W, et al. Phosphate removal in agro-industry: pilot-and full-scale operation considerations of struvite crystallization. *Water Res.* 1887-1892 (2009)
  13. Hartel RW. Advances in food crystallization. *Annu. Rev. food. Sci. technol.* 277-292 (2013)
  14. Tung, HH, et al. Crystallization of Organic Compounds: An Industrial Perspective. John Wiley & Sons (2023).
  15. Chen J, et al. Pharmaceutical crystallization. *Cryst. Growth Des.* 887-895 (2011)
  16. Peng, L, et al. A comprehensive review of phosphorus recovery from wastewater by crystallization processes. *Chemosphere* 768-781 (2018)
  17. Kiani H, DW Sun. Water crystallization and its importance to freezing of foods: A review. *Trends Food Sci. Technol.* 407-426 (2011)
  18. Carpenter K, WML Wood. Industrial crystallization for fine chemicals. *Adv. Powder Technol.* 657-672 (2004)
  19. Variankaval N, AS Cote, MF Doherty. From form to function: Crystallization of active pharmaceutical ingredients. *AIChE J.* 1682-1688 (2008)
  20. Ulrich J, P Frohberg. Problems, potentials and future of industrial crystallization. *Front Chem. Sci. Eng.* 1-8 (2013)
  21. Meldrum FC, C O'Shaughnessy. Crystallization in confinement. *Adv. Mater.* 2001068 (2020)
  22. Brown, CJ, et al. Enabling precision manufacturing of active pharmaceutical ingredients: workflow for seeded cooling continuous crystallizations. *Mol. Syst. Des. Eng.* 518-549 (2018)
  23. Chung, SH, DL Ma, RD Braatz. Optimal model-based experimental design in batch crystallization. *Chemometr Intell Lab Syst.* 83-90 (2000)
  24. Jha SK, S Karthika, TK Radhakrishnan. Modelling and control of crystallization process. *Resour. Effic. Technol.* 94-100 (2017)
  25. Randolph A, Theory of Particulate Processes: Analysis and Techniques of Continuous Crystallization. Elsevier (2012).
  26. Nagy ZK, et al. Recent advances in the monitoring, modeling and control of crystallization systems. *Chem Eng Res Des.* 1903-1922 (2013)
  27. He Y, et al. Seeding techniques and optimization of solution crystallization processes. *Org Process Res Dev.* 1839-1849 (2020)
  28. Li H, Y Kawajiri, MA Grover, RW Rousseau. Modeling of nucleation and growth kinetics for unseeded batch cooling crystallization. *Ind. Eng. Chem. Res.* 4060-4073 (2017)
  29. Cashmore A, et al. Rapid assessment of crystal nucleation and growth kinetics: comparison of seeded and unseeded experiments. *Cryst. Growth Des.* (2023)
  30. Mullin JW, Crystallization. Elsevier (2001)
  31. Wang J, AW Dowling. Pyomo.DOE: An open-source package for model-based design of experiments in Python. *AIChE J.* e17813 (2022)
  32. Hulburt HM, S Katz. Some problems in particle technology: a statistical mechanical formulation. *Chem Eng Sci* 555-574 (1964)
  33. Omar HM, S Rohani. Crystal population balance formulation and solution methods: a review. *Cryst. Growth Des.* 4028-4041 (2017)
  34. Yazdanpanah N, ZK Nagy. The Handbook of Continuous Crystallization. The Royal Society of Chemistry (2020)
  35. Szilagyi B, et al. Cross-pharma collaboration for the development of a simulation tool for the model-based digital design of pharmaceutical crystallization processes (CrySiV). *Cryst. Growth Des.* 6448-6464 (2021)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Machine Learning-Aided Process Design for Microwave-Assisted Ammonia Production

Md Abdullah Al Masud, Alazar Araia, Yuxin Wang, Jianli Hu, Yuhe Tian\*

West Virginia University, Department of Chemical and Biomedical Engineering, Morgantown, West Virginia, USA

\* Corresponding Author: [yuhe.tian@mail.wvu.edu](mailto:yuhe.tian@mail.wvu.edu).

## ABSTRACT

Machine learning (ML) has become a powerful tool to analyze complex relationships between multiple variables and to unravel valuable information from big datasets. However, an open research question lies in how ML can accelerate the design and optimization of processes in the early experimental development stages with limited data. In this work, we investigate the ML-aided process design of a microwave reactor for ammonia production with exceedingly little experimental data. We propose an integrated approach of synthetic minority oversampling technique (SMOTE) regression combined with neural networks to quantitatively design and optimize the microwave reactor. To address the limited data challenge, SMOTE is applied to generate synthetic data based on experimental data at different reaction conditions. Neural network has been demonstrated to effectively capture the nonlinear relationships between input features and target outputs. The softplus activation function is used for a smoother prediction compared to the Rectified Linear Unit activation function. Ammonia concentration is predicted using pressure, temperature, feed flow rate, and feed composition ratio as input variables. For point-wise prediction based on discrete operating conditions, the proposed SMOTE integrated neural network approach outperforms with 96.1% accuracy compared to neural networks (without SMOTE), support vector regression, and linear regression. The multi-variate prediction trends are also validated which are critical for design optimization.

**Keywords:** Process Design, Process Intensification, Machine Learning, Neural Networks, Ammonia Production

## INTRODUCTION

Machine Learning (ML) offers the capability to surpass the constraints of mechanistic modeling by enabling the learning of complex relationships between process variables and the target outputs, offering cost-effective model development, and proving advantageous for optimization [1-2]. To reliably transform the data into valuable predictions, ML aims to acquire rules and patterns from a sufficiently large number of samples [3-4]. However, in certain cases, data availability may be intrinsically limited such as during the early development stage of novel experimental technologies. The scarcity of the datasets may hinder the accuracy of ML predictions. The prediction error typically follows a consistent power-exponential decline as the dataset size increases [5]. To give a more intuitive idea, doubling the number of training samples can result in approximately 20% reduction in

prediction error [6]. As such, a key research challenge is how ML can successfully aid process design and optimization at early-stage process developments using sparse experimental data.

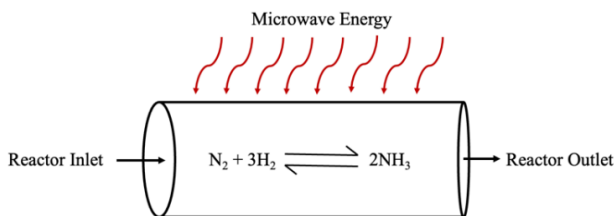
To address this challenge, several approaches have been developed in recent research work which can be categorized as: (i) Sampling, encompassing both over and under sampling techniques [7,8], (ii) Cost sensitivity, which involves modifying the cost function to prompt models to prioritize minority samples [9], (iii) Adversarial network generation, which can produce spurious data [10]. However, these methods show inefficacy in handling continuous target output prediction in the regression context which is essential for the data-driven modeling of chemical process systems.

In this work, we present a neural network-based method integrating synthetic minority oversampling technique (SMOTE). The proposed strategy is applied to

the design of a microwave-assisted ammonia production process with little experimental data. The remainder of this paper is organized as follows: Section 2 introduces the motivating case study of a microwave-assisted reactor for ammonia production. Section 3 details the SMOTE-integrated neural network method. Section 4 showcases this method for the data-driven modeling and optimization of the microwave reactor. Section 5 presents concluding remarks and ongoing work.

## CASE STUDY: MICROWAVE-ASSISTED AMMONIA PRODUCTION

The use of microwaves in chemical synthesis has sparked considerable interest in recent years. Compared with traditional thermal heating, microwave offers the capability of direct and selective volumetric heating which can result in higher reaction selectivity, shorter reaction times, and milder reaction conditions [11,12]. One of the key applications of microwave reactor is ammonia production under ambient/moderate pressure and moderate temperature [13]. This provides a promising technological alternative to the current ammonia synthesis route via the highly energy intensive Haber-Bosch process.



**Figure 1.** Schematic of a cylindrical microwave reactor.

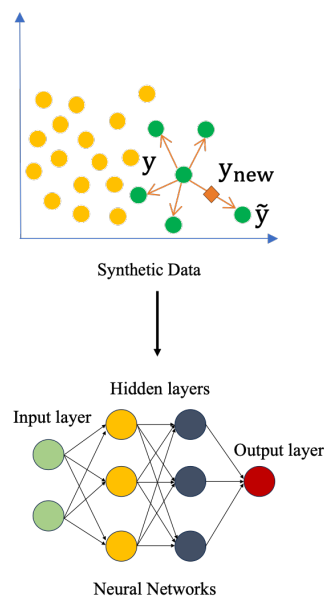
A schematic of the microwave-assisted ammonia reactor is shown in Fig. 1 which is adapted as the case study of this work [14]. Ru-based catalyst is used which has been demonstrated to exhibit high activity for ammonia synthesis under moderate reaction conditions, particularly when alkali metal Cs is used as a promoter. A total of 46 data points is collected from the microwave reactor at different operating conditions. Table 1 gives examples of such measured data. Our research objective is to predict the optimal reaction conditions for ammonia synthesis based on these experimental data.

**Table 1:** Experimental data – an indicative list.

T (°C)	F (ml/h·g <sub>cat</sub> )	NH <sub>3</sub> %	F (ml/h·g <sub>cat</sub> )	NH <sub>3</sub> %
280		0.56		0.27
320		1.57		1.02
340	6000	1.63	15000	1.21
360		1.55		1.25
400		1.28		1.06

## THE PROPOSED METHODOLOGY

As depicted in Fig. 2, a neural network approach is developed leveraging SMOTE and softplus activation function. Neural network is employed to capture the non-linear input-output relationship. SMOTE regression [6,7] is applied to overcome the challenge of limited data availability, which can generate new synthetic data to be used for ML training together with the original experimental data. Soft plus activation function is applied to obtain smooth and differentiable output predictions. In what follows, we briefly introduce the key components of the proposed methodology.



**Figure 2.** SMOTE integrated Neural Networks.

### SMOTE Regression

SMOTE [7] is commonly employed for data simulation to create distribution-dependent neighbor samples for the minority class. Unlike the random oversampling algorithms, SMOTE can dynamically generate the necessary number of samples for the minority class. The use of SMOTE aims to explore new samples that closely resemble the original data distribution. This is instrumental for generating experimentally relevant data with similarities to the existing dataset. In the data generation algorithm of SMOTE, several samples are randomly chosen from the  $k$  neighbors for each sample  $y$ , in the minority class. A new sample is constructed for each randomly selected neighbor from the original sample, as shown by Eq. 1. The new synthetic data points are then merged with the original experimental data to be fed to neural network training.

$$y_{new} = y + rand(0,1) \times (y - \tilde{y}) \quad (1)$$

## Neural Network with Soft plus Activation

Neural network models are comprised of nodes, where each node has the capacity to host several neurons. The neural network architecture used in this work comprises of one input layer, two hidden layers, and one output layer. Each neuron is associated with weights and biases which get updated to minimize the difference between the true labels and predicted labels of the input datasets. The output of a neuron becomes input for the consecutive neurons. The output is expressed by Eq. 2.

$$y^t = \text{softplus}(w^t x^t + b^t) \quad (2)$$

where  $w^t$  is the weight vector,  $x^t$  is the vector for input data,  $b^t$  is the bias.

The role of activation function in a neural network architecture is that it transforms the linear combination of summed weights and biases into a nonlinear output. Activation function normalizes the data that imposes a restriction to convert the output of a neuron into a specific bound. The Rectified Linear Unit (ReLU) and Softplus activation functions are compared as shown in Fig. 3. ReLU activation function is a non-differentiable function which may provide non-smooth predictions compared to Softplus, i.e.  $f(x) = \log(1 + e^x)$ . A major advantage of using Softplus is that it does not suffer from the “dying ReLU” problem which refers to the situation when some of the neurons become inactive, and they cannot update their weights and biases resulting in non-smooth predictions.

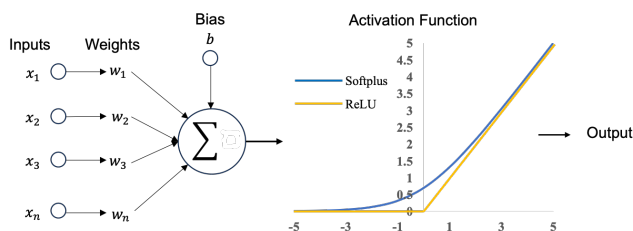


Figure 3: Comparison of activation functions.

## MACHINE LEARNING-AIDED MICROWAVE DRIVEN REACTION DESIGN

In this section, we apply the SMOTE-integrated neural network approach to model the microwave-assisted ammonia reaction. Specifically, we will investigate:

- Point-wise prediction for  $\text{NH}_3$  concentration based on discrete operating conditions (i.e., temperature, pressure, inlet  $\text{H}_2$  to  $\text{N}_2$  ratio, inlet gas flow rate).
- Continuous prediction to capture the variation trend of  $\text{NH}_3$  concentration under varying operating conditions.
- Reaction design optimization for maximum  $\text{NH}_3$  concentration.

## Point-wise Prediction

Two hidden layers are used for the neural networks with 128 and 64 nodes, respectively. 40,589 synthetic data are generated via SMOTE. The training and testing data are separated by an 80-20% split. The adam optimizer is applied utilizing the mean squared error as the loss function. The number of epochs is set as 100 with a learning rate of 0.0001. The average discrete prediction accuracy of SMOTE-integrated neural network is 96.10% as shown in Table. 1. This accuracy is superior to other regression methods such as neural networks (without SMOTE, 95.10%), support vector regression (88.70%), and linear regression (86.30%).

Table 2: SMOTE-integrated NN pointwise prediction.

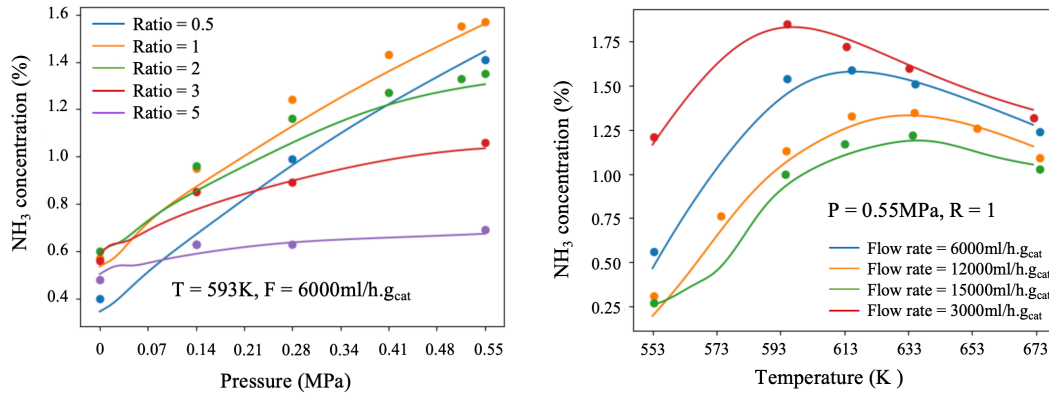
Actual $\text{NH}_3$ Concentration	Reaction Conditions [T(°C), P(psig), R, F (ml/h <sub>cat</sub> )]	Prediction Accuracy (%)
0.63	320, 20, 0.50, 6000	99.60
0.56	320, 0, 3, 6000	95.50
1.55	320, 75, 1, 6000	99.80
1.35	320, 80, 2, 6000	92.92
0.60	320, 0, 2, 6000	95.80
0.89	320, 40, 3, 6000	98.68
1.06	320, 80, 3, 6000	95.14
1.57	320, 80, 1, 6000	92.31
...	...	...
		Avg. = 96.10

## Continuous Prediction

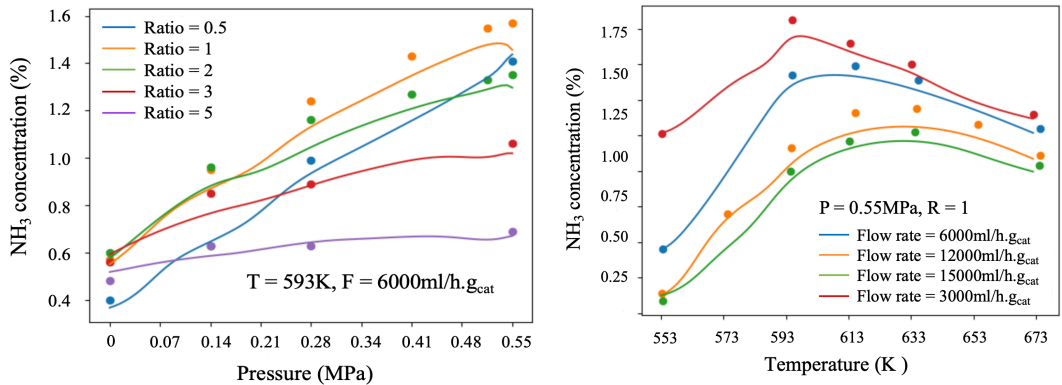
Herein, the primary goal is to verify that the resulting data-driven model can also capture the continuous variational trend of ammonia concentration versus operating conditions. We first develop a SMOTE-integrated neural network model for two inputs and one output. For example,  $\text{NH}_3$  concentration is computed from the two-input one-output data-driven model at varying pressures and inlet  $\text{H}_2$  to  $\text{N}_2$  ratios while temperature and inlet gas flow rate are kept constant. Two separate models are developed for the prediction of  $\text{NH}_3$  concentration, respectively using 11,340 and 4,180 synthetic data generated by SMOTE. The total number of synthetic data is reduced to avoid overfitting. The prediction results are validated against the original experimental data as shown in Fig. 4, in which the markers represent the original experimental data, and the solid lines depict the continuous prediction using the above trained SMOTE-integrated NN model.

On this basis, we proceed to build a single data-driven model considering all the four input variables (i.e., temperature, pressure,  $\text{H}_2$  to  $\text{N}_2$  ratio, and feed flow rate).

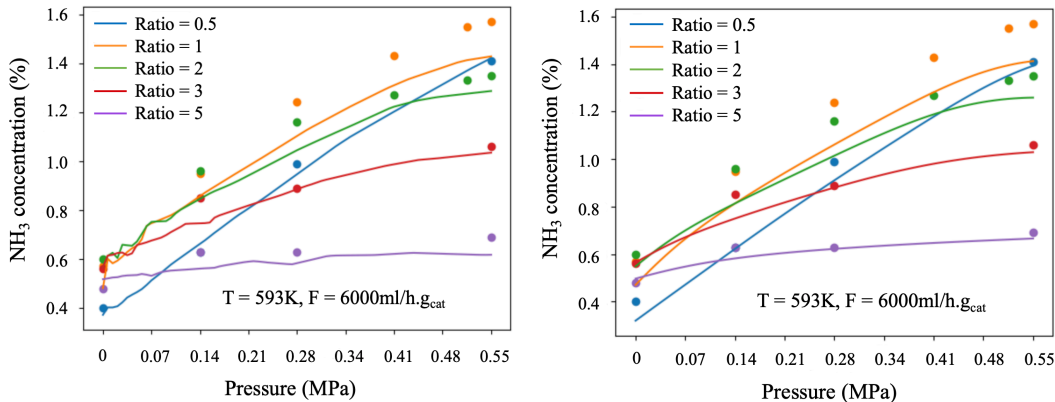




**Figure 4:** Continuous prediction with two-input one-output data-driven model. (Markers: Original experimental data, Solid lines: SMOTE integrated NN prediction)



**Figure 5:** Continuous prediction with four-input one-output data-driven model.



**Figure 6:** Role of activation function. (Left) ReLU. (Right) Soft plus.

As the input dimensions increase, the input-output relationship becomes more nonlinear and complex. Namely, the resulting data-driven model should simultaneously capture the nonlinear relationship between ammonia concentration versus temperature, pressure, inlet  $H_2$  to  $N_2$  ratio, and inlet gas flow rate. 23,000 synthetic data generated via SMOTE are used to train the four-input one-output neural network. As illustrated in Fig. 5, the

continuous prediction using a higher dimensional data-driven model can also effectively capture the variational trend while featuring larger deviations than the two-input one-output models. The role of activation functions is also investigated as shown in Fig. 6. It can be noted that the use of ReLU activation function may result in non-smooth predictions compared with that of Softplus activation function. This is consistent with the observations

reported in open literature [15].

## Comparison with SVR and NN

The continuous predictions from our proposed methodology were compared to the support vector regression (SVR) and NN (without SMOTE). For SVR, Radial Basis Function (RBF) is identified as the best kernel function for our datasets. As illustrated in Fig. 7, SVR identifies a pseudo-linear multi-variate relationship instead of capturing the correct nonlinear trends. The predictions using NN (without SMOTE) are depicted in Fig. 8. Even if NN offers a point-wise prediction accuracy of 95.10%, it introduces excessive nonlinear directional changes in the continuous prediction. In other words, NN cannot effectively capture the continuous variational trend. If this data-driven model is applied for design optimization, it may fail to identify the correct direction toward optimality.

## Design Optimization

Finally, we performed the design optimization to determine the optimum reaction conditions. A mesh grid of data points is created and fed to the afore-trained SMOTE-integrated neural networks (Fig. 9). The ammonia concentration for each of the input data is obtained from the data-driven model prediction and the index of maximum ammonia concentration is located. The optimum temperature, pressure, H<sub>2</sub> to N<sub>2</sub> ratio, and inlet gas flow rate are identified as 324.37°C, 80psig, 1, and 3000ml/h·g<sub>cat</sub> respectively. The optimum operating conditions obtained from experimental data are at the temperature of 320°C, pressure of 80psig, H<sub>2</sub> to N<sub>2</sub> ratio of 1, and inlet gas flow rate of 3000ml/h·g<sub>cat</sub> which well justifies the validity of data-driven model prediction.

## CONCLUSION

In summary, this work has developed a machine learning-aided method to design and optimize microwave-assisted ammonia reaction conditions with little experimental data. The efficacy of the proposed approach is demonstrated on the point-wise and continuous prediction of ammonia concentration under varying reaction temperature, pressure, H<sub>2</sub> to N<sub>2</sub> ratio, and feed flow rates. Ongoing work is investigating the systems-level analysis leveraging this data-driven microwave reactor model.

## ACKNOWLEDGEMENT

This research project acknowledges the funding resource of West Virginia University and WV Research Challenge Grant (RCG23-009): Metal-embedded carbon-based catalytic membranes for co-production of ammonia and ethylene.

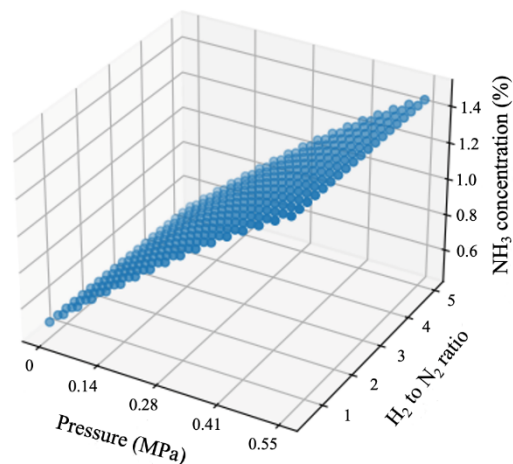


Figure 7: Continuous prediction using SVR.

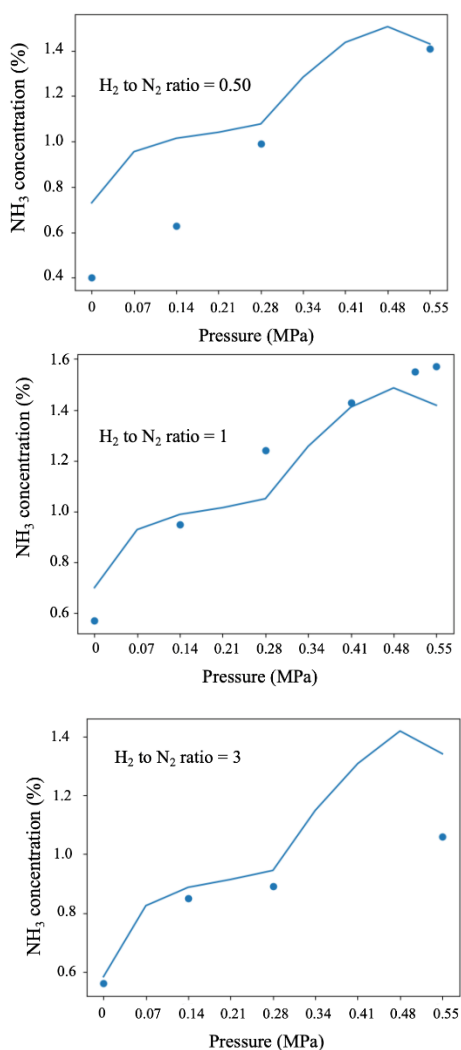
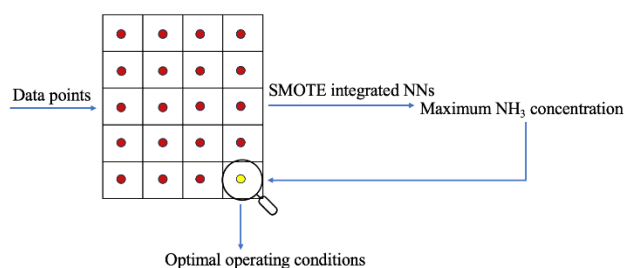


Figure 8: Continuous prediction using NN without SMOTE. (Markers: Original experimental data, Solid lines: NN prediction)



**Figure 9:** Illustration of design optimization.

## REFERENCES

- Schweidtmann, A. M., Esche, E., Fischer, A., Kloft, M., Repke, J. U., Sager, S., & Mitsos, A. (2021). Machine learning in chemical engineering: A perspective. *Chemie Ingenieur Technik*, *93*(12), 2029-2039.
- Lee, J. H., Shin, J., & Realff, M. J. (2018). Machine learning: Overview of the recent progresses and implications for the process systems engineering field. *Computers & Chemical Engineering*, *114*, 111-121.
- Golkarnarenji, G., Naebe, M., Badii, K., Milani, A. S., Jazar, R. N., & Khayyam, H. (2019). A machine learning case study with limited data for prediction of carbon fiber mechanical properties. *Computers in Industry*, *105*, 123-132.
- Zhang, Z., Schott, J. A., Liu, M., Chen, H., Lu, X., Sumpter, B. G., & Dai, S. (2019). Prediction of carbon dioxide adsorption via deep learning. *Angewandte Chemie*, *131*(1), 265-269.
- Faber, F. A., Lindmaa, A., Von Lilienfeld, O. A., & Armiento, R. (2016). Machine learning energies of 2 million elpasolite (ABC<sub>2</sub>D<sub>6</sub>) crystals. *Physical review letters*, *117*(13), 135502.
- Liu, X., Xu, Y., Li, J., Ong, X., Ibrahim, S. A., Buonassisi, T., & Wang, X. (2021). A robust low data solution: dimension prediction of semiconductor nanorods. *Computers & Chemical Engineering*, *150*, 107315.
- SMOTE-RSB\*: a hybrid preprocessing approach based on oversampling and undersampling for high imbalanced data-sets using smote and rough sets theory. *Knowledge and information systems*, *33*, 245-265.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, *16*, 321-357.
- Elkan, C. (2001, August). The foundations of cost-sensitive learning. In *International joint conference on artificial intelligence* (Vol. 17, No. 1, pp. 973-978). Lawrence Erlbaum Associates Ltd.
- Antoniou, A., Storkey, A., & Edwards, H. (2017). Data augmentation generative adversarial networks. *arXiv preprint arXiv:1711.04340*.
- Horikoshi, S., & Serpone, N. (2015). General Introduction to Microwave Chemistry. *Microwaves in Catalysis: Methodology and Applications*, 1-28.
- Hunt, J., Ferrari, A., Lita, A., Crosswhite, M., Ashley, B., & Stiegman, A. E. (2013). Microwave-specific enhancement of the carbon-carbon dioxide (Boudouard) reaction. *The Journal of Physical Chemistry C*, *117*(51), 26871-26880.
- Wang, Y., Khan, T. S., Wildfire, C., Shekhawat, D., & Hu, J. (2021). Microwave-enhanced catalytic ammonia synthesis under moderate pressure and temperature. *Catalysis Communications*, *159*, 106344.
- Hu, J., Wildfire, C., Stiegman, A. E., Dagle, R. A., Shekhawat, D., Abdelsayed, V., & Wang, Y. (2020). Microwave-driven heterogeneous catalysis for activation of dinitrogen to ammonia under atmospheric pressure. *Chemical Engineering Journal*, *397*, 125388.
- Rittig, J. G., Felton, K. C., Lapkin, A. A., & Mitsos, A. (2023). Gibbs-Duhem-informed neural networks for binary activity coefficient prediction. *Digital Discovery*, *2*, 1752-1767.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>.



# Fast, Accurate, and Robust Fault Detection and Diagnosis of Industrial Processes

Alireza Miraliakbar<sup>a</sup>, Zheyu Jiang<sup>a,\*</sup>

<sup>a</sup> School of Chemical Engineering, Oklahoma State University, Stillwater, Oklahoma, USA, 74078

\* Corresponding Author: [zheyu.jian@okstate.edu](mailto:zheyu.jian@okstate.edu)

---

## ABSTRACT

Modern industrial processes are continuously monitored by a large number of sensors. Despite having access to large volumes of historical and online sensor data, industrial practitioners still face challenges in the era of Industry 4.0 in effectively utilizing them to perform online process monitoring and fast fault detection and diagnosis. To target these challenges, in this work, we present a novel framework named “FARM” for Fast, Accurate, and Robust online process Monitoring. FARM is a holistic monitoring framework that integrates (a) advanced multivariate statistical process control (SPC) for fast anomaly detection of nonparametric, heterogeneous data streams, and (b) modified support vector machine (SVM) for accurate and robust fault classification. Unlike existing general-purpose process monitoring frameworks, FARM’s unique hierarchical architecture decomposes process monitoring into two fault detection and diagnosis, each of which is conducted by targeted algorithms. Here, we test and validate the performance of our FARM monitoring framework on Tennessee Eastman Process (TEP) benchmark dataset. We show that SPC achieves faster fault detection speed at a lower false alarm rate compared to state-of-the-art benchmark fault detection methods. In terms of fault classification diagnosis, we show that our modified SVM algorithm successfully classifies 17 out of 20 of the fault scenarios present in the TEP dataset. Compared with the results of standard SVM trained directly on the original dataset, our modified SVM improves the fault classification accuracy significantly.

---

**Keywords:** Fault Detection and Diagnosis, Process Monitoring, Statistical Process Control, Riemannian Manifold, Support Vector Machine

## INTRODUCTION

Safe and efficient operation of an industrial plant depends on effective, continuous process monitoring (e.g., fault detection and diagnosis), which is enabled by advanced sensory systems that continuously generate streams of data to dictate the state of the plant. Despite having access to large volumes of historical and online sensor data, challenges remain in how these data could be used for effective online process monitoring. Existing techniques for process monitoring are inadequate because (a) fault scenarios in industrial systems and plants are complex, (b) sensors continuously produce massive arrays of big data streams that are often nonparametric (i.e., data streams may not follow any specific distribution) and heterogeneous (i.e., data streams may not

follow the same distribution), and (c) there is an intrinsic trade-off between fault detection time and diagnostic accuracy.

To address this need, several process monitoring solutions have been developed over the past decades. Among them, dimensionality reduction techniques, such as principal component analysis (PCA), partial least squares (PLS) regression, as well as their different variations, are the most popular ones in the literature [1–3]. Dimensionality reduction techniques assume that the statistics characterizing the in-control profiles also span the subspace where out-of-control states (faults) lie in [4]. However, this assumption is generally invalid for industrial process monitoring as the process dynamics are quite complex and out-of-control states cannot be fully enumerated a priori. Also, plant operators often find it

difficult to interpret the results from PCA/PLS-based methods because the features are in the reduced space and do not have one-to-one mapping to the original sensor data sources. In addition, monitoring only the most significant subset of features often causes significant errors, as the fault may not be noticeable in the selected features. Lastly, dimensionality reduction techniques have no statistical guarantee on false alarm rate, making them unreliable for actual plant monitoring which requires false alarm to be low and controlled (e.g.,  $\leq 0.0027$ , the classic three-sigma limit) due to the significant money loss and safety issues of unplanned unit shutdown.

More recently, various machine learning (ML) tools such as support vector machine, decision tree, and deep neural network, have also been proposed and applied to process monitoring [5–8]. Nevertheless, existing ML methods still face problems such as overfitting and poor predictive accuracy. For example, while most published ML algorithms perform well during training and validation, their fault detection accuracies deteriorate and rarely exceed 90–95% in test sets. Considering the severe consequences in case of fault detection failure, such predictive accuracy is unacceptable. Furthermore, ML methods do not scale well with rare or new fault scenarios due to the lack of sufficient training data.

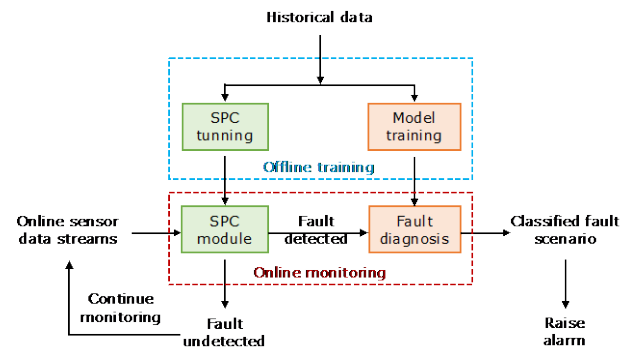
To target these challenges, in this work, we present a novel industrial process monitoring tool, which we named it as “FARM”, for fast, accurate, and robust online fault detection and diagnosis. FARM is a holistic monitoring framework that integrates (a) advanced multivariate statistical process control (SPC) for fast anomaly detection of nonparametric, heterogeneous data streams, and (b) a modified support vector machine (SVM) for accurate and robust fault classification. Unlike existing general-purpose process monitoring frameworks, FARM’s unique hierarchical architecture (see Figure 1) decomposes process monitoring into two fault detection and diagnosis, each of which is conducted by targeted algorithms. Only if a process anomaly is detected will the online data be sent to the fault classification/diagnosis module for accurate fault classification. Such hierarchical architecture successfully bypasses the intrinsic trade-off between fault detection speed and accuracy that is present in existing monitoring tools. Furthermore, using FARM, plant operators can choose a user-specified false alarm rate based on their expert knowledge of the process.

## STRUCTURE AND WORKFLOW OF FARM

As mentioned earlier, FARM consists of two distinct yet interconnected modules. The first module performs fault detection by adopting the state-of-the-art quantile-based non-parametric SPC proposed by Ye and Liu [9]. Quantile-based nonparametric SPC can detect any

process mean shift or anomaly from heterogeneous high-dimensional sensor data streams as early as possible while maintaining a pre-specified incontrol average run length. Inspired by the work of Smith et al. [10], the second module conducts fault classification through a modified SVM model. Both modules are connected as shown in Figure 1. FARM’s workflow contains two steps: (1) offline training with historical data, followed by (2) online monitoring of real-time sensor data streams. During offline training, the parameters of the SPC module to be used for online monitoring are obtained using the historical in-control data. Also, the modified SVM module is trained by treating the faulty data’s covariance matrices as features and the corresponding faulty scenario as labels.

Once offline training of FARM is complete, online sensor measurements will continuously be sent to FARM for simultaneous fault detection and diagnosis. First, they are monitored by the SPC module to detect any process anomaly in real time. Only if a process anomaly is detected will the online data be sent to the fault diagnosis module for accurate fault classification. Unlike general-purpose process monitoring frameworks, FARM’s hierarchical architecture decomposes process monitoring tasks into two subtasks (fault detection and diagnosis), each of which is accomplished by specialized techniques. This allows fast, accurate, and robust fault detection and diagnosis to be simultaneously accomplished by FARM.



**Figure 1.** FARM’s hierarchical structure consisting of fault detection and diagnosis modules.

### Fault Detection

The backbone of FARM’s fault detection module is the quantile-based non-parametric SPC algorithm proposed by Ye and Liu [9]. Jiang modified the original quantile-based SPC formulation of Ye and Liu [9] to monitor fully observable data streams [11]. Here, a brief description of the modified SPC formulation is presented. In offline training, the sensor measurements in each of the  $M$  historical in-control data streams  $X_j$  ( $j = 1, 2, \dots, M$ ) are sorted in ascending order and partitioned into  $d$  number of quantiles  $I_{j,1}, \dots, I_{j,d}$  defined as:



$$I_{j,1} = (-\infty, q_{j,1}], I_{j,2} = (q_{j,1}, q_{j,2}], \dots, I_{j,d} = (q_{j,d-1}, +\infty) \quad (1)$$

For each  $q_{j,i}$ , two intervals called positive and negative cumulative intervals are defined as:

$$CI_{j,i}^+ = [q_{j,i}, +\infty) \quad \text{and} \quad CI_{j,i}^- = (-\infty, q_{j,i}] \quad (2)$$

for every  $i = 1, \dots, d-1$  and  $j = 1, 2, \dots, M$ . With these positive/negative cumulative intervals identified from historical in-control data, one can detect anomalies in real time by detecting any upward/downward mean shift of online sensor data streams. To do this, for an online sensor data stream  $X_j(t)$  where  $t$  stands for time, we define a binary variable  $A_{j,i \in [1, \dots, d-1]}^+$  and  $A_{j,i \in [1, \dots, d-1]}^-$  to indicate which positive and negative cumulative interval  $X_j(t)$  lies in at time  $t$ , respectively:

$$A_{j,i \in [1, 2, \dots, d-1]}^+ = \begin{cases} 1 & \text{if } X_j(t) \in CI_{j,i}^+ \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

$$A_{j,i \in [1, 2, \dots, d-1]}^- = \begin{cases} 1 & \text{if } X_j(t) \in CI_{j,i}^- \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

With this, we obtain two vectors  $\mathbf{A}_j^+(t)$  and  $\mathbf{A}_j^-(t)$  as:

$$\mathbf{A}_j^+(t) = [A_{j,1}^+, A_{j,2}^+, \dots, A_{j,d-1}^+], \quad (5)$$

$$\mathbf{A}_j^-(t) = [A_{j,1}^-, A_{j,2}^-, \dots, A_{j,d-1}^-]. \quad (6)$$

One can show that  $\mathbb{E}[\mathbf{A}_j^+(t)] = [1 - \frac{1}{d}, 1 - \frac{2}{d}, \dots, 1 - \frac{d-1}{d}]$  and  $\mathbb{E}[\mathbf{A}_j^-(t)] = [\frac{1}{d}, \frac{2}{d}, \dots, \frac{d-1}{d}]$  for  $j = 1, \dots, M$  and  $i = 1, \dots, d$ . Therefore, by defining  $\mathbf{A}_j^+(t)$  and  $\mathbf{A}_j^-(t)$ , the idea is to convert the task of detecting any mean shift in the distribution of  $X_j(t)$  with respect to the distribution of historical in-control data into an equivalent task of detecting the upward (resp. downward) mean shift in the distribution of  $A_{j,i}^+$  (resp.  $A_{j,i}^-$ ) with respect to  $\mathbb{E}[A_{j,i}^+]$  (resp.  $\mathbb{E}[A_{j,i}^-]$ ). This transformation presents at least two major advantages. First, it has been shown that  $A_{j,i}^+$  (resp.  $A_{j,i}^-$ ) is more sensitive to upward (resp. downward) mean shifts than the original data streams themselves [9], thus allowing faster fault detection. And second, it allows nonparametric, heterogeneous data streams to be successfully monitored for the first time.

Quantile-based SPC implements the multivariate cumulative sum (CUSUM) procedure first proposed by Qiu and Hawkins [12, 13] to monitor multivariate big data streams of  $\mathbf{A}_j^+(t)$  and  $\mathbf{A}_j^-(t)$  for  $j = 1, \dots, M$ . This is achieved by defining  $C_j^+(t)$  and  $C_j^-(t)$  as:

$$C_j^\pm(t) = \left[ \left( \mathbf{S}_j^{\pm, \text{obs}}(t-1) + \mathbf{A}_j^\pm(t) \right) - \left( \mathbf{S}_j^{\pm, \text{exp}}(t-1) + \mathbb{E}[\mathbf{A}_j^\pm(t)] \right) \right]^T \cdot \left( \text{diag} \left( \mathbf{S}_j^{\pm, \text{exp}}(t-1) + \mathbb{E}[\mathbf{A}_j^\pm(t)] \right)^{-1} \cdot \left[ \left( \mathbf{S}_j^{\pm, \text{obs}}(t-1) + \mathbf{A}_j^\pm(t) \right) - \left( \mathbf{S}_j^{\pm, \text{exp}}(t-1) + \mathbb{E}[\mathbf{A}_j^\pm(t)] \right) \right] \right) \quad (7)$$

In Equation (7),  $\mathbf{S}_j^{\pm, \text{obs}}(t)$  and  $\mathbf{S}_j^{\pm, \text{exp}}(t)$  are four vectors of size  $d-1$  that are the CUSUM statistics initiated at  $\mathbf{S}_j^{\pm, \text{obs}}(t=0) = \mathbf{S}_j^{\pm, \text{exp}}(t=0) = 0$ :

$$\begin{cases} \mathbf{S}_j^{\pm, \text{obs}}(t) = 0, \mathbf{S}_j^{\pm, \text{exp}}(t) = 0, & \text{if } C_j^\pm(t) \leq k \\ \mathbf{S}_j^{\pm, \text{obs}}(t) = \frac{(C_j^\pm(t)-k)}{C_j^\pm(t)} \left( \mathbf{S}_j^{\pm, \text{obs}}(t-1) + \mathbf{A}_j^\pm(t) \right) \\ \mathbf{S}_j^{\pm, \text{exp}}(t) = \frac{(C_j^\pm(t)-k)}{C_j^\pm(t)} \left( \mathbf{S}_j^{\pm, \text{exp}}(t-1) + \mathbb{E}[\mathbf{A}_j^\pm(t)] \right) & \text{if } C_j^\pm(t) > k \end{cases} \quad (8)$$

In Equation (8),  $k$  is an allowance parameter that restarts the CUSUM procedure if no evidence of significant shift is detected after a while [14]. The value of  $k$  is obtained during offline training using historical in-control data. Then, one-sided local statistics  $W_j^+$  and  $W_j^-$  for respectively detecting upward and downward mean shifts of data stream  $j$  can be defined as:

$$W_j^+(t) = \max(0, C_j^+(t) - k), \quad (9)$$

$$W_j^-(t) = \max(0, C_j^-(t) - k). \quad (10)$$

If one wants to detect either upward or downward mean shifts, then a two-sided local statistic  $W_j(t)$  can be defined as the maximum of the two one-sided local statistics:

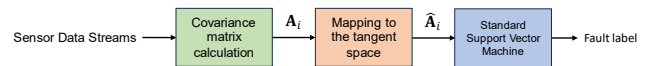
$$\begin{cases} W_j(t=0) = 0, \\ W_j(t > 0) = \max(W_j^+(t), W_j^-(t)). \end{cases} \quad (11)$$

Finally, to determine the stopping time  $T$  for raising the alarm by declaring the process is out-of-control, the top- $r$  approach proposed by Mei [15] is adopted. First, at each time step  $t$ , the values of individual local statistics  $W_j(t)$  for all data streams are ranked from largest to smallest:  $W_{(1)}(t) > \dots > W_{(k)}(t) > \dots > W_{(M)}(t)$ , in which  $W_{(k)}(t)$  corresponds to the  $k^{\text{th}}$  largest local statistic. Next, the top  $r$  of the local statistics at time  $t$  is calculated, and the stopping time  $T$ , also known as the out-of-control run length, is defined as:

$$T = \inf \{ t > 0 : \sum_{(k)=1}^r W_{(k)}(t) \geq h \}, \quad (12)$$

where  $h$  is a threshold value that corresponds to the pre-specified false alarm rate and can be obtained during offline training using historical in-control data. A commonly used  $h$  is obtained based on the false alarm rate of 0.27% (the classic  $3\sigma$  limit).

## Fault Classification and Diagnosis



**Figure 2.** Flowchart of the modified SVM algorithm for improved fault classification.

In this section, we discuss how accurate fault

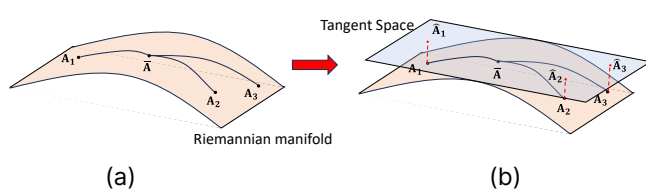
diagnosis can be achieved using a modified SVM module in FARM. Figure 2 illustrates how we modify standard SVM for fault classification by adding a data pre-processing step in the training step. To train the SVM model using the historical sensor data corresponding to different fault scenarios, we first compute the covariance matrix of the historical faulty data streams, followed by training the SVM model over the covariance matrix instead of the original faulty data streams. This modification is inspired by the fact that covariance matrices are symmetric and positive definite, and thus always lie on a Riemannian manifold. It has been recently shown that, by respecting this important geometric insight, one can greatly enhance the accuracy and interpretability of classification, regression, dimensionality reduction algorithms by conducting these computations on the tangent space of the manifold [10]. Inspired by this finding, we map the generated covariance matrices to their tangent space, which intersect the Riemannian manifold where these covariance matrices reside at the geometric mean of the covariance matrices (see Figure 3). This mapping is done through the logarithm operation as:

$$\hat{\mathbf{A}}_i = \log_{\bar{\mathbf{A}}}(\mathbf{A}_i), \quad (13)$$

where  $\mathbf{A}_i$  is the covariance matrix of sensor data streams for dataset  $i$  calculated as:

$$\mathbf{A}_i = \frac{1}{N-1} \mathbf{X}_i \mathbf{X}_i^T, \quad (14)$$

where  $\mathbf{X}_i$  is the original sensor data matrix containing  $M$  number of data streams values over  $N$  time steps.  $\bar{\mathbf{A}}$  is the geometric mean of covariance matrices ( $\mathbf{A}_i$ ), and  $\hat{\mathbf{A}}_i$  is the mapped matrix of matrix  $\mathbf{A}_i$  to the tangent space as shown on Figure 3. The reader is encouraged to read the main reference explaining this mathematical calculation if interested [10].



**Figure 3.** Illustration of (a) a Riemannian manifold and (b) the associated tangent space. The logarithmic map as well as the geodesic between the geometric mean  $\bar{\mathbf{A}}$  and each covariance matrix  $\mathbf{A}_i$  are also shown.

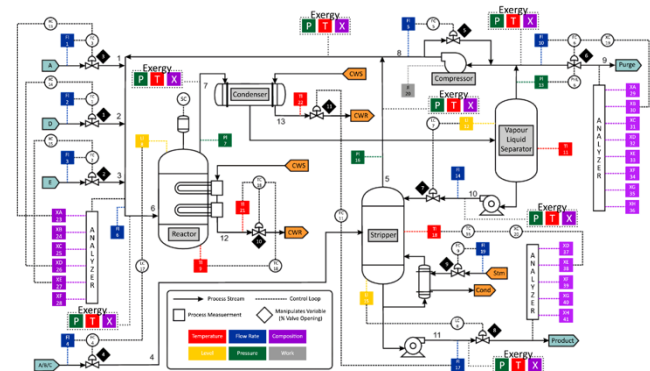
After this data preprocessing step, the mapped covariance matrices are used as input features, whereas the corresponding fault scenarios are used as labels to train a standard SVM model using a radial basis function (RBF) kernel.

During online monitoring stage, real-time sensor

data streams are processed in the fault/anomaly detection module first. Only when a process anomaly is detected will the data streams be sent to the fault classification/diagnosis module. Such an arrangement will further enhance the accuracy and reliability of fault diagnosis module, as the data streams are certain to be faulty. Next, the covariance matrix for the sensor data streams is calculated, mapped to the tangent space of the Riemannian manifold, and used as the input to the trained SVM model to classify its fault label.

## CASE STUDY: TENNESSEE EASTMAN PROCESS

Abstracted from a real chemical process, the Tennessee Eastman Process (TEP) is a nonlinear open-loop unstable process that has been widely used in various computational studies as benchmark case for plant-wide control, process monitoring, and data-driven optimization [16]. As shown in the schematic of Figure 4, the TEP consists of 4 major unit operations: a reactor, a stripping column, a separator, and a product condenser. The process involves the production of two liquid product components G and H from four gaseous reactants A, C, D and E with an additional inert B and a by-product F. The process is continuously monitored by a total of 52 process variables, including 11 manipulated and 41 measured variables.



**Figure 4.** Schematic of TEP (figure extracted from [17]).

### Fault Detection Module Performance

Table 1 lists the comparison results of our SPC module with respect to two benchmark fault detection algorithms, which are PCA-T<sup>2</sup> and SVM [11]. The data used for this study is obtained by the MATLAB graphical user interface (GUI) originally developed by Andersen et al. [18]. Overall, a total of 50 hours (simulation) of normal operation data were generated using this GUI to determine the threshold value  $h$  in Equation (12) and to construct the quantiles  $I_{j,1}, \dots, I_{j,d}$  as well as the cumulative intervals  $CI_{j,i}^{\pm}$ . In addition to normal operation (in-control) data, the

GUI can generate process data for 28 different fault scenarios. Here, we select three faults, namely IDV 2, 3, and 13 (see Table 1 for description), to compare the performance of the SPC algorithm with other benchmarks.

**Table 1.** Description of faults for comparison study of multiple fault detection benchmarks.

Fault #	Description	Fault Type
IDV 2	B composition in stream 4 with A/C ratio constant	Step
IDV 3	D feed temperature in stream 2	Step
IDV 13	Reaction kinetics	Slow drift

Table 2 summarizes the comparison results of fault detection speed and the corresponding false alarm rate of all three monitoring frameworks, quantified by out-of-control run length (i.e., how many additional observations are needed to declare out-of-control status and raise alarm after the actual fault is introduced) for each algorithm. As we can see, among the three monitoring frameworks, quantile-based SPC framework yields the fastest fault detection speed in all three fault scenarios, while maintaining the lowest false alarm rate. Given that a lower false alarm rate generally sacrifices fault detection speed due to more conservative monitoring behavior, the quantile-based SPC framework achieves a win-win situation compared to other benchmark algorithms.

**Table 2.** Fault detection results in terms of out-of-control run length (false alarm rate) for SPC, PCA-T<sup>2</sup> and SVM for TEP dataset [11].

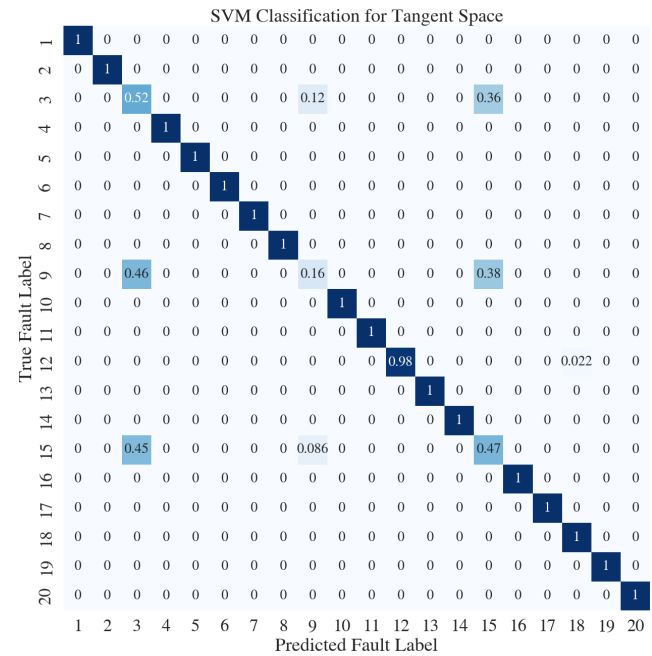
Fault #	SPC	PCA-T <sup>2</sup>	SVM
IDV 2	125 (0.27%)	216 (0.5%)	180 (0.8%)
IDV 3	95 (0.27%)	366 (0.5%)	16815 (83%)
IDV 13	128 (0.27%)	1131 (0.5%)	675 (12.7%)

### Fault Diagnosis Module Performance

For fault diagnosis, we experimented various classification algorithms using the TEP dataset developed by Rieth et al. [19], which consists of 500 simulation cases of normal (in-control) operation as well as 20 fault scenarios. To illustrate, we present three representative models here.

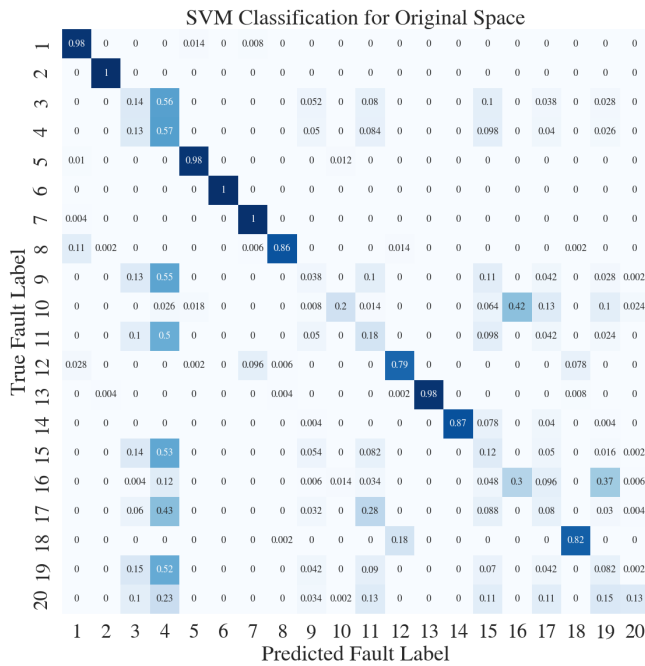
First, we highlight the “best model” obtained by following training procedure illustrated in Figure 2. Figure 5 shows the confusion matrix obtained through 10-fold cross-validation of these 20 faults. Clearly, the modified SVM model demonstrated outstanding classification performance for all faults except for faults IDV 3, 9, and 15. This result outperforms a number of fault diagnosis algorithms in the literature. It is worth noting that faults IDV 3, 9, and 15 correspond to “step change in temperature of

reactor feed D”, “random variation in temperature of reactor feed D”, and “sticking value failure for condenser cooling water valve”, respectively. And these three faults are well-known to be particularly challenging to differentiate due to the close similarity of their dynamic behaviors to the overall process. To tackle this longstanding challenge of successful differentiation of these faults, new, creative methodologies need to be developed.



**Figure 5.** Confusion matrix (after 10-fold cross validation) of fault diagnosis results for our proposed modified SVM model.

As a direct comparison, Figure 6 shows the confusion matrix for the case where standard SVM without any data pre-processing is used for training and validation. It is clear that fault classification accuracy deteriorates significantly in this case.

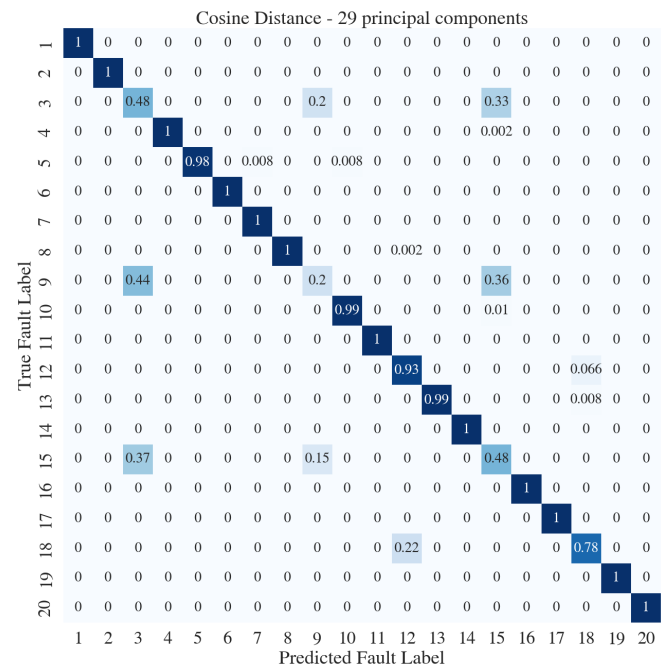


**Figure 6.** Confusion matrix (after 10-fold cross validation) of fault diagnosis results for standard SVM model without the introduced data pre-processing step.

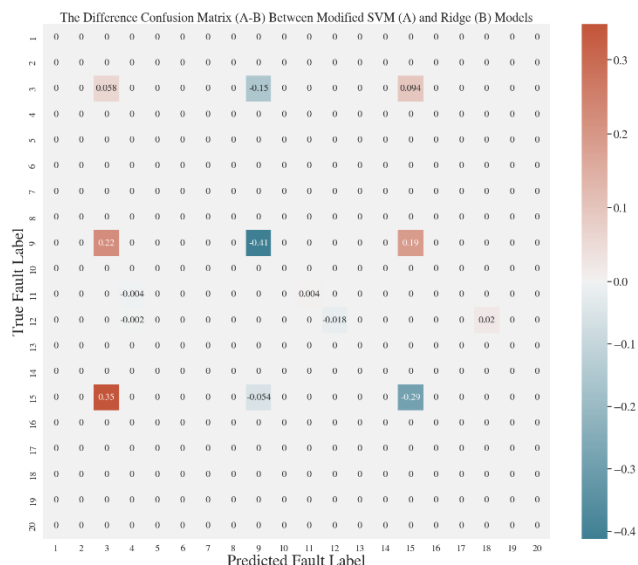
Finally, we present the results for another fault diagnosis algorithm based on principal geodesic analysis (PGA) discussed by Smith et al. [10]. PGA is a counterpart of principal component analysis (PCA) applied on the tangent space of the Riemannian manifold, as it identifies the geodesics that capture the most variance in the data. In other words, in PGA, we simply apply PCA technique to the mapped covariance matrices of faulty data streams for dimensionality reduction. To determine the number of principal geodesics (which are the “principal components” in PGA) needed, we perform sensitivity analysis and identify that 29 principal geodesics are required to capture 99% of the variance in the original dataset containing covariance matrices on the Riemannian manifold. Furthermore, four distance measures, namely Euclidean, Mahalanobis, Manhattan, and Cosine are tested and compared. Clustering is done by assigning a point to its closest cluster based on the distance measure used. We identify that, among these four measures, the cosine distance offers the best fault classification performance. Figure 7 shows the confusion matrix of PGA-cosine approach with the 10-fold cross validation. As we can see, in general, 12 out of the 20 faults can be fully classified, whereas faults IDV 3, 5, 9, 10, 12, 13, 15, and 18 cannot. Although its accuracy is yet to match with the best model, the PGA-Cosine algorithm performs much better than standard SVM without data pre-processing.

Since the modified SVM model showed superior performance over PGA-Cosine method, the confusion

matrix of this method is compared with 10-fold confusion matrix of the ridge classifier model presented by Smith et al. [10], which trained on the mapped covariances. Figure 8 depicts the difference of confusion matrices between the modified SVM (A) and Ridge (B) classifiers. Positive numbers show that the prediction probability values of modified SVM model were higher than Ridge classifier. Conversely, the negative values indicate that the Ridge model's prediction probabilities were more than the modified SVM. Lastly, zero means that both models had the same prediction probability. As can be seen, both models have the same accuracy for all faults excepts faults 9 and 15, which the ridge classifier had a better accuracy than the modified SVM model by looking at the diagonal values of the difference matrix.



**Figure 7.** Confusion matrix (after 10-fold cross validation) of fault diagnosis results for PGA-Cosine classification algorithm with 29 principal components being selected.



**Figure 8.** The difference confusion matrix between 10-fold cross validation matrices of the modified SVM classifier presented in this study (A) and the ridge classifier presented by Smith et al. (B) [10].

## CONCLUSION

In this work, we present a fast, accurate, and robust algorithmic framework named FARM for industrial process monitoring. FARM is a holistic framework that synergistically performs fault detection and diagnosis tasks to improve monitoring performance. The fault detection module inside FARM adopts an advanced quantile-based SPC approach that can detect any mean shift of non-parametric and heterogenous multivariate data streams as soon as possible while maintaining a pre-specified false alarm rate. Meanwhile, the fault diagnosis module inside FARM implements a modified SVM algorithm for fault classification. Compared to standard SVM approach, our modified SVM algorithm includes an important data pre-processing step that makes use of the manifold insight of covariance matrix to greatly enhance classification accuracy. By validating and evaluating the performance of our FARM framework using the TEP dataset, we observe that 1) our fault detection module can achieve fast anomaly detection speed at a low false alarm rate, and 2) our fault diagnosis module successfully classifies 17 out of 20 fault scenarios at 100% accuracy. Unfortunately, faults IDV 3, IDV 9, and IDV 15 of the TEP dataset, which are known to be hard to classify, still face challenges in differentiating among one another with high accuracy. Our future work involves revamping the FARM framework to improve the classification accuracy of these hard-to-differentiate faults.

## ACKNOWLEDGEMENTS

We gratefully acknowledge financial support from Oklahoma State University College of Engineering, Architecture, and Technology's Startup Fund No. 1-155160 and from the National Science Foundation award TI-2331080.

## REFERENCE

1. Jackson, J.E., Mudholkar, G.S.: Control Procedures for Residuals Associated with Principal Component Analysis. *Technometrics*. 21, 341–349 (1979). <https://doi.org/10.2307/1267757>
2. Geladi, P., Kowalski, B.R.: Partial least-squares regression: a tutorial. *Anal. Chim. Acta*. 185, 1–17 (1986). [https://doi.org/10.1016/0003-2670\(86\)80028-9](https://doi.org/10.1016/0003-2670(86)80028-9)
3. Fezai, R., Mansouri, M., Taouali, O., Harkat, M.F., Bouguila, N.: Online reduced kernel principal component analysis for process monitoring. *J. Process Control*. 61, 1–11 (2018). <https://doi.org/10.1016/j.jprocont.2017.10.010>
4. Woodall, W.H., Spitzner, D.J., Montgomery, D.C., Gupta, S.: Using Control Charts to Monitor Process and Product Quality Profiles. *J. Qual. Technol.* 36, 309–320 (2004). <https://doi.org/10.1080/00224065.2004.11980276>
5. Zhao, H., Hu, Y., Ai, X., Hu, Y., Meng, Z.: Fault detection of Tennessee Eastman process based on topological features and SVM. *IOP Conf. Ser. Mater. Sci. Eng.* 339, 012039 (2018). <https://doi.org/10.1088/1757-899X/339/1/012039>
6. Onel, M., Kieslich, C.A., Pistikopoulos, E.N.: A nonlinear support vector machine-based feature selection approach for fault detection and diagnosis: Application to the Tennessee Eastman process. *AIChE J.* 65, 992–1005 (2019). <https://doi.org/10.1002/aic.16497>
7. Chebel-Morello, B., Malinowski, S., Senoussi, H.: Feature selection for fault detection systems: application to the Tennessee Eastman process. *Appl. Intell.* 44, 111–122 (2016). <https://doi.org/10.1007/s10489-015-0694-6>
8. Heo, S., Lee, J.H.: Fault detection and classification using artificial neural networks. *10th IFAC Symp. Adv. Control Chem. Process. ADCHEM 2018*. 51, 470–475 (2018). <https://doi.org/10.1016/j.ifacol.2018.09.380>
9. H. Ye, K. Liu: A Generic Online Nonparametric Monitoring and Sampling Strategy for High-Dimensional Heterogeneous Processes. *IEEE Trans. Autom. Sci. Eng.* 19, 1503–1516 (2022). <https://doi.org/10.1109/TASE.2022.3146391>
10. Smith, A., Laubach, B., Castillo, I., Zavala, V.M.: Data analysis using Riemannian geometry and



applications to chemical engineering. *Comput. Chem. Eng.* 168, 108023 (2022).  
<https://doi.org/10.1016/j.compchemeng.2022.108023>

11. Jiang, Z.: Online Monitoring and Robust, Reliable Fault Detection of Chemical Process Systems. In: Kokossis, A.C., Georgiadis, M.C., and Pistikopoulos, E. (eds.) *Computer Aided Chemical Engineering*. pp. 1623–1628. Elsevier (2023)
12. Qiu, P., Hawkins, D.: A Rank-Based Multivariate CUSUM Procedure. *Technometrics*. 43, 120–132 (2001)
13. Qiu, P., Hawkins, D.: A Nonparametric Multivariate Cumulative Sum Procedure for Detecting Shifts in All Directions. *J. R. Stat. Soc. Ser. Stat.* 52, 151–164 (2003)
14. Xian, X., Zhang, C., Bonk, S., Liu, K.: Online monitoring of big data streams: A rank-based sampling algorithm by data augmentation. *J. Qual. Technol.* 53, 135–153 (2021).  
<https://doi.org/10.1080/00224065.2019.1681924>
15. Y. Mei: Quickest detection in censoring sensor networks. In: *2011 IEEE International Symposium on Information Theory Proceedings*. pp. 2148–2152 (2011)
16. Downs, J.J., Vogel, E.F.: A plant-wide industrial process control problem. *Ind. Chall. Probl. Process Control*. 17, 245–255 (1993).  
[https://doi.org/10.1016/0098-1354\(93\)80018-I](https://doi.org/10.1016/0098-1354(93)80018-I)
17. Hu, M., Hu, X., Deng, Z., Tu, B.: Fault Diagnosis of Tennessee Eastman Process with XGB-AVSSA-KELM Algorithm. *Energies*. 15, (2022).  
<https://doi.org/10.3390/en15093198>
18. Andersen, E.B., Udugama, I.A., Gernaey, K.V., Khan, A.R., Bayer, C., Kulahci, M.: An easy to use GUI for simulating big data using Tennessee Eastman process. *Qual. Reliab. Eng. Int.* 38, 264–282 (2022).  
<https://doi.org/10.1002/qre.2975>
19. Rieth, C.A., Amsel, B.D., Tran, R., Cook, M.B.: Additional Tennessee Eastman Process Simulation Data for Anomaly Detection Evaluation, <https://doi.org/10.7910/DVN/6C3JR1>, (2017)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Development of Steady-State and Dynamic Mass-Energy Constrained Neural Networks using Noisy Transient Data

Angan Mukherjee<sup>a</sup>, and Debangsu Bhattacharyya<sup>a\*</sup>

<sup>a</sup> West Virginia University, Department of Chemical and Biomedical Engineering, Morgantown, WV, USA

\* Corresponding Author: [Debangsu.Bhattacharyya@mail.wvu.edu](mailto:Debangsu.Bhattacharyya@mail.wvu.edu).

---

## ABSTRACT

This paper presents the development of algorithms for mass-energy constrained neural network (MECNN) models that can exactly conserve the overall mass and energy of distributed chemical process systems, even though the noisy steady-state/transient data used for optimal model training violate the same. For developing dynamic mass-energy constrained network models for distributed systems, hybrid series and parallel dynamic-static neural networks are used as candidate architectures. The proposed approaches for solving both the inverse and forward problems are validated considering both steady-state and dynamic data in presence of various noise characteristics. The proposed network structures and algorithms are applied to the development of data-driven models of a nonlinear non-isothermal reactor that involves an exothermic reaction making it significantly challenging to exactly satisfy the mass and energy conservation laws of the system only by using the available input and output boundary conditions.

---

**Keywords:** Mass Conservation, Energy Conservation, Equality Constraints, Inverse Problem, Forward Problem, Noisy Data

## INTRODUCTION

Complex first-principles models are often needed to represent the complicated physics and chemistry associated with many chemical engineering applications. While the first-principles models can exactly satisfy mass and energy conservation, their development can consume considerable time, and resources, and require good understanding of the physics/chemistry that may not be available or may be difficult to obtain. One alternative is the data-driven / artificial intelligence (AI) modeling tools, that are relatively easier to develop, simulate, and adapt online for many chemical engineering applications [1,2]. Recent advances of machine learning (ML) have significantly influenced a wide range of scientific and engineering fields with applications to thermofluidic processes [3,4], computational fluid dynamics [5,6], as well as varieties of other complex industrial applications [7]. More recently, physics informed neural networks [8] (PINNs) have gained popularity due to their effectiveness in solving realistic practical problems with noisy data and often partially missing physics, with implementations ranging from solving systems of partial differential equations

[9,10] to modeling different chemical processes such as photochemical systems [11], hydrofluorocarbon refrigerant mixtures [12], biomass pyrolysis process [13], heat transfer problems [7], etc. In all these works, the physics conservation equations have been either augmented in the loss (objective) function as additional penalty terms to penalize the violation of such constraint(s) or by integrating them into sequential conservation layers following the data-driven model. If the system conservation laws are included in the objective function for training the neural network (NN) models, they serve as soft constraints and do not ensure an 'exact' satisfaction of the physics constraints. However, in numerous chemical engineering applications, it is desired that certain mathematical relationships (such as mass and energy balances) are exactly conserved and not only approximated<sup>3</sup>. In very few recent publications [3,14] that seek to satisfy the physics constraints exactly, the proposed algorithms are either highly system specific, or require rigorous understanding of the chemical process to impose appropriate physics conservation laws pertaining to the system. These approaches can lead to higher errors, especially during the forward problem. Another common approach

in the open literature is to send the NN model outputs through a linear transformation block to satisfy mass and energy balances. However, physics-conserving equations are not necessarily linear. Furthermore, the corrections applied to the network output in the conservation layer, if applied simultaneously while training the network rather than sequentially, can affect the training of the NN model.

It is also desired that the physics constraints be exactly satisfied not only during solving the inverse problem, but also during the forward problem for unknown inputs. In this work, 'inverse' problem refers to the training approaches for optimal NN models, whereas 'forward' problem is represented by model validation / simulation. Most variants of typical PINNs that seek to conserve the mass [3,4,13] and energy [7] of chemical systems are found to satisfy the conservation constraints only approximately through penalty terms included in the objective function. Furthermore, these models may not satisfy the mass/energy balances during the forward problem. Exactly satisfying mass and energy conservation of chemical process systems is critical to ensure that model predictions remain realistic and bounded by physical laws.

Another aspect which poses to be a significant challenge during optimal model synthesis and parameter estimation is to develop accurate constrained network models when the data are noisy [15]. The presence of noise in the training data can corrupt the gradient calculations during estimation of network parameters [16] and can result in biased estimates of parameters during model training. Though variants of PINNs such as Bayesian PINNs [17], fractional PINNs [18], etc. address existence of noise in data while solving systems of partial differential equations, existing algorithms that seek to exactly conserve system physics, especially mass and energy balances, do not consider the presence of uncertainties in training data at all [3,4,15]. However, to the best of our knowledge, although a few papers exist till date trying to exactly conserve mass of a system [3,4], there is currently no work in existing literature focused on exactly conserving the energy of distributed chemical process systems, perhaps due to the complexities associated with distributed nature of energy constraints, energy addition/removal at different spatial locations as well as loss or generation of heat due to complicated mechanisms including endothermic/exothermic reactions in the system. Furthermore, such approaches consider only static feedforward NNs and obviously only steady-state data for model development. Developing a fully data-driven dynamic model that can satisfy system physics, especially mass and energy balance constraints, can be challenging since mass/energy conservation can be adequately applied only to the steady-state data especially if the system is distributed and there is no

measurement available for the system holdup. Though one may desire to satisfy practically any type of physics-based constraint equation through such approaches, in this work we are primarily focused on exactly conserving the mass and energy of a system, hence denoting the corresponding NNs as Mass-Energy Constrained Neural Networks (MECNNs). The primary objective of the paper is to develop optimal NN models that exactly satisfy mass and energy conservation of chemical processes, even in the presence of noise in transient training data. In summary, contributions of this work are as follows:

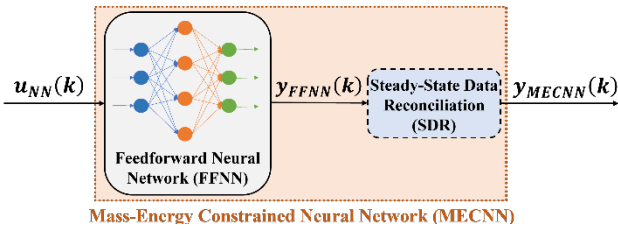
- Both training (inverse problem) and validation (forward problem) algorithms are developed for MECNNs, where the mass (atom) and energy (enthalpy) balance equations are imposed as equality constraints during parameter estimation to ensure that mass and energy conservation laws are exactly satisfied in a fully data-driven approach.
- It is desired that the proposed approach accommodates nonlinear transformations for constraint satisfaction and is also applicable when there is no close-form equation available for satisfying the desired constraints.
- Different uncertainty characteristics, represented by presence of a constant (time-invariant) and random (time-varying) bias in addition to a zero-mean (known variance) Gaussian noise are simulated.
- In addition to steady-state modeling of nonlinear chemical process systems, our approach also addresses dynamic data-driven model development using noisy transient data. The proposed algorithms are demonstrated for modeling the non-isothermal Van de Vusse reactor system [19].

## MECNN STRUCTURE AND FORWARD / INVERSE PROBLEMS

### Steady-State MECNN Implementation

Novel model structures have been developed for both steady-state and dynamic implementations of MECNNs. The block-oriented structure for steady-state MECNNs, as shown in Figure 1, considers a typical feedforward neural network (FFNN) model followed by a steady-state data reconciliation (SDR) step, where  $\mathbf{y}_{MECNN}$  represent the final outputs from the optimal MECNNs,  $\mathbf{y}_{FFNN}$  refer to the intermediate outputs obtained from the FFNN corresponding to the model inputs ( $\mathbf{u}_{NN}(k)$ ), and  $k$  denotes the indices of steady-state training data. The FFNN model is a single hidden-layered feedforward NN involving logistic sigmoid activation functions and the Levenberg-Marquardt (LM) algorithm for parameter estimation. The choice of FFNN model is completely arbitrary since the proposed approaches can be applied irrespective of the specific network model architecture. More details about various FFNN models can

be found in numerous references [1,20]. The SDR block provides the additional degrees of freedom necessary for solving the equality constrained optimization problem, thus facilitating the simultaneous estimation of parameters for both FFNN and SDR blocks through the inverse problem formulation described by Equation (1). The mass and energy balance equations of the system represent the equality constraints imposed on the MECNN during both inverse and forward problems to guarantee the exact conservation of system physics. The proposed architecture is also flexible to accommodate any linear or nonlinear close-form transformation equations that may be used for constraint satisfaction by substituting the SDR block which requires solving an optimization problem for convergence.



**Figure 1.** General Architecture of Steady-State MECNNs

The inverse problem for steady-state modeling using MECNNs combines optimization problems representing the regression of FFNN model as well as reconciliation of the neural network outputs. Therefore, in such approaches, not only the FFNN model parameters (referred to as  $\theta_{FFNN}$ ) are estimated, but also the outputs from the FFNN model ( $y_{NN}$ ) are reconciled simultaneously, in presence of the mass and energy conservation constraints represented by Equations (1c) and (1d).

$$\min_{\mathbf{y}_{MECNN}, \theta_{FFNN}} \sum \left( (\mathbf{y}_{tar}(k) - \mathbf{y}_{MECNN}(k))^T R^{-1} (\mathbf{y}_{tar}(k) - \mathbf{y}_{MECNN}(k)) \right) + \sum \left( (\mathbf{y}_{MECNN}(k) - \mathbf{y}_{FFNN}(k))^T R^{-1} (\mathbf{y}_{MECNN}(k) - \mathbf{y}_{FFNN}(k)) \right) \quad (1a)$$

$$s. t. \quad \mathbf{y}_{FFNN}(k) = \mathbf{f}_{FFNN}(\mathbf{u}_{NN}(k), \theta_{FFNN}) \quad (1b)$$

$$\mathbf{h}_M(\mathbf{u}_{NN}, \mathbf{y}_{MECNN}, k) = 0 \quad (1c)$$

$$\mathbf{h}_E(\mathbf{u}_{NN}, \mathbf{y}_{MECNN}, k) = 0 \quad (1d)$$

where,  $\mathbf{y}_{tar}$  denotes the target training data for MECNNs.

Although the FFNN model parameters (weights and biases) can be regressed to guarantee the exact satisfaction of mass and energy balance constraints for all data on which the MECNN is trained on, it can still not be ensured that that the optimal FFNN ( $FFNN_{opt}$ ) will continue to exactly conserve mass and energy of a system when subjected to unknown inputs different from those encountered during model training. Therefore, for the forward problem (i.e., model simulation) as well, the SDR block exists following the optimal FFNN model, in a similar configuration as Figure 1, to ensure the exact conservation of system mass and energy even during model simulation (validation). However, for this case, the SDR

block simply acts as a post-processing step leading to the following formulation of the forward problem:

$$\min_{\mathbf{y}_{MECNN}} \sum \left( \left( \mathbf{y}_{FFNN_{opt}}(k) - \mathbf{y}_{MECNN}(k) \right)^T R^{-1} \left( \mathbf{y}_{FFNN_{opt}}(k) - \mathbf{y}_{MECNN}(k) \right) \right) \quad (2a)$$

$$s. t. \quad \mathbf{y}_{FFNN_{opt}}(k) = \mathbf{f}_{FFNN}(\mathbf{u}_{NN}(k), \theta_{FFNN_{opt}}) \quad (2b)$$

$$\mathbf{h}_M(\mathbf{u}_{NN}, \mathbf{y}_{MECNN}, k) = 0 \quad (2c)$$

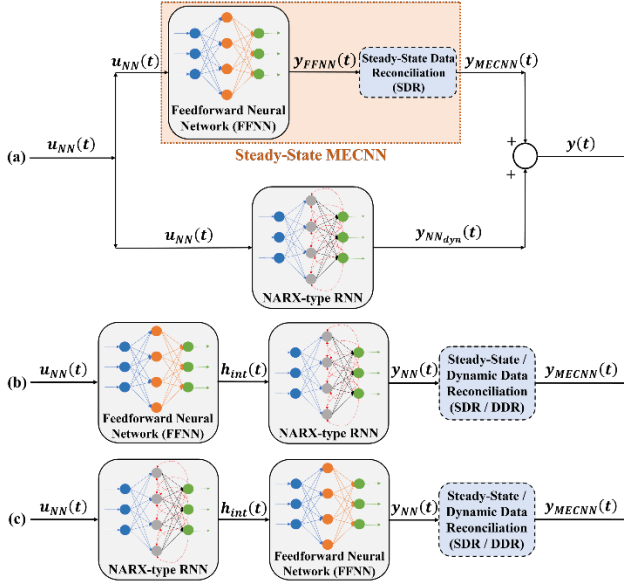
$$\mathbf{h}_E(\mathbf{u}_{NN}, \mathbf{y}_{MECNN}, k) = 0 \quad (2d)$$

## Dynamic MECNN Implementation

Although typical FFNNs feature sufficiently satisfactory performance while modeling various steady-state systems, it may be significantly challenging for static networks to accurately model highly complex nonlinear temporal systems with time-lagged input variables. One of our previous papers [1] has shown that hybrid series and parallel all-nonlinear static-dynamic network models exhibit superior performance than many state-of-the-art recurrent neural networks (RNNs), especially while handling large-sized dynamic training data containing complex nonlinearities. Such hybrid models consider typical FFNNs representing the nonlinear static (NLS) network, while the nonlinear dynamic (NLD) network is characterized by the Nonlinear Auto-Regressive with eXogenous inputs (NARX) type RNN which provide efficient approximations while modeling temporal systems due to presence of feedback connections eradicating the necessity of data windows for time-lagged inputs [21]. The NARX-type RNN models considered in this work also consist of a single hidden layer and the logistic as well as hyperbolic tangent sigmoid activation functions, where the parameters are estimated by the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm. Furthermore, for different transient systems, local measurements depicting holdup information may or may not be available. Therefore, the algorithms for developing accurate dynamic MECNNs may differ based on whether holdup measurements are available. Since both hybrid parallel and series network models can be considered during development of dynamic MECNNs, two different configurations are possible, namely the hybrid parallel MECNN and the hybrid series MECNN respectively, as shown in Figure 2.

The hybrid parallel MECNN operates under the assumption that the entire time-series data can be distinctly partitioned into steady-state and dynamic zones. This approach of partitioning training data also facilitates the flexibility of solving a sequential optimization problem [1,23] for the overall inverse problem of the hybrid parallel MECNN. In such an approach, the steady-state MECNN is first developed based on only the steady-state zones, i.e., the formulation of the inverse problem

remains exactly same as given by Equation (1), followed by constructing a deterministic NLD deviation model ( $NN_{dyn}$ ) represented by a NARX-type RNN to match the residuals / deviations obtained with respect to the training data and steady-state MECNN model outputs. The deviation data thus generated serve as target data for training the NARX-type RNN model independently during the sequential approach while minimizing the typical squared error objective function as given by Equation (3).



**Figure 2.** General Architecture of (a) hybrid parallel and (b), (c) hybrid series MECNNs

$$\min_{\theta_{NARX}} \left( \mathbf{y}_{tar}(t) - \mathbf{y}_{MECNN}(t) - \mathbf{y}_{NN_{dyn}}(t) \right)^T R^{-1} \left( \mathbf{y}_{tar}(t) - \mathbf{y}_{MECNN}(t) - \mathbf{y}_{NN_{dyn}}(t) \right) \quad (3a)$$

$$s. t. \mathbf{y}_{NN_{dyn}}(t) = \mathbf{g}_{NARX}(\mathbf{u}_{NN}(t), \theta_{NARX}) \quad (3b)$$

Unlike hybrid parallel MECNNs, the hybrid series MECNNs are flexible to accommodate the entire time-series dynamic data, regardless of whether there is access to system holdup information or not. When holdup measurements are not available, then the mass and energy conservation constraints can only be applied to the steady-state zones while formulating the SDR block in the inverse problem due to insufficient information about the system holdup during transients. On the other hand, for the case when the system holdup information is available during formulation of the inverse problem for hybrid series MECNNs, the approach discussed above can readily be extended to exactly satisfy the mass ( $\mathbf{h}_M(\cdot)$ ) and energy ( $\mathbf{h}_E(\cdot)$ ) balance constraints at each time step in the transient profile, thus leading to substituting the SDR step with a dynamic data reconciliation (DDR) based transformation [22]. For systems with no or negligible

spatial distribution, holdup measurements may be available, such as continuous stirred tank reactors (CSTRs). The corresponding forward problems for dynamic MECNNs follow similar formulations based on the sequential architectures involving the SDR / DDR post-processing step.

## Error Characterizations

Measurement data considered for developing optimal accurate data-driven models often contain uncertainties stemming from various factors. The two types of error characterizations (Equations (4) and (5)) that have been considered in this work during both steady-state and dynamic implementations of MECNNs include adding constant (time-invariant) and random (time-varying) biases ( $\boldsymbol{\beta}$ ) respectively, to the true data,  $\mathbf{y}_{true}$  (generated from simulation) in presence of Gaussian noise ( $\eta(0, \sigma^2)$ ). The primary goal is to assess how the presence or absence of noise and its characterization impact the performance of the proposed MECNNs and their capabilities to exactly conserve the mass and energy of a system. It is to be noted that for the case when the measurements are corrupted with a random (time-varying) bias, the inverse problem formulation as given by Equation (1) may lead to lack of identifiability while constructing accurate data-driven models which yield results close to the system truth. Therefore, in this work, additional parametric (linear, quadratic, etc.) noise models have been incorporated into the inverse problem formulation while addressing random (time-varying) bias in measurement data.

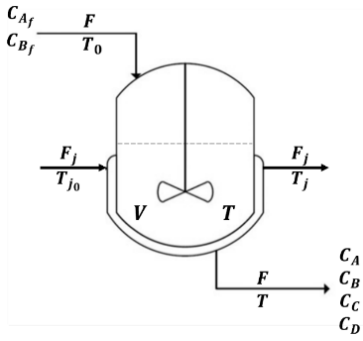
$$\mathbf{y}_{tar}(t) = \mathbf{y}_{true}(t) + \boldsymbol{\beta} + \eta(0, \sigma^2) \quad (4)$$

$$\mathbf{y}_{tar}(t) = \mathbf{y}_{true}(t) + \boldsymbol{\beta}(t) + \eta(0, \sigma^2) \quad (5)$$

## CASE STUDY: NON-ISOTHERMAL VAN DE VUSSE REACTOR SYSTEM

The nonisothermal model of the Van de Vusse reactor system is one of the most vastly studied multi-input-multi-output (MIMO) processes due to its high complex nonlinearities [19]. The reaction involves the production of cyclopentenol (B) from cyclopentadiene (A) by acid-catalyzed electrophilic addition of water in dilute aqueous solution. The side products of this reaction are cyclopentanediol (C) and dicyclopentadiene (D). Cooling water flows through the jacket surrounding the reactor to remove excess heat generated due to the highly exothermic reaction within the system, as shown in Figure 3. The detailed mathematical model consisting of the overall material and energy balance equations and nominal values of the involved parameters can be found in the literature [19].





**Figure 3.** Schematic of the Nonisothermal Van de Vusse Reactor for MECNN Implementation

The feed stream contains predominantly component A, with small quantities of B. Following variables are considered to be the inputs to the MECNN model-volumetric flowrate ( $F$ ) and inlet temperature ( $T_0$ ) of the feed, volumetric flowrate ( $F_j$ ) of the jacketing fluid, along with the concentrations of A ( $C_{A_f}$ ) and B ( $C_{B_f}$ ) in the feed stream. Following variables are considered to be the outputs for the model-concentration of all reaction species, i.e., A ( $C_A$ ), B ( $C_B$ ), C ( $C_C$ ), D ( $C_D$ ) in the product stream and the outlet temperatures of the product stream ( $T_{MECNN}$ ) and jacketing fluid ( $T_{j_{MECNN}}$ ). It is assumed that the inlet and outlet volumetric flowrates are the same across the reactor system as well as across the cooling jacket. The steady-state mass balance constraints in terms of elemental atom (C, H, and O) balance equations and the energy balance constraint have been provided in Equations (6) through (9), where  $n_{C_i}$ ,  $n_{H_i}$ , and  $n_{O_i}$  represent the number of C, H, and O atoms respectively in the  $i^{th}$  species,  $\rho_r$  denotes the density of the reactor mixture,  $h_{in}$  and  $h_{out}$  refer to the specific enthalpies of the reactant and product mixtures as functions of inlet ( $T_0$ ) and outlet ( $T_{MECNN}$ ) temperatures,  $m_j$ ,  $V_j$  and  $C_{p_j}$  refer to the mass, volume and specific heat capacity of cooling water flowing through the jacket, and  $T_{j_0}$  represent the inlet temperature of the jacketing fluid (cooling water).

$$\sum_{i=A,B,H_2O} n_{C_i} * F(k) * C_{i_f}(k) = \sum_{i=A,B,C,D,H_2O} n_{C_i} * F(k) * C_{i_{MECNN}}(k) \quad (6)$$

$$\sum_{i=A,B,H_2O} n_{H_i} * F(k) * C_{i_f}(k) = \sum_{i=A,B,C,D,H_2O} n_{H_i} * F(k) * C_{i_{MECNN}}(k) \quad (7)$$

$$\sum_{i=A,B,H_2O} n_{O_i} * F(k) * C_{i_f}(k) = \sum_{i=A,B,C,D,H_2O} n_{O_i} * F(k) * C_{i_{MECNN}}(k) \quad (8)$$

$$F(k) * \rho_r * (h_{in}(T_0(k)) - h_{out}(T_{MECNN}(k))) = \frac{F_j(k)}{V_j} * m_j * C_{p_j} * (T_{j_{MECNN}}(k) - T_{j_0}) \quad (9)$$

Since this case study considers a CSTR system with no spatial distribution, the system holdup information can be explicitly modeled using the outlet boundary conditions and reactor instantaneous volume, since the accumulation term for concentrations / temperature of the

reaction species in a CSTR is exactly same as the corresponding outlet values at every time instance during dynamic modeling. For example, the mass (in terms of C atom) and energy balance constraints which can be imposed at every time step ( $t$ ) for this case study can be expressed in terms of inlet / outlet flowrates ( $F_{in}$ ,  $F_{out}$ ) and instantaneous reactor volume ( $V$ ) as:

$$\frac{d(V C_i)}{dt} - \left[ \sum_{i=A,B,H_2O} n_{C_i} * F_{in}(t) * C_{i_f}(t) - \sum_{i=A,B,C,D,H_2O} n_{C_i} * F_{out}(t) * C_{i_{MECNN}}(t) \right] = 0 \quad (10)$$

$$\frac{d(\rho_r V h)}{dt} - \left[ \rho_r * (F_{in}(t) h_{in}(T_0(t)) - F_{out}(t) h_{out}(T_{MECNN}(t))) - \frac{F_j(t)}{V_j} * m_j * C_{p_j} * (T_{j_{MECNN}}(t) - T_{j_0}) \right] = 0 \quad (11)$$

## RESULTS AND DISCUSSIONS

### Steady-State Modeling

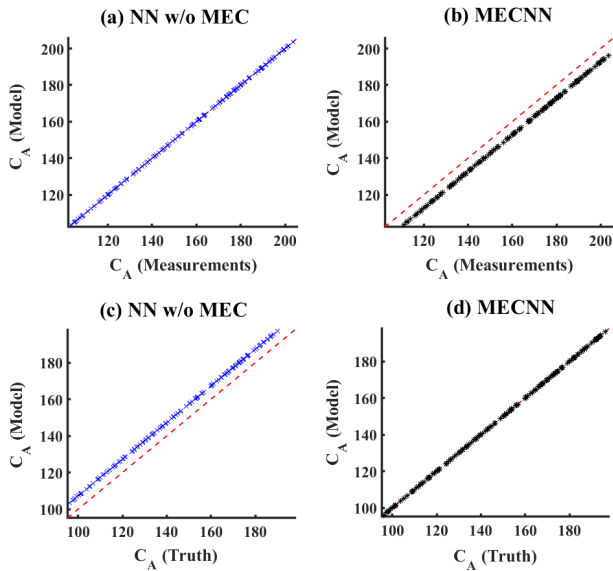
The different types of error characterizations considered for generating measurement data for the steady-state modeling of the non-isothermal Van de Vusse reactor system have been separately analyzed. For presenting both steady-state and dynamic results, the data used for training MECNN models are referred to as 'measurements', irrespective of whether they had been generated from a first-principles model or collected from experiments. Furthermore, the violation of mass and energy conservation constraints by the network models have been quantified in terms of absolute percentage error (APE) in atom and energy (Q) balances defined as:

$$APE_{atom_i} = \left| \frac{\sum_{in} F_{m_{atom_i}} - \sum_{out} F_{m_{atom_i}}}{\sum_{in} F_{m_{atom_i}}} \right| * 100\% \quad (12)$$

$$APE_{energy} = \left| \frac{Q_R - Q_j}{Q_R} \right| * 100\% \quad (13)$$

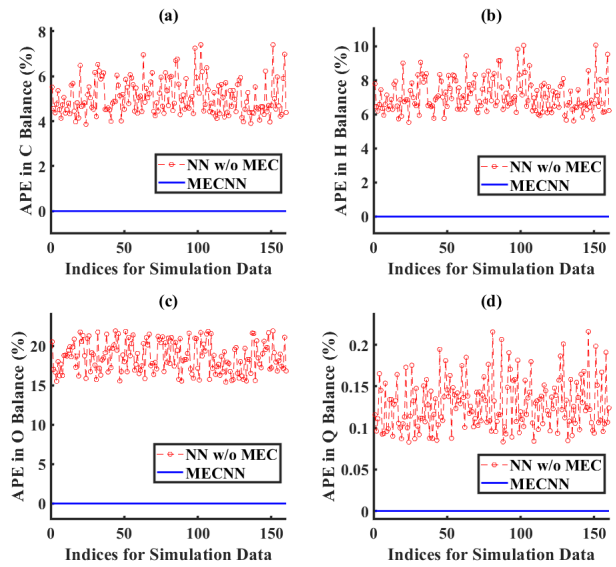
where,  $\sum_{in} F_{m_{atom_i}}$  and  $\sum_{out} F_{m_{atom_i}}$  respectively denote the total mass flowrate of the  $i^{th}$  atom in all inlet and outlet streams, and  $Q_R$  and  $Q_j$  refer to the total heat lost by the reactor mixture and that gained by the jacketing fluid (cooling water) respectively. While the root mean squared error (RMSE) calculated between the measurements / truth and model predictions account for the predictive capabilities of the proposed data-driven models for nonlinear steady-state / dynamic data, the APE in atom and energy balances represent the magnitude of the maximum (worst-case) violation of the mass and energy balance constraints respectively. The specific enthalpy terms involved in the energy balance constraint (as given by Equation (9)) have been appropriately expressed in terms of specific heat capacity of the reactor mixture and the heat of reaction calculated from the

respective heats of formation of the reactants and products. For brevity, only the steady-state results obtained when the measurement data are corrupted with a constant bias along with Gaussian noise, have been included here. Constant biases equal to 5% of the mean of the true concentrations and a constant positive value of 3.5°C for each temperature variables have been added to the respective outputs in presence of a Gaussian noise with  $\mu = 0$  and  $\sigma^2 = 0.25$ . Figure 4 compares the model results vs truth and measurements for MECNN and NN w/o mass-energy constraints (MEC) for simulation data of  $C_A$  ( $\frac{kmol}{m^3}$ ).



**Figure 4.** Comparison of results between MECNN and NN w/o MEC for simulation data of  $C_A$  (noise in measurement data represented by Equation (4))

The unconstrained NN yields an excellent match with about 0.03% RMSE with respect to the measurements but results in biased estimates compared to the truth, thus leading to around 5.3% RMSE with respect to the true data. On the contrary, the MECNN accurately captures the system truth (characterized by approximately 0.01% RMSE) but shows a clear mismatch with the measurements leading to around 4.9% RMSE. The imposition of mass and energy conservation laws as equality constraints, along with modification of the objective function as given in Equation (1a) results in the outputs from MECNN that minimize error with respect to the system truth, as opposed to minimizing error with respect to the measurements (training data). Furthermore, the MECNN exactly satisfies both mass and energy balance constraints during the forward problem as well, as evident from Figure 5, whereas the NN w/o those constraints show maximum APE in C, H, O and Q balance as 7.4%, 10.1%, 21.8%, and 0.22% respectively.



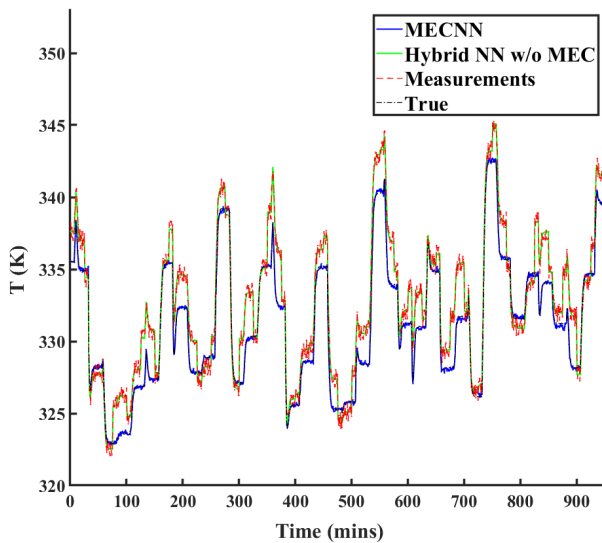
**Figure 5.** Comparison between MECNN and NN w/o MEC in terms of violating (a) C, (b) H, (c) O and (d) Q balance constraints during simulation (noise in measurement data represented by Equation (4))

## Dynamic Modeling

The performance of dynamic MECNNs is evaluated in this paper for the case when a time-varying bias is added to the system truth to generate measurements in presence of Gaussian noise, as given in Equation (5). As discussed before, different parametric forms of a noise (error) model, such as linear, quadratic, etc. have been considered during optimal model synthesis when the measured data are corrupted with a time-varying bias. In this case, time-varying biases in the range of -4% to +6% of the true outlet concentrations have been added to the concentration variables and that within -1°C to 6°C have been incorporated into the true temperature variables to generate training data, along with an additional Gaussian noise distribution with  $\mu = 0$  and  $\sigma^2 = 0.25$ . The linear parametric form of the noise (error) model yielded superior results in terms of percentage RMSE with respect to true data as well as minimum values of the corrected Akaike Information Criteria [1,23] ( $AIC_c$ ) evaluated during model training, as compared to the quadratic form as well as the case when no separate error model was considered during parameter estimation. Figure 6 shows the comparison between the results obtained from the hybrid series MECNN vs standalone series NN without mass-energy constraints (MEC) for simulation data of  $T$ .

While the unconstrained hybrid NLS – NLD model shows an excellent match with the measurements (represented by around 0.09% RMSE), the results obtained violate the mass and energy conservation of the system leading to approximately 2.5% RMSE with the true data. On the other hand, the hybrid series MECNN shows a significant mismatch with the measured data (characterized

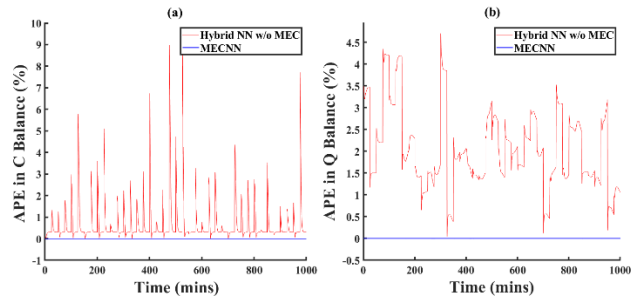
by around 2.8% RMSE) but accurately captures the system truth recording an overall percentage RMSE of just 0.04%. The MECNN has also been seen to exactly satisfy the C, H, O, and energy balances at steady state while showing some mismatch during transients. However, the unconstrained NN model show APE in the mass and energy conservation equations as high as 22%, 30%, 60%, and 0.5% respectively. It is to be noted that for process control applications, the dynamic models for MECNNs can be used as soft sensors, especially for those involving an economic measure where such mass / energy imbalances can lead to inaccurate accounting of the control objective. The impact of inaccurate measurements on the key economic measures of the process such as efficiency has been extensively investigated by other researchers [24,25].



**Figure 6.** Comparison of results between hybrid series MECNN and NN w/o MEC for simulation data of  $T$  (noise in measurement data represented by Equation (5))

When holdup measurements are available, the proposed algorithms can be extended to exactly satisfy mass and energy balance constraints even during the transients. Since this case study example is essentially a CSTR, the concentrations of the reaction species and the instantaneous temperature of the reaction mixture inside the reactor at any given point of time can be assumed to be exactly equal to the outlet concentrations and temperature of the same at that time instant. The plots comparing results from unconstrained vs constrained network models when system holdup information is available have not been discussed here for brevity. However, unlike previous cases, the MECNN exactly satisfies the mass and energy conservation even during the transients due to the imposition of mass and energy balance constraints at every time step of the dynamic training / simulation data. The unconstrained NN model, for this case,

shows significantly higher APEs in C, and Q balances recording maximum errors of 9% and 4.5% respectively, as shown in Figure 7.



**Figure 7.** Comparison between hybrid series MECNN and NN w/o MEC in terms of violating (a) C (b) Q constraints (holdup info available) during simulation (noise in measurement data represented by Equation (4))

## CONCLUSIONS

In this study, we have developed algorithms for synthesizing steady-state and dynamic neural network models capable of exactly satisfying mass and energy conservation laws, even when the training data used for training the networks violate these laws. Furthermore, we have also demonstrated that if system holdup information is available during process transience, these algorithms can also exactly satisfy physics constraints during process dynamics. The proposed approaches can be individually applied to different equipment items in interconnected systems assuming that the required boundary information is measured / available. For interconnected reactive systems and systems involving mass / energy integration / losses, application of the proposed approach to the plant-wide system needs further investigation. For all results presented in this paper, the steady-state and dynamic MECNNs consistently exhibited RMSE of less than 1% when compared to the true data for both inverse and forward problems. On the contrary, the corresponding unconstrained networks resulted in significantly higher errors. The MECNN structures developed as part of this work can be applied not only for steady-state modeling, but also in effectively representing dynamic process systems represented by noisy measurements. However, since the proposed approaches work on a minimum-bias principle during estimation of optimal parameters in presence of different types of uncertainties in training data, it is important to clarify that these algorithms do not necessarily guarantee perfect predictions of the true data as expected but do ensure that mass and energy balance constraints are exactly satisfied during both inverse (training) and forward (simulation) problems.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge the U.S. DOE grant for the project titled “Boiler Health Monitoring Using a Hybrid First Principles – Artificial Intelligence Model” (Grant #: DE – FE0031768).

## REFERENCES

1. Mukherjee A, Bhattacharyya D. Hybrid Series / Parallel All-Nonlinear Dynamic-Static Neural Networks: Development, Training, and Application to Chemical Processes. *Ind Eng Chem Res.* 62:3221–3237 (2023).
2. Venkatasubramanian V. The promise of artificial intelligence in chemical engineering: Is it here, finally? *AIChE J.* 65:466–478 (2019).
3. Carranza-Abaid A, Jakobsen JP. Neural network programming: Integrating first principles into machine learning models. *Comput Chem Eng.* 163: 107858 (2022).
4. Ihunde TA, Olorode O. Application of physics informed neural networks to compositional modeling. *J. Pet. Sci. Eng.* 211:110175 (2022).
5. Cai S, Mao Z, Wang Z, Yin M, Karniadakis GE. Physics-informed neural networks (PINNs) for fluid mechanics: a review. *Acta Mech. Sin.* 37:1727–1738 (2021).
6. Zheng H, Huang Z, Lin G. PCNN: A physics-constrained neural network for multiphase flows. *Phys Fluids.* 34:102102 (2022).
7. Cai S, Wang Z, Wang S, Perdikaris P, Karniadakis GE. Physics-informed neural networks for heat transfer problems. *J. Heat Transfer.* 143:1–15 (2021).
8. Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J Comput Phys.* 378:686–707 (2019).
9. Blechschmidt J, Ernst OG. Three ways to solve partial differential equations with neural networks — A review. *GAMM Mitteilungen.* 44:1–29 (2021).
10. Subraveti SG, Li Z, Prasad V, Rajendran A. Physics-Based Neural Networks for Simulation and Synthesis of Cyclic Adsorption Processes. *Ind Eng Chem Res.* 61:4095–4113 (2022).
11. Sturm PO, Wexler AS. Conservation laws in a neural network architecture: enforcing the atom balance of a Julia-based photochemical model (v0.2.0). *Geosci Model Dev.* 15:3417–3431 (2022).
12. Befort BJ, Garciadiego A, Wang J, Wang K, Franco G, Maginn EJ, Dowling AW. Data science for thermodynamic modeling: Case study for ionic liquid and hydrofluorocarbon refrigerant mixtures. *Fluid Phase Equilib.* 572:113833 (2023).
13. Xing J, Kurose R, Luo K, Fan J. Chemistry-Informed Neural Networks modelling of lignocellulosic biomass pyrolysis. *Bioresour Technol.* 355:127275 (2022).
14. Beucler T, Pritchard M, Rasp S, Ott J, Baldi P, Gentile P. Enforcing Analytic Constraints in Neural Networks Emulating Physical Systems. *Phys Rev Lett.* 126:098302 (2021).
15. Karniadakis GE, Kevrekidis IG, Lu L, Perdikaris P, Wang S, Yang L. Physics-informed machine learning. *Nat Rev Phys.* 3:422–440 (2021).
16. Psychogios DC, Ungar LH. A hybrid neural network-first principles approach to process modeling. *AIChE J.* 38:1499–1511 (1992).
17. Yang L, Meng X, Karniadakis GE. B-PINNs: Bayesian physics-informed neural networks for forward and inverse PDE problems with noisy data. *J Comput Phys.* 425:109913 (2021).
18. Pang G, Lu L, Karniadakis GE. fPINNs: Fractional Physics-Informed Neural Networks. *SIAM J Sci Comput.* 41:A2603–A2626 (2019).
19. Ridlehoover GA, Seagrave RC. Optimization of Van de Vusse Reaction Kinetics Using Semibatch Reactor Operation. *Ind Eng Chem Fundam.* 12:444–447 (1973).
20. Himmelblau DM. Accounts of experiences in the application of artificial neural networks in chemical engineering. *Ind Eng Chem Res.* 47:5782–5796 (2008).
21. Du KL, Swamy MNS. *Neural Networks and Statistical Learning.* 2nd Edition. Springer London (2014).
22. Mukherjee A, Bhattacharyya D. On the Development of Steady-State and Dynamic Mass-Constrained Neural Networks Using Noisy Transient Data. *Comput Chem Eng.* (2024).
23. Ferguson JM, Taper ML, Zenil-Ferguson R, Jasieniuk M, Maxwell BD. Incorporating Parameter Estimability Into Model Selection. *Front Ecol Evol.* 7:1–15 (2019).
24. Paul P, Bhattacharyya D, Turton R, Zitney SE. Sensor network design for maximizing process efficiency: an algorithm and its application. *AIChE J.* 61(2):464–476 (2015).
25. Paul P, Bhattacharyya D, Turton R, Zitney SE. Dynamic model-based sensor network design algorithm for system efficiency maximization. *Comput Chem Eng.* 89:27–40 (2016).

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Model Based Process Development and Operation of a Fluid Bed Granulation Unit to Manufacture Pharmaceutical Tablets.

Salvador García Muñoz\*, Maitraye Sen, Shashwat Gupta, Ronald Ruff

<sup>a</sup> Synthetic Molecule Design and Development, Eli Lilly and Co. Indianapolis IN. 46074 USA

\* Corresponding Author: sal.garcia@lilly.com

## ABSTRACT

A hybrid model for the fluid bed granulation operation was built. The deterministic component focuses on the mass and energy balances representing the water ingress and egress from the powder bed. The empirical one does on granule growth. Estimability techniques were used to determine which parameters to regress from the available data. A partial least squares approach was used to better understand the impact of the model parameters onto key model responses and sensitivity plots were made to aid operational decisions and support pharmaceutical development.

**Keywords:** Process Operations

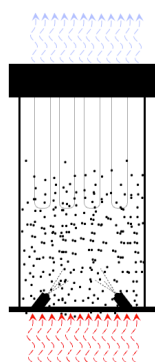
## INTRODUCTION

Fluid bed granulation is a common operation in the manufacture of pharmaceutical oral solid dosage forms (i.e. tablets and capsules). The aim of the operation is to agglomerate the blend of primary particles in the incoming powder blend to increase particle size and density. This operation becomes necessary when simpler approaches (e.g. direct compression) are unfeasible due to the poor flowability of the unprocessed powder mixture. The operation typically consists of a closed chamber where powder is loaded, an aqueous solution of a binding polymer

is sprayed onto the powder as drying air is fed through the bottom of the chamber fluidizing the powder and aiding the incorporation of the spray into the powder bed. The operation consists of two main phases after the powder is loaded, a *spray* phase where a pre-defined amount of solution is sprayed onto the powder and a *drying* phase where the spray is stopped, and air continues to be fed to fluidize and dry the agglomerates (granules) to a desired level of humidity (Figure 1).

The common manipulated variables (MV) in this operation are i) the amount of spray added ii) the spraying rate, iii) the temperature of the drying air, iv) the air flow during the spraying phase and v) airflow during the drying phase. A typical constraint is the composition of the powder blend since this is determined to satisfy a desired dose in the final product.

The aim of this study is to approach the design of the operating conditions for this unit for a specific product, utilizing scientific understanding rather than solely through empiricism. We do this through the formulation of a mathematical expression (a model) that captures the combined effect of the multiple driving forces affecting the in the system. Once the model is formulated, we use in-silico studies to define a design space for commercial manufacturing.



**Figure 1.** Schematic of a fluid bed granulation chamber



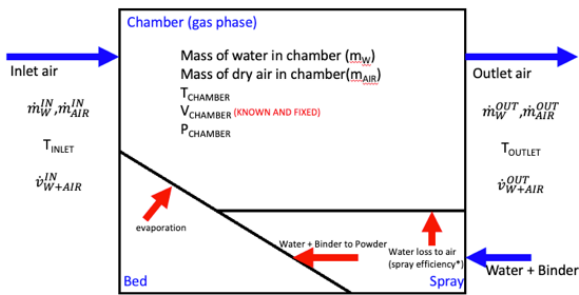
## MODEL DESCRIPTION

Detailed studies [1] demonstrate the strong coupling between granule growth and drying mechanism. Both mechanisms are tied since the granule size will determine the particle specific surface area available for water intake and drying. Some researchers have focused on studying the granule growth solely [2], while others focus the study on the water intake and egress [3] under the hypothesis that controlling one mechanism will result in control of the other.

Our work approaches the drying phenomena through mechanistic modeling while addressing the granule size through an empirical model.

### Mass and Energy Conservation

Our approach follows that of Ochsenbein et. al [3] with some differences that are discussed. The considered flows of water across the system are illustrated in figure 2. Three distinct phases are considered: spray, powder bed, and chamber (gas phase). Liquid water ingresses the system through the spray and most of it assumed to be immediately transferred to the powder bed. Water vapor is assumed to be lost directly from the spray to the gas phase through an efficiency and transferred between the powder bed and the gas phase through a drying expression.



**Figure 2.** Mass flows across phases

Several key differences separate our work from that of Ochsenbein et al. [3]. The way we calculate the outlet airflow is discussed first. The only inlet of bone-dry air is the inlet air flow. We assume that the volume of humid air leaving the chamber is a natural response to equalizing the pressure in the chamber as it can build up or down depending on the temperature or the water content in the gas phase. This approach is known as a pressure-driven simulation and would typically require complex momentum transfer calculations, we took a short-cut approach. In this work we calculate the pressure in the chamber as a function of temperature and water content using an ideal gas approach (the volume of the chamber is fixed). We determine the necessary volumetric flow of humid air leaving the chamber necessary to keep the pressure in

the chamber steady at atmospheric conditions. This is calculated using the equations of a fictitious PI controller (Equations. 1 and 2) tuned to be fast but stable.

$$\dot{v}_{W+AIR}^{OUT} = K_p \varepsilon + \frac{1}{\tau_i} \int \varepsilon + \dot{v}_{W+AIR}^{IN} \quad (1)$$

$$\varepsilon = P_{CHAMBER} - P_{SP} \quad (2)$$

A second difference is the assumed driving force for the water transfer from powder to the gas phase. In our work we assume that the difference between the absolute humidity in the powder and the equilibrium humidity of the powder when exposed to the current conditions in the gas phase ( $\Delta Hum = \gamma_{bed}^{abs} - \gamma_{bed}^{@eq w air}$ ) is the drying driving force. This equilibrium condition can be obtained in the laboratory using a dynamic vapor sorption instrument. Typically, these equilibrium curves will be produced at ambient conditions, which are not reflective of the powder bed temperatures in the process. In our work we chose to consider the dependency of this equilibrium with respect to temperature. We include this in the model by measuring the water sorption isotherms at multiple temperatures and following the approach by Quirijns et. al. [4] to modify the GAB equation and thus be able to include the temperature effect on the equilibrium condition. This equilibrium is also a function of composition since each different ingredient has a different sorption curve and the composition of the bed changes as binder is added through the spray; a simple mass fraction weighted average is used.

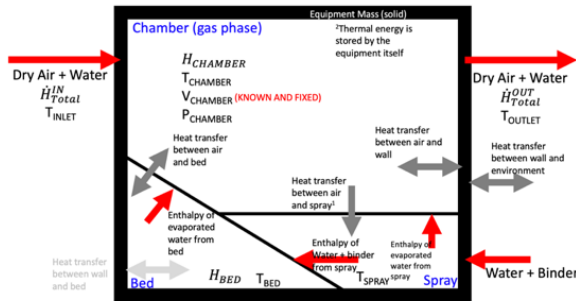
We also consider the inclusion of a spraying efficiency ( $\gamma$ ) equation (Equation 3) to capture the losses of water directly from the spray into the gas phase as a function of the absolute humidity of the powder bed. This expression introduces into the model the phenomena known as “wetting” where water permeation onto a dry powder is more difficult than onto a wet powder. We also consider the inclusion of a modifying scalar  $\omega$  to the available surface area for drying (Equation 4). This scalar is less than one during the spray phase and equal to one during drying. During the spray phase, water is entering the powder bed, and it is assumed that the portion of the total surface area of the powder that is used for water intake cannot be used for drying. The evaporation rate coefficient ( $k_{evap}$ ) is assumed to follow a first order exponential function with respect to the driving force. This allows  $k_{evap}$  to be a constant during the constant drying phase; and to decay when the drying is limited by mass transfer. This approximation Eq (5) enables the drying calculations without formally solving the intra-particle mass transfer problem.

$$\gamma = \left(1 - e^{(-\tau_s \gamma_{bed}^{abs})}\right) \quad (3)$$

$$\dot{m}_{evap bed} = SA_{particles} k_{evap} * \omega \quad (4)$$

$$k_{evap} = \begin{cases} = K_{evap} & \text{if } \Delta Hum \geq \Delta Hum_{crit} \\ = K_{evap}(1 - e^{-\tau_{evap}\Delta Hum}) & \text{if } \Delta Hum < \Delta Hum_{crit} \end{cases} \quad (5)$$

The exchanges of energy across the unit are illustrated in Figure 3, they are the same transfers considered by Ochsenbein [3].



**Figure 3.** Energy transfers across phases

One final parameter to discuss is that of a correction factor for the measured absolute humidity of the outlet air. After discussions with the equipment vendor, it became clear that value reported by this sensor needed a multiplier to be reflective of the true value. This is an additional parameter considered in the model.

### Parameter Estimability

The mass and energy portion of the model consists of 9 model parameters plus the initial wall temperature as quantities that need to be estimated from data.

**Table 1:** Model Parameters

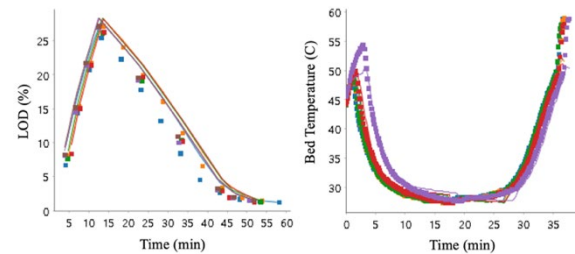
Parameter	Estimability Measure	Rank
u wall_air	6.498	1
k_evap_rate.K	3.313	2
temp wall init	1.245	3
u wall_bed	1.028	4
omega_factor	0.551	5
u bed_air	0.372	6
spraying_efficiency.tau	0.155	7
k_evap_rate.tau	0.003	8
abs.humidity out correction factor	-0.004	9
u wall_env	-1.380	10

The fluid bed granulator unit is fitted with sensors that provide electronic data captured throughout the operation. Specifically for the airflow, the temperatures and

absolute humidity for the air before and after the chamber, the temperature of the powder bed and the spray rate. Additional to these real-time measurements, samples are taken from the unit and analyzed for water content. This is reported in percent weight loss after drying (LOD). With these data available, we analyzed the estimability [5] of the model parameters to determine which parameters to estimate and which to fix. The absolute humidity correction factor and the heat transfer from wall to environment are the least estimable parameters from the data. The low estimability for the heat transfer coefficient from wall to environment was expected since the wall temperature is not measurable in the unit.

Three sets of parameters were obtained for nominal conditions that would incrementally create larger granule sizes (referred to as *under-granulation*, *nominal* and *over-granulation conditions*).

Obtained predicted vs measured values for the LOD and the bed temperature are shown in figure 4 for illustration purposes.



**Figure 4.** Predicted vs observed profiles

### Hybrid component for total particle growth

Through all the mass and energy calculations a fundamental assumption is made that the particle size of the powder is constant. This also fixes the available particle specific surface area which is essential in the calculations of drying kinetics and heat transfers to/from the bed. This mimics the approach taken by Ochsenbein et. al. [3].

A hybrid component is added to the model to account for the effect of the process conditions onto the final granule particle size. This extension adds predictability to the granule size and particle specific surface area, bonding the granule growth and drying mechanisms. A Partial Least Squares (PLS) model is built to predict the final granule properties (including particle specific surface area) as a function of the recipe parameters that also dictate the drying behaviour. The particle specific surface area is predicted with an  $R^2$  value of  $\sim 90\%$ . The hybrid model proves to be adequate to predict the time varying values for the LOD, the powder bed temperature and the humidity of the outlet air. The PLS model was built with open-source code (<https://github.com/salvadorgarciamunoz/pyphi>).

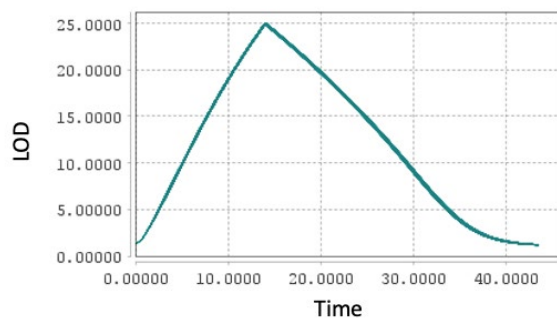
The PLS model was implemented along with the

mass and energy balance using gPROMS (Siemens Industry Software, Ltd).

## MODEL PARAMETERS EFFECT

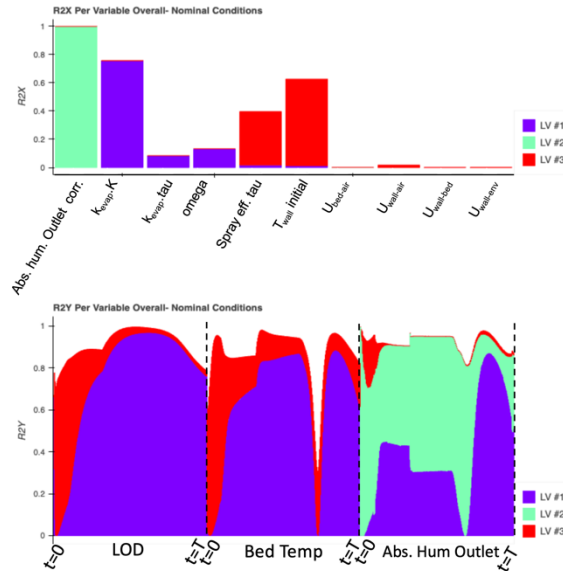
Once a model is assembled and parametrized a natural question to further the understanding of the model is to explore the effect of the model parameters onto the predictions. We explore this dependency in two ways: a) the effect of parameter uncertainty and b) The effect of true statistically significant changes in the model parameters (beyond uncertainty) onto the responses of interest.

To explore the propagation of parameter uncertainty, a sensitivity analysis is done simultaneously varying the values of each model parameter by three standard deviations above and below the optimal value obtained at the parameter estimation step. The variation of LOD due to this parametric uncertainty is shown in figure 5. The variation is  $\sim 0.125\%$  which is acceptable for the application at hand.



**Figure 5.** Variation in LOD prediction due to model parameter uncertainty.

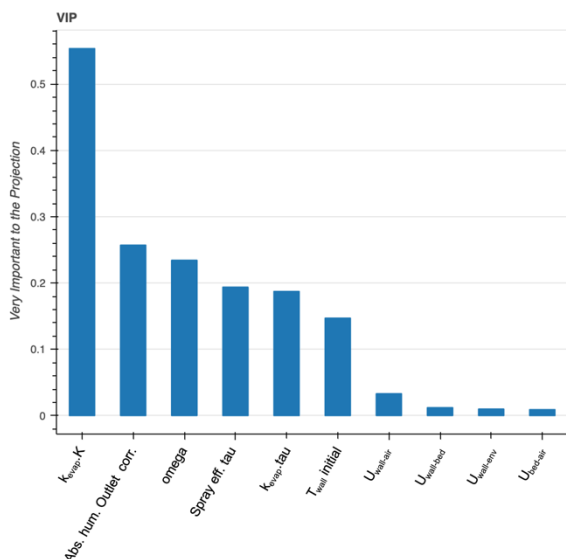
A second sensitivity study was carried out to understand the numerical effect from meaningful changes to the model parameters onto the predictions of interest: the LOD, the absolute humidity of the outlet air, and the temperature of the powder bed. This analysis was done by simultaneously varying the 10 estimated parameters to random values within 20% (above and below) of its estimated one. Ten thousand simulations were conducted for each of the three scenarios of interest (under, nominal and over granulation). A PLS model was fitted to better understand the multivariate interactions between the model parameters and the time varying profiles for the responses. Three significant latent variables were found across the three scenarios. From the coefficients of correlation (Figure 6) one can conclude that the model parameters that cause the largest change across the time varying profiles for the three model outputs are: the time constant and the gain for the evaporation rate



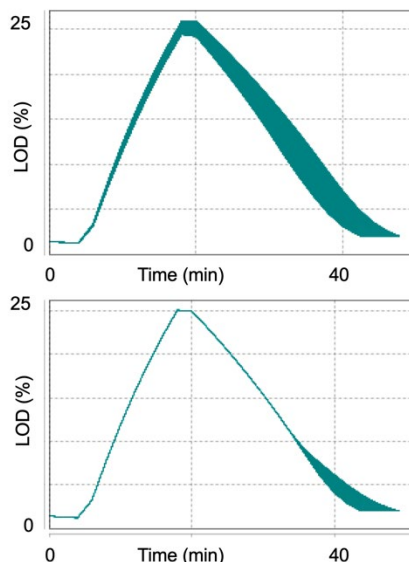
**Figure 6.**  $R^2X$  and  $R^2Y$  for the PLS model at nominal conditions fitted to meaningful variation in model parameters and its effect onto responses.

Coefficient (Equation 5), and the omega factor that controls the fraction of surface area available for drying during the spray phase. The second latent variable captures the influence of the time constant for the spray efficiency equation (Equation 3) and the initial temperature of the bed onto the LOD and the bed temperature. The third latent variable represents the effect of the outlet absolute humidity correction factor. Interestingly, this analysis also seems to indicate the almost negligible effect of the heat transfer coefficients. The trends and conclusions from Figure 6 are shared (with minor differences) across the three scenarios (*under*, *nominal* and *over* granulation). The overall ranking of importance for the model parameters can be summarized in the Very Important to the Projection (VIP) plot shown in Figure 7. This diagnostic considers the importance for each regressor accounting for the multivariate consequence onto the response. The combined effect on the LOD from of these meaningful variations in model parameters with the necessary adjustments in the process parameters to go from under to nominal to over granulation are shown in Figure 8.

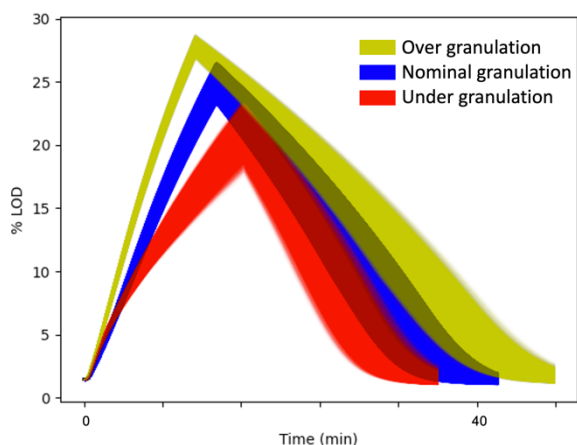
Beyond the correlations illustrated in Figure 6, the PLS model diagnostics could be further analyzed to also uncover the directionality in these correlations. This understanding is useful to determine necessary direction and magnitude for a change in a model parameter to create a positive or negative change in a model response. This information in the PLS model can be embedded with an optimization routine for example to find improved initial guesses for the model parameters for a new product.



**Figure 7.** Very Important to the Projection (VIP) plot for the PLS model fitted to meaningful variation in model parameters and its effect onto responses



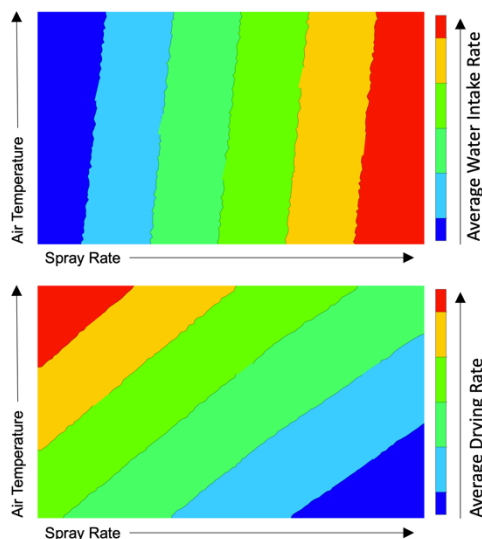
**Figure 9.** Variation in LOD due to changes in Airflow during the spraying phase (top) and drying phase (bottom)



**Figure 8.** Variation in LOD prediction due to meaningful changes in model parameters at different operating regimes.

## PROCESS PARAMETERS EFFECT

The pharmaceutical development process requires the study of the impact of all the MV in the operation of a process onto the quality of the final product. Regulatory agencies encourage practitioners to take a risk-based approach to optimize development efforts [6]. Although the granules produced from the fluid bed granulation operation can be the end-product [7]; more commonly these are intermediate materials to the final oral solid dosage form. As such, the study of this operation is customarily coupled with the study of all downstream operations in a systematic way [8]. Through a combination of experimentation and modeling described elsewhere [9] two intermediate process characteristics were identified



**Figure 10.** Sensitivity of the average water intake and drying by the powder bed with changes in air temperature and spray rate.

as key for this operation: The average rate of water intake into the powder bed and the average rate of drying. These are as indicative of the downstream performance of the granule.

Univariate studies per MV are conducted to help identify initial criticality and isolate the effect per MV. For illustrative purposes, Figure 9 shows the LOD result from variations in volumetric airflow during spraying, and during drying (as two independent MV). Remarkably, the effect of the airflow during *spraying* has a stronger effect on the LOD during the *drying* period than the airflow during *drying*. A similar effect emerges in Figure 10 showing



the effect of the air temperature (which remains constant throughout the operation) and the spray rate onto the average rate of water ingress (top) and the average drying rate (bottom). These figures show that while the spray rate dominates the effect of the water ingress to powder bed, it also has a strong effect onto the drying rate. The strong impact of the manipulated variables defining the water ingress onto the drying rate is because these variables define the initial conditions for the drying phase. This follows what other researchers have also postulated: that the dynamics of water intake and egress from the powder bed are intrinsically coupled with the dynamics of granule growth and thus the final attributes of the granule [1]. The model described here was utilized to better understand the effect of the MV in the process as they affected these two rates. These sensitivity plots can be used in the definition of design spaces once constraints are given for the rates of water intake and drying.

## CONCLUSIONS

The foundations for the assembly of a useful mathematical representation of a process are the initial articulation of the assumptions made, and the systemic enumeration of the driving forces to be considered. In a systemic approach, the latter exercise needs to be done in consideration to the available measurements that can help identify the individual contributions from the sub-systems of the model, and their behavior upon interaction. After the euphoric task of postulating equations and equations to represent a process; the modeler must subject itself to the sobering exercise of statistics, such as performing an Estimability analysis [5] to better understand what level of rigor can be afforded by the available data for the parametrization of the model. This intersection of statistical and deterministic approaches is the key to the construction of a useful model with optimal level of parsimony.

Once a model is built and well parametrized two natural questions arise. One resulting from the need to make operational decisions, and one from the desire to better understand the interaction of the different sub-systems built onto the model. *What is the effect of the process MV onto key model outputs?* and *what is the impact of the model parameters onto the key responses?* And with coupled and non-linear effects at play, the answer to both questions is often of a multivariate nature. Herein we used simple graphical approaches to support operational decisions, these of course are limited to a small number of factors. We present the use of PLS as a useful tool to better understand the multivariable relations between model parameters and responses. In the authors opinion, the visual aids and diagnostics obtained by PLS are richer and provide a more intuitive way to uncover the complex relations represented by the model structure

and its parameters, which ultimately represent the innermost fundamental understanding of the process.

## REFERENCES

1. Hayashi, K., H. Nakamura and S. Watano. Numerical study on granule aggregation and breakage in fluidized bed granulation by a novel PBM with DEM-CFD coupling approach. *Powder Technol.* 360: 1321-1336.(2020)
2. Cronin, K. and F. J. G. Ortiz. Understanding the evolution of particle size dispersion with time in a fluidised bed granulation process. *Powder Technol.* 423: 118502.(2023)
3. Ochsenbein, D. R., M. Billups, B. Hong, E. Schäfer, A. J. Marchut and O. K. Lyngberg . Industrial application of heat-and mass balance model for fluid-bed granulation for technology transfer and design space exploration. *Int. J. Pharm.: X* 1: 100028.(2019)
4. Quirijns, E. J., A. J. Van Boxtel, W. K. van Loon and G. Van Straten. "Sorption isotherms, GAB parameters and isosteric heat of sorption." *J. Sci.Food Agric.* 85(11): 1805-1814.(2005)
5. Wu, S., K. A. McLean, T. J. Harris and K. B. McAuley . Selection of optimal parameter set using estimability analysis and MSE-based model-selection criterion. *Int. J. of Adv. Mech. Syst.* 3(3): 188-197.(2011)
6. Food and Drug Administration, Pharmaceutical CGMP Initiative for the 21st Century – a Risk Based Approach, *Department of Health and Human Services.* (2004)
7. Lee, H. S., J.-J. Lee, M.-G. Kim, K.-T. Kim, C.-W. Cho, D.-D. Kim and J.-Y. Lee. Sprinkle formulations—a review of commercially available products. *Asian J. Pharm. Sci.* 15(3): 292-310. (2020)
8. Bano, G., R. M. Dhenge, S. Diab, D. J. Goodwin, L. Gorringer, M. Ahmed, R. Elkes and S. Zomer. Streamlining the development of an industrial dry granulation process for an immediate release tablet with systems modelling. *Chem.Eng. Res. Des.* 178: 421-437. (2022)
9. Sen, M. Gupta, S. Butikofer S. Developing a pharmaceutical fluid bed granulation & drying process via design of experiments based on a multivariate model. *Chem Eng.Res. Des.* Submitted.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Modeling hiPSC-to-Early Cardiomyocyte Differentiation Process using Microsimulation and Markov Chain Models

Shenbageshwaran Rajendiran<sup>a</sup>, Francisco Galdos<sup>b</sup>, Carissa Anne Lee<sup>b</sup>, Sidra Xu<sup>b</sup>, Justin Harvell<sup>a</sup>, Shireen Singh<sup>a</sup>, Sean M. Wu<sup>b</sup>, Elizabeth A. Lipke<sup>a</sup>, and Selen Cremaschi<sup>a\*</sup>

<sup>a</sup> Department of Chemical Engineering, Auburn University, Auburn, Alabama, USA

<sup>b</sup> Stanford Cardiovascular Institute, Stanford Medicine, Stanford, California, USA

\* Corresponding Author: [szc0113@auburn.edu](mailto:szc0113@auburn.edu).

## ABSTRACT

Cardiomyocytes (CMs), the contractile heart cells that can be derived from human induced pluripotent stem cells (hiPSCs). These hiPSC derived CMs can be used for cardiovascular disease drug testing and regeneration therapies, and they have therapeutic potential. Currently, hiPSC-CM differentiation cannot yet be controlled to yield specific heart cell subtypes consistently. Designing differentiation processes to consistently direct differentiation to specific heart cells is important to realize the full therapeutic potential of hiPSC-CMs. A model that accurately represents the dynamic changes in cell populations from hiPSCs to CMs over the differentiation timeline is a first step towards designing processes for directing differentiation. This paper introduces a microsimulation model for studying temporal changes in the hiPSC-to-early CM differentiation. The differentiation process for each cell in the microsimulation model is represented by a Markov chain model (MCM). The MCM includes cell subtypes representing key developmental stages in hiPSC differentiation to early CMs. These stages include pluripotent stem cells, early primitive streak, late primitive streak, mesodermal progenitors, early cardiac progenitors, late cardiac progenitors, and early CMs. The time taken by a cell to transit from one state to the next state is assumed to be exponentially distributed. The transition probabilities of the Markov chain process and the mean duration parameter of the exponential distribution were estimated using Bayesian optimization. The results predicted by the MCM agree with the data.

**Keywords:** Process Design, Derivative-free optimization, Biosystems, hiPSC cardiac differentiation

## INTRODUCTION

Cardiovascular diseases are the leading cause of death worldwide [1]. Almost 695,000 people in the US died of cardiovascular disease in 2022 [2]. Cardiomyocytes (CMs), i.e. contracting heart cells, can be derived from human induced pluripotent stem cells (hiPSCs) to be used for drug testing, regenerative therapies, and disease modeling for cardiovascular diseases [3]. The modulation of the Wnt signaling pathway using small molecules and growth factors provides the standard CM differentiation protocol, yielding CMs with up to 98% purity [4]. However, this benchmark protocol produces mostly left ventricular cells, with a little to no atrial cells, or right ventricular cells [5]. To model specific diseases in the heart requires specific heart cell types. For example,

atrial diseases like atrial fibrillation and Ebstein's anomaly, or right ventricular diseases, such as Tetralogy of Fallot or arrhythmogenic right ventricular cardiomyopathy, require atrial CMs and right ventricular CMs respectively [6].

Designing a differentiation process to consistently yield specific heart cells is essential to fully realize the therapeutic, drug testing, and disease modeling potential of CMs. The first step in designing a differentiation process that directs hiPSCs to specific heart cell subtypes is understanding the temporal changes in the cell type populations during differentiation.

Several mathematical models have been developed to capture cell differentiation process [7-16]. Stiehl et al [7] investigated if stemness (characterization of stem cells) can be defined at the single-cell level, and what

properties are suitable for defining stem cells. The authors proposed a multi-compartment deterministic model, which consists of discrete sets of ordinary differential equations, to describe the dynamics of cell differentiation and self-renewal regulated by extracellular signaling feedback. The cell behavior is characterized by parameters including the proliferative activity, the probability of differentiation, and the probability of dying. The model was developed to predict the cell type population and signaling molecule concentration over time. The model, calibrated with clinical data from multiple myeloma patients undergoing high-dose chemotherapy and stem cell transplantation, aimed to observe stem cell differentiation and determine cell type populations. Each differential equation in the model describes a distinct differentiation stage, reflecting the conventional notion that within each lineage of blood cell precursors, there is a discrete sequence of maturation stages traversed sequentially. Models presented in this paper characterize stem cells as the population that is most sensitive to environmental signals. The stem cell differentiation characterization developed in this paper identified stem cells and other cell types as a population. However, the model could not capture the differentiation mechanism happening in a single cell. The drawback of this deterministic model is that it does not capture the stochastic nature of the differentiation process.

Paździorek [8] investigated the stochastic stability of the model presented by Mar-Czohara and examined its response to noise. The deterministic model from [7] was transformed into a stochastic process using Ito differential equations. Ito calculus addresses processes with evolving random variables over time, where the random variable here is the cell type population. The random behavior of cells, whether differentiating, maturing, dying, or proliferating was modeled as a Wiener process, which is a continuous-time stochastic process. The increments of the Wiener process across disjoint time intervals exhibit a normal distribution. Premslaw demonstrated that the stochastic model achieves asymptotic stability, signifying that over time, stem cells tend to reach stable and well-defined states.

Pisu et al [9] introduced a novel mathematical model to simulate the differentiation of mesenchymal stem cells into specialized cells to study the effect of growth factors on cell proliferation/differentiation mechanisms. The model was built on material balances for extracellular matrix compounds, growth factors, and nutrients, along with a mass-structured population balance that describes cell growth, proliferation, and differentiation. The DNA content and the glycosaminoglycans (GAG) content present in the cells at different time points were used as indicators to identify the cell types. The model incorporated several parameters, with key ones being the kinetic constant of GAG synthesis, kinetic constant of collagen

synthesis, time rate of change of cells, maximum collagen and GAG concentration, concentration of  $O_2$  in saturation condition, and number of cell types. Regression analysis of the model predictions demonstrates that the average error for different differentiation pathways is below 20 percent.

The models presented in [10-16] also incorporate macroscopic analysis, lacking representation of mechanisms occurring at the single-cell level. Capturing the differentiation mechanism at the single-cell level will offer insights into interactions at the cellular level. This paper introduces a microsimulation model (MSM) to study the temporal changes in the cell type populations during cardiac differentiation at cell-level analysis. A microsimulation model, which is a stochastic model, aims to simulate individual entities of the system through stochastic parameters. Each cell is modeled as an individual entity of the MSM, and the differentiation process in each cell is represented with a continuous-time Markov chain model (CTMCM). The states of the CTMCM are cell subtypes representing the developmental stages in hiPSC differentiation to early CMs. They include pluripotent stem cells, early primitive streak, late primitive streak, mesodermal progenitors, early cardiac progenitors, late cardiac progenitors, and early CMs. The transition probabilities define the probability of a cell transitioning from one cell subtype to another or an absorbing state. The holding times, representing the time a cell spends in each state, are modeled using exponential distributions. The MSM is embedded in a Bayesian optimization framework to estimate the CTMCM parameters, i.e., transition probabilities and mean of the exponential distributions.

This paper is structured as follows: Next section discusses the modeling of cell differentiation using the Markov chain model and modeling the duration of state transitions using exponential distribution. The application of the microsimulation model for simulating the differentiation process is introduced next. The results for the microsimulation model parameters and a comparison of the model predictions to experimental data are in Results and Discussion section. Concluding remarks and future directions are given in the last section.

## MODELING DIFFERENTIATION

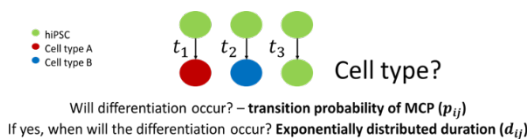
### Modeling cell differentiation

The data used to build the MSM model was obtained from the single-cell RNA sequencing analysis of 2D cardiac differentiation of two cell lines (WTC cell line and SCVI cell line) [5]. The differentiation was carried out for 6 days and samples were collected from day 1 to day 6 daily. The collected samples were captured, and single-cell RNA sequencing was performed. The gene expression matrix was analyzed and labeled to get the cell type population during differentiation from day 1 to day 6.

Based on the results obtained from 2D cardiac differentiation, the developmental trajectory of hiPSCs to early CMs follows a path. The path includes a series of states: which are pluripotent stem cells, early primitive streak, late primitive streak, mesodermal progenitors, early cardiac progenitors, late cardiac progenitors, and early CMs. At late primitive streaks and mesodermal progenitors state, cells may differentiate into non-CM cells. When a cell begins differentiation, it either moves into the next state in sequence or becomes a non-CM cell type, or remains in the same pluripotent state that it started in. There are two possible states from which cells can become non-CMs; cells from the late primitive streak can differentiate into hepatic endoderm and definitive endoderm, and cells from the mesodermal progenitor state can differentiate into endothelial cells, epicardial progenitors, and epicardial cells. Differing cell types have differing timings for differentiation, and, therefore, the time a cell spends in one state is dependent on the cell type being made. Based on this explanation, to model the cell differentiation, two questions need to be answered:

- 1) What is the next state in the differentiation path?
- 2) If a cell differentiates to the next state in sequence, how long will it stay between those two states?

The important assumptions here are the cells do not die during differentiation and they do not split or divide during the differentiation. Although we recognize the limitations of these assumptions, given the lack of data for modeling cell division and death rates for this differentiation protocol, this preliminary model allows making progress towards modeling hiPSC to early CM differentiation process.



**Figure 1.** Modeling cell differentiation.

### Continuous time Markov chain model

The first question can be answered by transition probabilities of a Markov chain model (MCM). A Markov chain model is a mathematical model that represents a sequence of events in a process in which the probability of transitioning from one state to another depends solely on the current state [17,18,19]. Markov chain models are memoryless, meaning the probability of transitioning to future states depends only on the current state and is independent of the sequence of states that led to the current state. Markov chains are used for modeling dynamic systems like financial markets, weather patterns, and language processing. In a Markov chain, the transition probability describes the probability of moving from one state

to another in the chain at each time step. The general representation of transition probability,  $P_{ij}$ , for the transition from state  $i$  to  $j$ , is given in Equation (1).

$$P_{ij} = P(X_{n+1} = j | X_n = i) \quad (1)$$

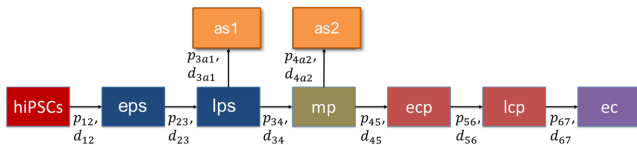
In Equation 1,  $X_n$  and  $X_{n+1}$  represent two consecutive states in a Markov chain. The RHS of Equation (1) represents the conditional probability that, given the current state  $X_n = i$ , the system will transition to state  $X_{n+1} = j$ . Markov chains are characterized by a transition matrix that represents the probabilities of moving from one state to another. The transition matrix enables predictions and analysis of system behavior over time.

Markov chain models can be classified into different types based on time homogeneity (homogenous Markov chain and non-homogenous Markov chain), state continuity (finite state Markov chains and continuous state Markov chains), time continuity (discrete-time Markov chains and continuous time Markov chains), absorbing or non-absorbing states (absorbing Markov chains and non-absorbing Markov chains). [20]. The classifications related to cell differentiation based on scRNA-seq data are discrete-time and absorbing Markov chains. In a discrete-time Markov chain, the transitions between states occur at discrete, evenly spaced time intervals. An absorbing Markov chain is used to model systems where certain states, known as absorbing states, act as final destinations from which there is no escape. This means that in a Markov chain, the transitions between states occur with probabilities, however, once the system reaches an absorbing state, it remains there indefinitely. In other words, the probability of transition from an absorbing state to any other non-absorbing state is zero. These chains can be characterized by their probability of transition between transient states (non-absorbing states) and absorbing states, leading to absorption, which represents the final resting places of the system.

A continuous-time absorbing Markov chain model is developed to represent hiPSC to CM cell differentiation. Figure 2 demonstrates the Markov chain model of cell differentiation. Late primitive streak can differentiate into endoderm cells, and mesodermal cells can differentiate into epithelial and epicardial cells. When cells commit to becoming endodermal, epicardial, or epithelial cell types, they can no longer differentiate into CMs. Hence, the transition probability from these cell types to any other cell types in the cardiac trajectory is zero. Two absorbing states, one from the late primitive streak and the other from the mesodermal progenitor state, were included in the MCM to denote cells that differentiate into non-CMs. Absorbing state 1 consists of hepatic endoderm cells and definitive endoderm cells, and absorbing state 2 consists of epithelial cells, epicardial progenitors, and epicardial cells.

The probability  $P_{ij}$  denotes the probability of

transition of the MCM from state  $i$  to state  $j$  in the cardiac differentiation trajectory and  $p_{ij}$  denotes the transition probability in a specific instance of simulating a Markov chain.  $P_{ij}$  is the Markov chain model parameter and  $p_{ij}$  is the probability of transitioning from state  $i$  to state  $j$  during a simulation run. In the simulation, for cell types other than late primitive streak and mesodermal progenitors, if  $p_{ij} \leq P_{ij}$ , the cell will differentiate from cell type  $i$  to cell type  $j$ . If a cell is initially in the late primitive streak or mesodermal progenitors state, there are three potential outcomes for the cell's eventual destination. If  $p_{ij} \leq P_{ij}$ , the cell will differentiate into cell type  $j$ , or if  $P_{ij} < p_{ij} \leq P_{ij} + P_{ia}$ , the cell will differentiate into absorbing state, or if  $P_{ij} + P_{ia} < p_{ij}$ , the cell will stay in the same state  $i$ . Based on the transition probability parameters of the MCM, the cells in the simulation will differentiate into different cell types. The Markov chain stops when the cell reaches either an absorbing state or an early CM state.



**Figure 2.** Markov chain representation of cell differentiation (Table 1 defines cell type abbreviations).

**Table 1:** Expansion of the cell types in the MCM.

Abbreviations	Cell types
hiPSCs	Human induced pluripotent stem cells
eps	Early primitive streak
lps	Late primitive streak
as1	Absorbing state 1
mp	Mesodermal progenitors
as2	Absorbing state 2
ecp	Early cardiac progenitors
lcp	Late cardiac progenitors
ec	Early cardiomyocytes

### Modeling state duration in the MCM with exponential distribution

The exponential distribution is a continuous probability distribution commonly used to model the time between events in a Poisson process, where events occur at a constant average rate and are independent of each other [21]. The probability density function  $f(x)$  of the exponential distribution is given by Equation 2.

$$f(x) = \lambda \exp(-\lambda x) \quad (2)$$

In Equation 2,  $x$  is a non-negative variable representing the time between events, and  $\lambda$  is the rate parameter, which is a positive constant. It determines the

average number of events occurring per unit of time and is also equal to the inverse of the average time between events. The rate parameter quantifies the event rate, the frequency, on average, of the events occurring. One of the notable properties of the exponential distribution is memorylessness. It suggests that the probability of the next event occurring in the next time increment is the same, regardless of how much time has already passed. This property of exponential distribution helps modeling systems with no memory of past events.

In the cell differentiation process, the time a cell spends between two states is random, with a constant average rate of occurrence and the events are independent of each other. Therefore, the duration spent by a cell between two states can be modeled using the exponential distribution. The average duration  $d_{ij}$ , spent by the cells between state  $i$  and state  $j$  is the inverse of the rate parameter used to model the time spent by the cells between states.

## MODELING CELL POPULATIONS

### Microsimulation models

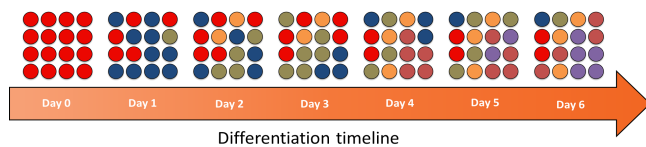
Microsimulation models (MSM) are computational models designed to simulate individual-level behavior using individual entities within a population [22,23]. Each entity is represented with specific attributes. In the case of modeling cell populations, the individual entity denotes each cell in the differentiation process, and the attributes denote transition probability and the average duration parameters. MSM consists of three components, the agents, rules, and environment. The agents represent individual cells. The rules denote the behavior and decision-making processes for each cell, which means the transition of cells into different states that are controlled by the Markov chain probabilities. The environment represents the context in which the cells interact. The constraint in the MSM is the sum of probability in each state should add up to one.

In the developed MSM (Figure 3), each cell in the simulation will move into the next state based on the transition probability of the Markov chain model, and whenever the cell moves to the next state, the duration the cell spends in that state is sampled from the exponential distribution. For transition from states  $i$  to  $j$ , excluding absorbing states, the sampled time interval is calculated based on the average duration parameter,  $\lambda_{ij}$ . The age associated with the cell in a state is obtained by adding the sampled time for that state to the age of the cell in the previous state. Since we are modeling the cell population at the end of each differentiation day, the result obtained from the simulation is converted into a binary matrix, in which the rows represent the differentiation days, and the column represents the cell states. For example, let us consider that the age of a cell when it

**Table 2:** Cell population matrix from WTC cell line ( $X_1$ ) used in the optimization problem (Equation 3). All numerical values are expressed as percentages.

Days	hiPSCs	eps	lps	as1	mp	as2	Ecp	lcp	Ec
Day 1	18.9	80.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0
Day 2	0.6	8.3	90.8	0.0	0.3	0.0	0.0	0.0	0.0
Day 3	0.8	1.8	22.8	17.5	50.6	0.0	6.5	0.0	0.0
Day 4	0.4	0.0	1.9	30.2	20.7	4.1	42.6	0.2	0.0
Day 5	0.8	0.0	0.8	31.7	5.8	7.2	36.5	17.3	0.0
Day 6	0.3	0.0	0.0	39.7	0.8	9.2	17.0	31.3	1.7

reaches the early primitive streak is 0.78 days, and the age of a cell when it reaches the late primitive streak is 2.4 days and it stays in the late primitive streak. In the given example, at the end of day 1, the cell state is the early primitive streak, so in the binary matrix, in the day 1 row, the eps column will have a value of 1. From 0.78 days to 2.4 days, the cell remains in the early primitive streak state, meaning at the end of day 2, the cell in the eps column will have a value of one, and the rest of the day 2 row will have a value of 0. From 2.4 days up to day 6, the cell will stay in the late primitive streak, and therefore, for the rows from day 3 to day 6, the lps column will have a value of 1, and the rest equal to zero. Five thousand cells were run in the simulation to generate the cell type matrix for each cell. The number of cells in each cell type at the end of each differentiation day was added and then normalized by the total number of cells. By this procedure, the cell population percentage matrix can be obtained using the microsimulation model with the Markov chain process.



**Figure 3.** Modeling cell population using a microsimulation model.

### Parameter estimation

The MSM parameters, transition probabilities, and the average durations were estimated using Bayesian optimization by minimizing the mean squared error between the model output and the experimental data. Bayesian optimization, driven by Bayesian inference and surrogate modeling, is a versatile and efficient approach for solving complex optimization problems [24,25]. Bayesian optimization uses a surrogate model, typically a Gaussian process (GP) model, to estimate the unknown objective function. Bayesian optimization balances exploration (sampling in uncertain regions), and exploitation (sampling in the regions with the highest estimated

objective value), to find the global optimum efficiently. The objective function (Equation 3) used is the mean squared error between the model output and the experimental data.

$$f = \frac{1}{n}((Y - X_1)^2 + (Y - X_2)^2) \quad (3)$$

$$Y = MSM(P_{ij}, d_{ij}) \quad (4)$$

$$P_{lps,mp} + P_{lps,a1} \leq 1 \quad (5)$$

$$P_{mp,a2} + P_{mp,ecp} \leq 1 \quad (6)$$

$$0 \leq P_{ij} < 1; d_{ij} \geq 0 \quad (7)$$

$Y$  is the MSM output. The variables  $X_1$  and  $X_2$  represent the cell population matrices obtained from the single-cell RNA sequencing of the cell samples. These samples were collected from two cell lines during differentiation days 1-6, as mentioned in the 'Modeling cell differentiation' section. The cell population matrix for WTC cell line,  $X_2$ , is shown in Table 2 as an example. Equation 4 states that the model output is a function of the decision variables  $P_{ij}$  and  $d_{ij}$ . To ensure the probabilities at the cell states having two states always less than or equal to 1, constraints (Equations 5, 6 and 7) were included in the model.

Bayesian optimization begins with generating an initial set of samples from the model. The Gaussian Process (GP) model is built based on the initial samples. The acquisition function selects the next sample for evaluation, balancing the exploration and exploitation. The selected point is evaluated in the true objective function. The surrogate model is updated with new data and the process iterates.

## RESULTS AND DISCUSSION

Five thousand cells were simulated in the Markov chain model embedded within the microsimulation model. The parameters obtained using Bayesian optimization are shown in Table 3. Based on the data, the optimization result showed that the probability of transitioning from psc state to eps state, and eps state to lps state is 1. During these two transitions, the cells spend an



average duration of 0.45 days and 1.32 days, respectively. At the lps state, 16% of cells transition into hepatic endoderm and definitive endoderm cells. These 16% of cells never become CMs. The remaining 84% of the cells take about 1.3 days to enter the mesoderm state, where 42% of these cells (35% of the initial number of cells) become endothelial cells, epicardial progenitors, and epicardial cells. The remaining 42% of the cells take about one day to enter the early cardiac progenitor state. After reaching the early progenitor state, 69% of these cells become late cardiac progenitors. Following this, 31% of the late progenitors become early cardiomyocytes. These cells spend about 3.58 days to become late cardiac progenitors and 2.98 days to become early CMs.

**Table 3:** MSM parameters obtained using Bayesian optimization.

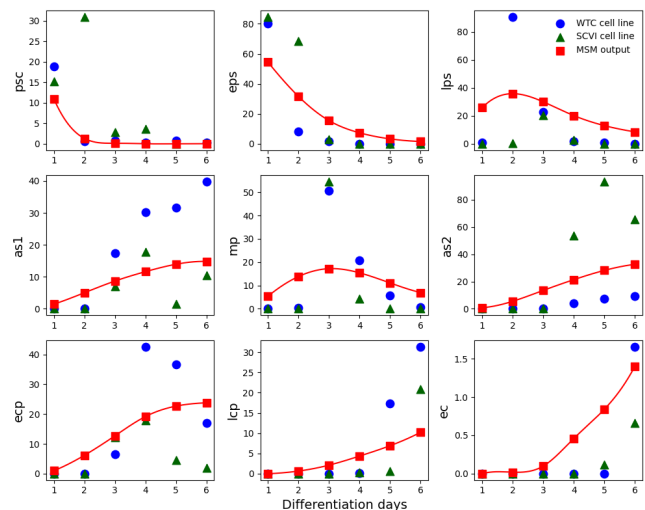
Current state	Next state	Transition probability	Average duration
hiPSCs	eps	1	0.45
eps	lps	1	1.32
lps	mp	0.81	1.30
lps	as1	0.16	-
mp	ecp	0.52	0.98
mp	as2	0.42	-
ecp	lcp	0.69	3.58
lcp	ec	0.31	2.98

Figure 4 shows the predicted cell population versus the experimental data. From the plot, we can see that there are differences between the WTC cell line data and the SCVI cell line data. Since the aim of this work is to develop a model that is independent of the cell line, the effect of the cell line is not introduced in the model. This can also be observed in Figure 4, where the predicted cell type populations by the MSM model are between the two datasets. The mean squared error and the  $R^2$  between the model and the data are 0.004 and 0.43, respectively. The performance of the developed model will undergo validation with further experimental data in the future.

## CONCLUSION AND FUTURE DIRECTIONS

Using the cardiac cell type population obtained from 2D cardiac differentiation, a microsimulation model with a Markov chain model was developed to predict the cardiomyocyte cell type population during differentiation. The model parameters were obtained using Bayesian optimization. The results showed that the simulation model predicted the cell type population with a low mean square error. The developed model can act as a digital twin to the cardiac differentiation experiment. The data for training this model was generated in a 2D environment

following the modulated Wnt signaling differentiation protocol, which mostly yields ventricular cells. This protocol will be tested to analyze if the ventricular cells can be produced on a large-scale using 3D techniques. The model will also be trained with data using the 3D differentiation protocol. If different sets of experimental parameters are used to obtain different cell-type populations, the model can be trained on those datasets, and a set of transition probabilities and average duration parameters can be estimated for each set of experimental parameters. Finding the relationship between the experimental parameters and the model parameters will help tune the experimental parameters to get the desired CM percentage on a certain differentiation day.



**Figure 4.** MSM predicted cell population vs differentiation experiment.

## ACKNOWLEDGEMENT

This work was funded by NSF grant # 2135059.

## REFERENCES

- Ahmed, R. E., Anzai, T., Chanthra, N., & Uosaki, H. A Brief Review of Current Maturation Methods for Human Induced Pluripotent Stem Cells-Derived Cardiomyocytes. *Front Cell Dev Biol*, 8, 178 (2020).
- <https://www.cdc.gov/mmwr/volumes/71/wr/mm717e1.html>
- Li, J., Hua, Y., Miyagawa, S., Zhang, J., Li, L., Liu, L., & Sawa, Y. hiPSC-derived cardiac tissue for disease modeling and drug discovery. *International journal of molecular sciences*, 21(23), 8893 (2020).
- Tani, H., Tohyama, S., Kishino, Y., Kanazawa, H., & Fukuda, K. Production of functional cardiomyocytes and cardiac tissue from human induced pluripotent stem cells for regenerative therapy. *J Mol Cell Cardiol*, 164, 83-91 (2022)

5. Galdos, F. X., Lee, C., Lee, S., Paige, S., Goodyer, W., Xu, S., ... & Wu, S. (2023). Combined lineage tracing and scRNA-seq reveals unexpected first heart field predominance of human iPSC differentiation. *Elife*, 12, e80075.
6. Basso, C., Corrado, D., Marcus, F. I., Nava, A., & Thiene, G.). Arrhythmogenic right ventricular cardiomyopathy. *The Lancet*, 373(9671), 1289-1300 (2009).
7. Stiehl, T., & Marciniak-Czochra, A. (2011). Characterization of stem cells using mathematical models of multistage cell lineages. *Mathematical and Computer Modelling*, 53(7-8), 1505-1517.
8. Paździorek, P. R. Mathematical model of stem cell differentiation and tissue regeneration with stochastic noise. *Bulletin of mathematical biology*, 76, 1642-1669 (2014).
9. Pisu, M., Concas, A., & Cao, G. A novel simulation model for stem cells differentiation. *Journal of biotechnology*, 130(2), 171-182 (2007).
10. Marciniak-Czochra, A., Stiehl, T., Ho, A. D., Jäger, W., & Wagner, W. (2009). Modeling of asymmetric cell division in hematopoietic stem cells—regulation of self-renewal is essential for efficient repopulation. *Stem cells and development*, 18(3), 377-386.
11. Till, J. E., McCulloch, E. A., & Siminovitch, L. (1964). A stochastic model of stem cell proliferation, based on the growth of spleen colony-forming cells. *Proceedings of the National Academy of Sciences*, 51(1), 29-36.
12. Yakovlev, A. Y., Mayer-Proschel, M., & Noble, M. (1998). A stochastic model of brain cell differentiation in tissue culture. *Journal of mathematical biology*, 37, 49-60.
13. Prokharau, P. A., Vermolen, F. J., & García-Aznar, J. M. A mathematical model for cell differentiation, as an evolutionary and regulated process. *Computer Methods in Biomechanics and Biomedical Engineering*, 17(10), 1051-1070 (2014).
14. Jilkine, A. Mathematical models of stem cell differentiation and dedifferentiation. *Current Stem Cell Reports*, 5, 66-72 (2019).
15. Doumic, M., Marciniak-Czochra, A., Perthame, B., & Zubelli, J. P. A structured population model of cell differentiation. *SIAM Journal on Applied Mathematics*, 71(6), 1918-1940 (2011).
16. Tu, X., Zhang, Q., Zhang, W., & Zou, X. Single-cell data-driven mathematical model reveals possible molecular mechanisms of embryonic stem-cell differentiation. *Mathematical Biosciences and Engineering*, 16(5), 5877-5896 (2019).
17. Gagniuc, P. A. *Markov chains: from theory to implementation and experimentation*. John Wiley & Sons (2017).
18. Tolver, A. *An introduction to Markov chains*. Department of Mathematical Sciences, University of Copenhagen. (2016).
19. Wasserman, L. *All of statistics: a concise course in statistical inference* (Vol. 26, p. 86). New York: Springer (2004).
20. Kampen, N. G. edition 3. *Stochastic processes in physics and chemistry* (2007).
21. Balakrishnan, K.. *Exponential distribution: theory, methods and applications*. Routledge (2019).
22. O'Donoghue, C. (Ed.). *Handbook of microsimulation modelling*. Emerald Group Publishing (2014).
23. Krijkamp, E. M., Alarid-Escudero, F., Enns, E. A., Jalal, H. J., Hunink, M. M., & Pechlivanoglou, P. Microsimulation modeling for health decision sciences using R: a tutorial. *Medical Decision Making*, 38(3), 400-422 (2018).
24. Frazier, P. I. A tutorial on Bayesian optimization. arXiv preprint arXiv:1807.02811 (2018).
25. Frazier, P. I. Bayesian optimization. In *Recent advances in optimization and modeling of contemporary problems* (pp. 255-278). *Informatics* (2018).

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Integrating Hybrid Modeling and Multifidelity Approaches for Data-Driven Process Model Discovery

Suryateja Ravutla<sup>a</sup>, and Fani Boukouvala<sup>a\*</sup>

<sup>a</sup> Department of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA

\* Corresponding Author: [fani.boukouvala@chbe.gatech.edu](mailto:fani.boukouvala@chbe.gatech.edu)

---

## ABSTRACT

Modeling the non-linear dynamics of a system from measurement data accurately is an open challenge. Over the past few years, various tools such as SINDy and DySMHO have emerged as approaches to distill dynamics from data. However, challenges persist in accurately capturing dynamics of a system especially when the physical knowledge about the system is unknown. A promising solution is to use a hybrid paradigm, that combines mechanistic and black-box models to leverage their respective strengths. In this study, we combine a hybrid modeling paradigm with sparse regression, to develop and identify models simultaneously. Two methods are explored, considering varying complexities, data quality, and availability and by comparing different case studies. In the first approach, we integrate SINDy-discovered models with neural ODE structures, to model unknown physics. In the second approach, we employ Multifidelity Surrogate Models (MFSMs) to construct composite models comprised of SINDy-discovered models and error-correction models.

---

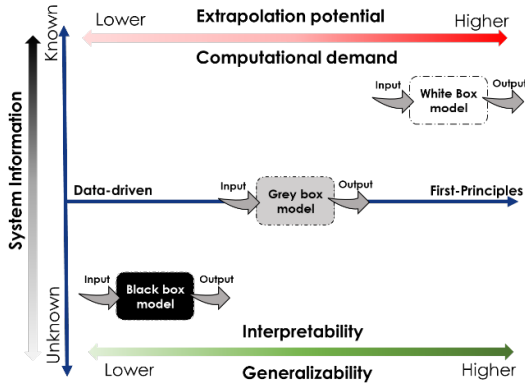
**Keywords:** Data-driven modeling, Model identification, Hybrid modeling, Multifidelity, Sparse regression

## INTRODUCTION

Present modeling approaches of intricate dynamic systems rely on ordinary and/or partial differential equations (ODEs, PDEs) to describe their behaviors. These governing equations are conventionally obtained from rigorous first principles like conservation laws or derived from phenomenological knowledge-based approaches. However, many dynamic systems remain unexplored, lacking comprehensive analytical descriptions. Exploiting advances in data acquisition, digitization, and storage, data-driven 'black box' models have emerged as an alternative [1, 2]. These data-driven methods excel in regression and classification tasks, yet their resultant black-box nature commonly lacks physical insight and exhibits limitations in extrapolation beyond the training data's boundaries. Certain systems exist, such as biological processes or intricate design problems, that necessitate a deeper understanding of governing equations. Consequently, recent developments focus on the integration of data-driven techniques into modeling system dynamics. Early attempts utilized symbolic regression [3] and genetic programming algorithms [4]. However,

challenges like overfitting and the computational demands arising from their combinatorial nature limited their applicability to low-dimensional systems and small initial candidate sets of symbolic expressions. An alternative approach called SINDY [5] was proposed, that reconstructs the underlying equations based on a large-space library of candidate terms and transforming the discovery problem to sparse regression and over time, several extensions to SINDY and alternative approaches on similar ideas were also introduced [6-9]. More recently, another such approach (DySMHO) [10] was proposed that uses moving horizon optimization for identifying the governing equations. Nevertheless, the effectiveness of these techniques relies on the identification and selection of terms from an array of potential candidate terms. While an exhaustive pool covering all potential terms might aid in building a highly accurate mechanistic model, this presents a challenge, especially in cases involving complex nonlinear systems and when the system dynamics are not fully known. A potential solution is to use hybrid modeling techniques with these methods. These hybrid models (HMs), also referred to as 'grey-box' models have the advantages of both mechanistic and the

black-box models [1, 11]. Figure 1 shows the comparison between the mechanistic, grey- and black-box models. Early implementations of these hybrid models can be found in works that date back to 1980's.



**Figure 1:** A comparison between white-box/mechanistic, grey-box and black-box models.

In this work, we combine a hybrid modeling paradigm with sparse regression, with the goal of simultaneous hybrid model development and model identification. Specifically, in our study, we investigate two distinct methods. In the first method, we employ the SINDy formulation to establish the initial physics-based model from the incomplete candidate library. Subsequently, we integrate this with neural ordinary differential equations (NODEs) [12-14], to improve the accuracy of the final model. Leveraging the NODE formulation to model unknown or missing physics within a mechanistic model has previously proven to be successful [12, 15, 16]. For the second approach, we employ composite structures known as Multifidelity Surrogate Models (MFSMs) [17, 18] using true data/high-fidelity (HF) and the low-fidelity (LF) data. The model output from SINDy constructed using the incomplete candidate library is treated as the LF model. In the subsequent step, we develop MFSMs that use HF and LF data to refine the model accuracy.

In the following sections of this paper, we introduce the methods for our proposed HM approaches and a workflow to build the HMs. In the subsequent sections, we utilize two non-linear case studies to test our HM models and show their prediction accuracy. Furthermore, we show our analysis on the impact of sampled data density and noise on the accuracy of the HMs, as well as the HMs extrapolation capabilities.

## 2. METHODS

### 2.1 Sparse identification of nonlinear dynamics

In [6], the authors leveraged the fact that most physical systems have only a few relevant terms that

define the dynamics, making the governing equations sparse in a high-dimensional nonlinear function space. Consider the example first-order ODE system with  $n$  states which are denoted by  $\mathbf{X}$ . We denote the derivative with respect to time for  $\mathbf{X}$  as  $\mathbf{X}'$ . The right-hand side of the ODE is given by the derivative with respect to time,  $t$ , denoted by the function  $f$ ,

$$\mathbf{X}' = \frac{d\mathbf{X}}{dt} = f(\mathbf{X}) \quad (1)$$

Next, a library  $\theta(\mathbf{X})$ , consisting of candidate nonlinear functions of the columns of  $\mathbf{X}$  is constructed. For example, the library may consist of constant, polynomial, and trigonometric terms. Finally, a sparse regression problem is set up to determine the sparse vectors of coefficients  $\boldsymbol{\varepsilon} = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n]$ , which determines the nonlinearities that are active.

$$\theta(\mathbf{X}) = \begin{bmatrix} | & | & | & | & \dots & | & | & | \\ 1 & \mathbf{X} & \mathbf{X}^2 & \mathbf{X}^3 & \dots & \sin(\mathbf{X}) & \cos(\mathbf{X}) & \dots \\ | & | & | & | & & | & | & | \end{bmatrix} \quad (2)$$

$$\mathbf{X}'_{SINDy} = \theta(\mathbf{X})\boldsymbol{\varepsilon} \quad (3)$$

$$\mathbf{X}_{SINDy} = \text{ODEsolver}(\theta(\mathbf{X})\boldsymbol{\varepsilon}, \mathbf{X}_0, t) \quad (4)$$

Discovering the mechanistic equations from state data with SINDy is subject to a) estimating the derivatives accurately with data limitations and b) formulating the candidate library  $\theta(\mathbf{X})$  to contain all the terms that could potentially form an accurate mechanistic equation. While several extensions to SINDy have been proposed over the years to address the former issue, the latter is still a challenge. This is especially a major issue when we are dealing with state data from non-linear systems where the potential terms in the candidate library are not known. To address this, we propose to use the following hybrid modeling approaches.

### 2.2 Neural Ordinary Differential Equations for Error Correction

Neural networks (NNs) are one of the widely used ML models in data-driven modeling due to their ability to approximate complex nonlinear relationships. Since the early 90s, NNs have been used to model dynamic systems within differential equations [19]. Recently, NODEs have emerged, integrating NNs with automatic differentiation tools [12-14]. NODEs predict system derivatives directly during training, capturing both state and derivative data. The potential of this approach was shown in better capturing curvature in dynamic data when compared to data-driven models that ignore derivative information. A neural ODE is essentially a NN used to model  $f(\mathbf{X})$  from Eq (1). We denote the modeled derivative by  $\mathbf{X}'_{NN}$ . We can obtain the predicted ODE solution  $\mathbf{X}_{NN}$  by employing any preferred ODE solver.

$$\mathbf{X}'_{NN} = \frac{d\mathbf{X}_{NN}}{dt} = NN(\mathbf{X}) \quad (5)$$

$$\mathbf{X}_{NN} = \text{ODEsolver}(\text{NN}, \mathbf{X}_0, t) \quad (6)$$

Training the NODE can be done either by minimizing the error between the predicted and true states  $\text{Error} = \text{MSE}(\mathbf{X}, \mathbf{X}_{NN})$  by passing the NN through an ODE solver [14] or, utilize a collocation-based approach [13] and avoid using explicit ODE solver by minimizing the loss function  $\text{Error} = \text{MSE}(\mathbf{X}'_{est}, \text{NN}(\mathbf{X}_{est}))$ . In [13], the authors show that by using the collocation based approach, training a NODE is faster. We utilize this collocation-based approach to train our HM with NODE formulation for model identification from state data. In the first step, we generate a low-fidelity (LF) mechanistic model using SINDy by assuming few terms in the candidate library  $\theta(\mathbf{X})$ . In the next step, we correct the error resulting from the LF model with a NODE. We utilize the following formulation shown in Eq (7).

$$\mathbf{X}'_{HM} = \frac{d\mathbf{X}_{HM}}{dt} = \theta(\mathbf{X})\boldsymbol{\varepsilon} + \text{NN}(\mathbf{X}) \quad (7)$$

$$\mathbf{X}_{HM} = \text{ODEsolver}(\theta(\mathbf{X})\boldsymbol{\varepsilon} + \text{NN}(\mathbf{X}), \mathbf{X}_0, t) \quad (8)$$

To train the NODE we take the collocation-based approach and minimize the error between the derivatives directly. In the next step, we calculate the difference between the derivatives values between the HF and the LF derivatives at the sampled data. This derivative difference corresponds to the mismatch between the true model and the LF model. A NN model is then trained to predict this difference when given the true state data. Finally, the hybrid model states  $\mathbf{X}_{HM}$  are estimated by using an ODE solver.

### 2.3 Multifidelity Surrogate Models for Error Correction

In recent years, hybrid composite structures that can learn from both HF and LF data were proposed to improve the LF model predictions by correcting the error between the HF and LF data. These composite structures are referred to as multi-fidelity surrogate models (MFSMs) [17, 18, 20]. A widely used structure of MFSMs is  $y_H = \rho(x)y_L + \delta(x)$ , where  $y_L, y_H$  represent the low and high-fidelity data respectively,  $\rho(x)$  is multiplicative correlation surrogate and  $\delta(x)$  is the additive surrogate. It can be re-written as  $y_H = F(x, y_L)$ . To establish a connection between the HF and LF data, it's necessary to have both a HF model and a LF model that can produce this data. Similar to the NODE model correction approach, we use SINDy to obtain the LF model. The LF model from SINDy is then integrated using an ODE solver to obtain  $\mathbf{X}_{SINDy}$  using Eq (4). We then formulate the MFSM by utilizing a NN to model the error. The NN model is trained to minimize the loss function  $\text{Error} = \text{MSE}(\mathbf{X}, \mathbf{X}_{MFSM})$ , and takes both time and LF states as input to predict  $\mathbf{X}_{MFSM}$ .

$$\mathbf{X}_{MFSM} = \text{NN}(t, \mathbf{X}_{SINDy}) \quad (9)$$

### 2.3 Workflow for constructing hybrid models

We utilize the workflow shown in Table (1) for constructing the hybrid models under approach 1 and approach 2. Here  $\mathbf{X}_{HF}$  represent the true state data.  $\beta_1, \beta_2, \beta_3$  are the regularization coefficients.

**Table 1:** Workflow for constructing the hybrid models to correct error for NODE and MFSM formulations.

Let $[t, \mathbf{X}_{HF}]$ be the complete HF dataset	
Generate SINDy model	<ol style="list-style-type: none"> <li>1. Set <math>t \leftarrow</math> Input and <math>\mathbf{X}_{HF} \leftarrow</math> HF output</li> <li>2. Estimate the derivatives <math>\mathbf{X}'_{HF}</math> from <math>[t, \mathbf{X}_{HF}]</math> data.</li> <li>3. Generate the feature library <math>\theta(\mathbf{X})</math> and set the optimizer</li> </ol> <p><i>While</i> termination criteria <b>not true</b>:</p> $\min( \mathbf{X}'_{HF} - \theta(\mathbf{X})\boldsymbol{\varepsilon} ^2) + \beta_1 (\boldsymbol{\varepsilon})$ <p>Tune <math>\boldsymbol{\varepsilon}</math></p>
Set $\mathbf{X}'_{LF} = \theta(\mathbf{x})\boldsymbol{\varepsilon}$ and generate data, $\mathbf{X}_{LF} \leftarrow$ LF output.	
<i>If</i> Approach 1: Correct LF model error with NODE formulation	<ol style="list-style-type: none"> <li>1. Initialize NN, Set <math>\mathbf{X}_{HF} \leftarrow</math> Input and <math>\Delta\mathbf{X}' = \mathbf{X}'_{HF} - \mathbf{X}'_{LF} \leftarrow</math> output</li> <li>2. <i>While</i> termination criteria <b>not true</b>:</li> </ol> $\min( \Delta\mathbf{X}' - \mathbf{X}'_{NN} ^2) + \beta_2 \ \Phi_{NN}\ _2$ <p>Set <math>\mathbf{X}'_{HM} = \theta(\mathbf{X})\boldsymbol{\varepsilon} + \text{NN}(\mathbf{X}) \leftarrow</math> final model</p>
<i>If</i> Approach 2: Correct LF model error with MFSM formulation	<ol style="list-style-type: none"> <li>1. Initialize a NN, Set <math>[t, \mathbf{X}_{LF}] \leftarrow</math> Input and <math>\mathbf{X}_{HF} \leftarrow</math> output</li> <li>2. <i>While</i> termination criteria <b>not true</b>:</li> </ol> $\min( \mathbf{X}_{MFSM} - \mathbf{X}_{HF} ^2) + \beta_3 \ \Phi_{NN}\ _2$ <p>Set <math>\mathbf{X}_{MFSM} = \text{NN}(t, \mathbf{X}_{LF}) \leftarrow</math> final model</p>

## 3. RESULTS AND DISCUSSION

In this section, we first introduce the case studies that we use to test and compare the approaches from 2.1 – 2.3. In the subsequent sections, we show the analysis on the model accuracy, effect of density of data and noise on the model accuracy and the extrapolation ability of the models.

### 3.1 Case Studies

To test the two approaches, we utilize two non-linear case studies. The first case study is a non-isothermal continuously stirred tank reactor (CSTR) problem, and the second case study is a penicillin biosynthesis problem.

#### 3.1.1 The non-isothermal CSTR problem

The CSTR's governing equation, Eq (11), represents mass conservation. But, in non-isothermal operating scenarios which is common in practical applications, the energy balance must also be considered. Thus, temperature is added as an extra state variable, along with obtaining temperature measurements in addition to composition data, as depicted in Eq (12). We utilize the model equations and assumptions from [10].

$$\frac{dC_A}{dt} = \frac{q}{V}(C_{A,i} - C_A) - k_0 e^{-\frac{E_a}{RT}} C_A \quad (11)$$



$$\frac{dT}{dt} = \frac{q}{V}(T_i - T) + \frac{(-\Delta H_R)}{\rho c} k_0 e^{-\frac{E_a}{RT}} C_A(t) + \frac{U_A}{V\rho c}(T_c - T) \quad (12)$$

### 3.1.2 The penicillin biosynthesis problem

For the second case study, we chose to model the production of penicillin via yeast fermentation. The level of nonlinearity in the system differs significantly between state variables. The process is modeled by four differential equations on the following states: volume ( $V$ ), concentrations of biomass ( $B$ ), product ( $P$ ) and substrate ( $S$ ). The system of ODEs is defined as shown in Eq (14) and Eq (15). The parameter values have been taken from [12].

$$\begin{aligned} \frac{dB}{dt} &= B(\mu - D - c_L) \\ \frac{dS}{dt} &= -\sigma B + (S_f - S)D \\ \frac{dP}{dt} &= q_p B - P(D + c_1 k) \\ \frac{dV}{dt} &= F \end{aligned} \quad (13)$$

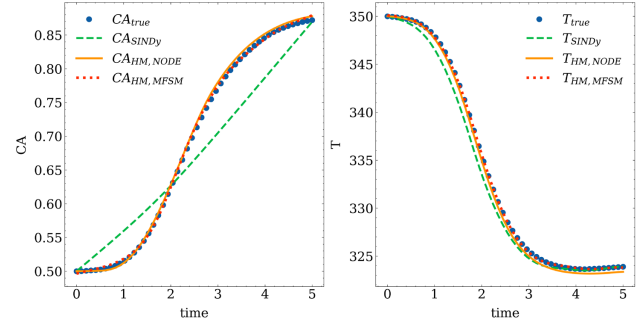
$$\begin{aligned} \mu &= \frac{\mu_m S}{k_x X + 10} \\ \sigma &= \frac{\mu}{Y_x} + \frac{q_p}{Y_p} + m_x \\ q_p &= \frac{1.5 q_{pm} S B}{4k_p + B S \left[1 + \frac{S}{3k_i}\right]} \\ c_L &= \frac{c_{Lmax} B \exp\left(-\frac{S}{100}\right)}{K_L + B + 1} \\ m_x &= m_{xm} \frac{B}{B + 10} \end{aligned} \quad (14)$$

### 3.2 Analysis with no noise in true state data

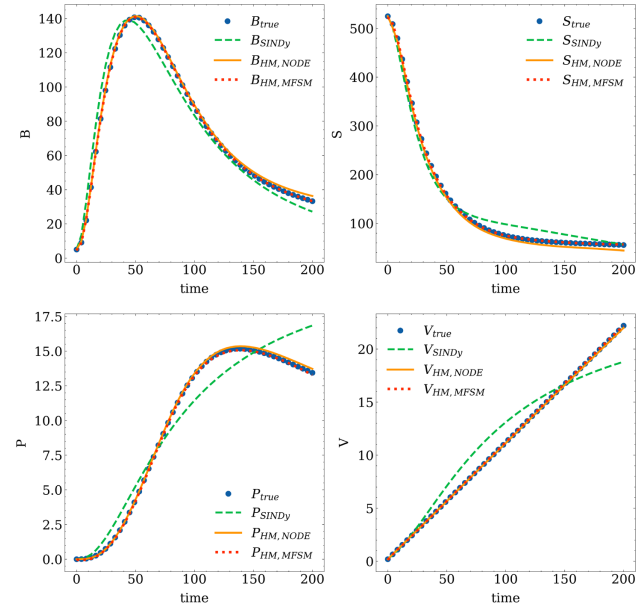
The case study shown in 3.1.1 and 3.1.2 were simulated with the initial conditions  $[C_{A,0}, T_0] = [0.5 \frac{mol}{L}, 350K]$  and  $[B_0, S_0, P_0, V_0] = [5 \frac{g}{L}, 525 \frac{g}{L}, 0 \frac{g}{L}, 0.2L]$  respectively, to generate 50 HF data points for each case. In the next step, the workflow shown in Table (1) was utilized to generate the Hybrid model with NODE and MSFM formulations to correct the error from SINDy – LF model. The results for both the approaches are shown below. Figure (2) shows the prediction using hybrid models with NODE MFSM formulation for case study 3.1.1 and Figure (3) shows the results for case study 3.1.2. In both Figures (2) and (3), the solid blue dots represent true (HF) state data. The green dashed line represents SINDy-LF model predictions. The solid orange line and the dotted red lines represent the HM model predictions with NODE and MFSM formulations respectively.

Figures (2) and (3) illustrate the limitation of the constructed SINDy model in accurately predicting the true states. This mismatch between the HF state data and the SINDy model stems from the candidate library's inability to comprehensively encompass all potential nonlinear terms contributing to the final mechanistic model equation. Consequently, the mechanistic model derived lacks accuracy in predicting the HF states. However, the HMs built using this SINDy model as the LF model,

employing both NODE and MFSM formulations, demonstrate the ability to accurately predict the HF state profiles. This HM structure effectively compensates for mismatch and missing terms within the LF model.



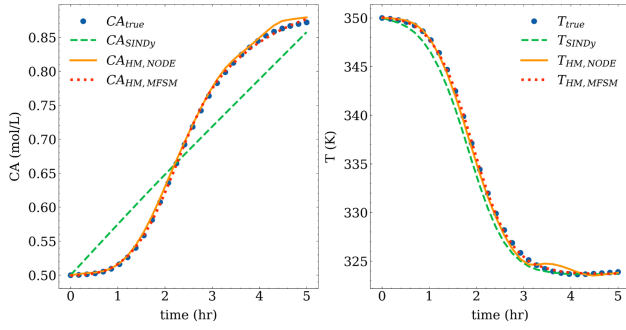
**Figure 2:** Concentration ( $C_A$ ) and temperature ( $T$ ) profiles from SINDy model, HM-NODE and HM-MFSM formulation for case study 3.1.1, compared with true state data.



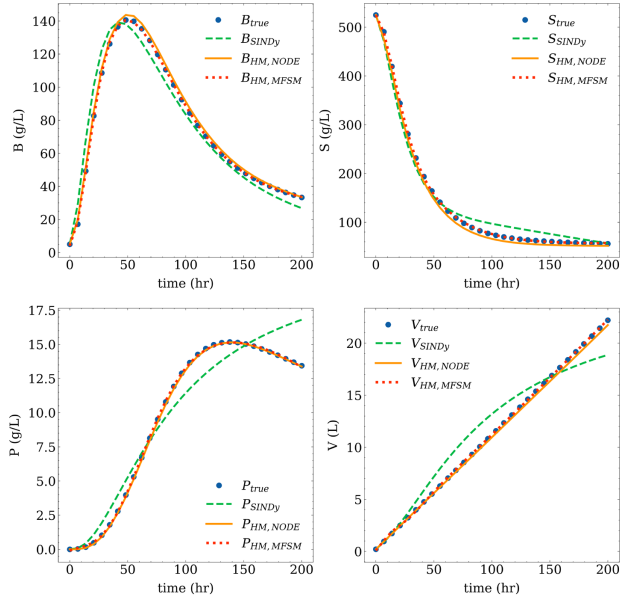
**Figure 3:** Volume ( $V$ ), Biomass ( $B$ ), product ( $P$ ) and substrate ( $S$ ) profiles from SINDy model, HM-NODE and HM-MFSM formulation for case study 3.1.2 with 50 HF data points, compared with true state data.

### 3.3 Effect of density of sampled data on HMs

To analyze the effect of density of sampled data on the HM model, we repeated the experiment from 3.2, by decreasing the amount of HF data available. For this analysis, we reduced the data size to 30 HF samples. The case study shown in 3.1.1 and 3.1.2 were simulated with the initial conditions in 3.2 to generate 30 HF data points. In the next step, using the workflow from Table (1) HMs are built. Figures (4) and (5) show the prediction from the SINDy and HMs.

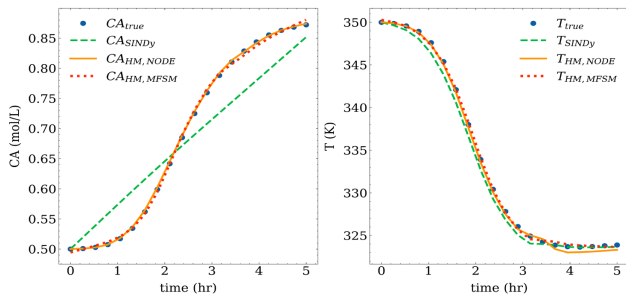


**Figure 4:** Concentration ( $C_A$ ) and temperature ( $T$ ) profiles from SINDy model, HM-NODE and HM-MFSM formulations for case study 3.1.1, with 30 HF datapoints.



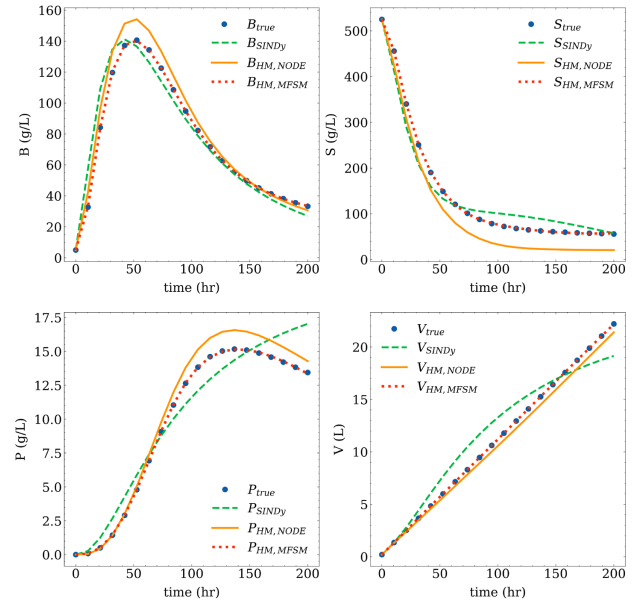
**Figure 5:** Volume ( $V$ ), Biomass ( $B$ ), product ( $P$ ), substrate ( $S$ ) profiles from SINDy model, HM-NODE and HM-MFSM formulations for case study 3.1.2, with 30 datapoints

We repeat the analysis once again, but this time by reducing the data size to 20 HF samples. We simulate the case studies from 3.1.1 and 3.1.2 with the same initial conditions to generate 20 HF data points to train the SINDy and HM models. Figures (6) and (7) show the prediction from the SINDy model and HMs.



**Figure 6:** Concentration ( $C_A$ ) and temperature ( $T$ ) profiles from SINDy model, HM-NODE and HM-MFSM formulations for case study 3.1.1, with 20 HF datapoints

formulations for case study 3.1.1, with 20 HF datapoints.



**Figure 7:** Volume ( $V$ ), Biomass ( $B$ ), product ( $P$ ), substrate ( $S$ ) profiles from SINDy model, HM-NODE and HM-MFSM formulations for case study 3.1.2, with 20 HF datapoints

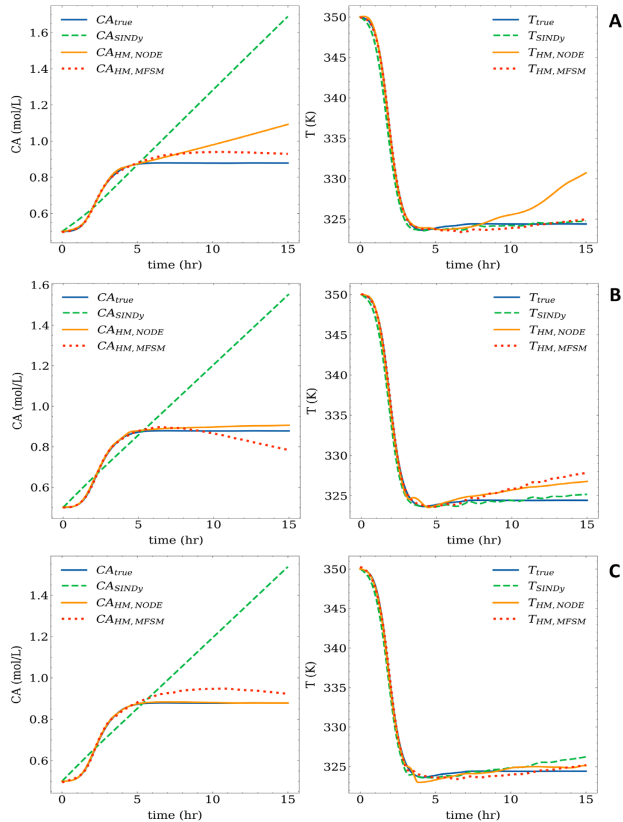
Figures (4-7) show that the HMs constructed with both NODE and MFSM approaches are robust to the low densities of sampled data and can still predict the true states with decent accuracy and capture the profile trends better than the SINDy model. We can also notice that the decrease in density of the data affects the model accuracy, and we start to observe a slight mismatch in HM predictions, and this increases as the density of sampled data decreases. Among the two HM approaches, the HM-MFSM approach performs marginally better than the HM-NODE approach. This is because we take a two-step approach for correcting the error, and decreasing the density of data affects the LF model in step 1. Consequently, the LF model directly affects the NODE approach which is aimed at correcting the derivative space error and results in low errors as we integrate it forward in time for the state profiles. On the other hand, we see an improved fit with MFSMs in the training range. This is because the MFSM approach corrects the error in the state space and does not need integrating the states forward in time.

### 3.3 Extrapolation of the HMs

To analyze the extrapolation ability of the constructed HMs, we show the model predictions with a test dataset that contains input to the HM model from outside the training dataset region. We also compare the results with the SINDy model we obtained in the first step and with decreasing data density. Figure (8) shows three scenarios aimed at illustrating extrapolation using HMs for the case study 3.1.1. Specifically, Figures (8A, 8B, 8C)

correspond to instances with 50, 30, 20 HF samples, respectively. It is evident that HMs utilizing both NODE and MFSM formulations demonstrate better extrapolation capabilities. On the other hand, predictions obtained from the SINDy model show limited extrapolative power.

We can also notice that as data density decreases, the extrapolation performance of these HM with NODE formulation is slightly better than the MFSM formulation. This observation can be attributed to the manner in which error correction is implemented in both approaches. As the NODE approach corrects error in the derivatives, it has an edge in capturing profiles compared to MFSMs when extrapolating the corresponding HM.



**Figure 8:** Extrapolated ( $C_A$ ) and ( $T$ ) profiles from SINDy model, HM-NODE and HM-MFSM formulations for case study 3.1.1 with 50, 30, and 20 HF samples.

Figure (9) shows the same analysis for the case study 3.1.2. Figures (9A, 9B, 9C) correspond to instances with 50, 30, 20 HF samples, respectively. We observe a similar behavior for this case study as well. HMs with NODE and MFSM formulations demonstrate better extrapolation capabilities than the SINDy. We can also observe a similar trend that shows HM with NODE formulation is slightly better than the MFSM formulation.

### 3.4 Effect of noise on the HMs prediction

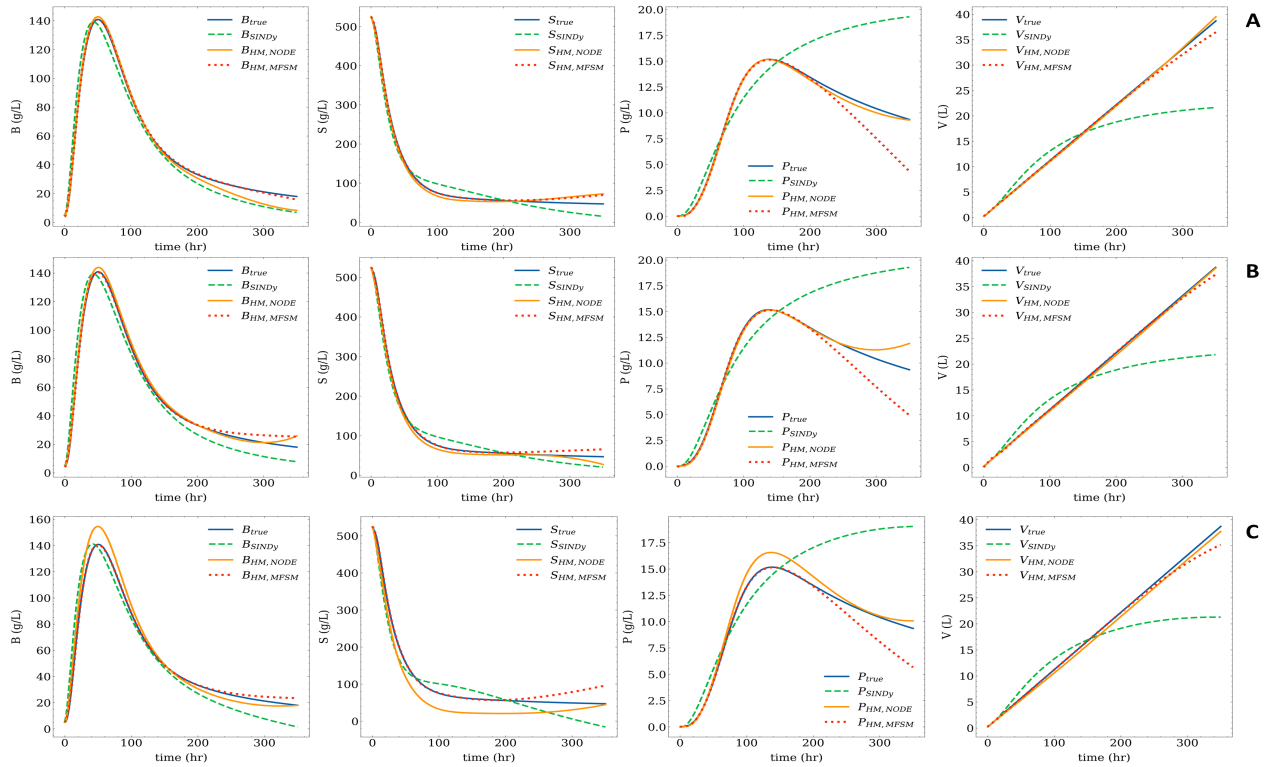
In most practical applications, there is noise associated with the true state measurements. It is necessary to

evaluate the model performance with noise in the true state data and check for robustness. To test our approach, we replicate a practical scenario by adding noise to the simulated data. The true state dataset was modified by adding Gaussian-distributed noise to each state variable data for the two case studies 3.1.1 and 3.1.2. The noise was simulated using a normal distribution, where the mean was set to zero and the standard deviation was adjusted to represent 3%, 5% of the range of the uncorrupted data. To test the methods at challenging scenarios, we present results with 20 HF samples and varying levels of noise.

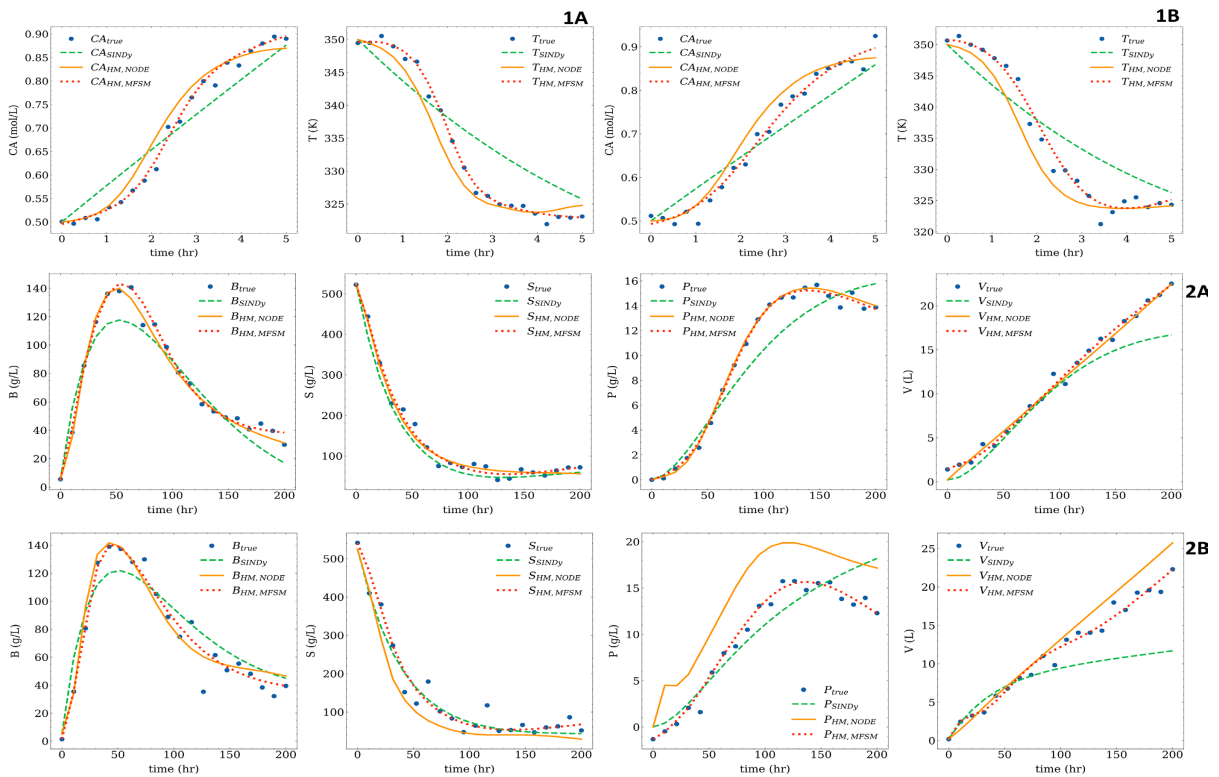
Figure (10) shows the results for this analysis. The legend (A) in Figure (10) represents the 3% noise and (B) represents 5% noise in data. We can notice the robustness of the HM approach towards the noise in data and both HM approaches are able to still predict the profiles accurately in both cases studies 3.11 and 3.1.2. But we can also observe that increasing the noise for the same sparsity reduces the accuracy of predicted profiles. This is expected because modeling accuracy becomes a challenge with increasing sparsity and noise. When comparing the two methods, the HM-MFSM approach performs better because the HM-MFSM structure directly fits the state values for each variable and proper tuning of the model helps in making it more robust towards noise. On the other hand, the HM-NODE fits the derivative values and estimating the derivatives in the presence of noise and sparsity becomes a challenge. While there exist many noise filtering techniques to smoothen the noisy data and mitigate this issue, we intend to include those as the future work.

## 4. CONCLUSIONS

In this work, we address the complexities inherent in data-driven model identification, especially when the underlying physics of the model remains unknown. We investigate the limitations of current state-of-the-art tools like SINDy when the available candidate library inadequately covers all potential terms contributing to the final mechanistic model equation. To tackle this challenge, we propose combining hybrid modeling techniques with sparse regression with the goal of simultaneous hybrid model development and model identification. To achieve this, we outline a workflow that employs two distinct methods: a) utilizing a NODE formulation, and b) employing an MFSM formulation as hybrid modeling approaches integrated with SINDy. Our study demonstrates the effectiveness of this hybrid model architecture in constructing accurate models capable of predicting profiles that predict the true data accurately, while integrating the mechanistic model knowledge.



**Figure 9:** Extrapolated Volume ( $V$ ), Biomass ( $B$ ), product ( $P$ ), substrate ( $S$ ) profiles from SINDy model, HM-NODE and HM-MFSM formulations for case study 3.1.2 with 50, 30, and 20 HF samples.



**Figure 10:** Predicted profiles from SINDy model, HM-NODE and HM-MFSM formulations for case study 3.1.1 (1A, 1B), and 3.1.2 (2A, 2B) with noise in the data. A, B correspond to 3% noise and 5% noise respectively.

We also showcase the robustness of hybrid models in handling low densities in sampled data and their ability to extrapolate. At lower amounts of sampled data, HM-NODE formulation can extrapolate better and HM-MFSM formulation can predict the states more accurately. Furthermore, we investigated the effect of noise in the true state data on these hybrid models. While the HM formulations could still predict the true profiles with a good accuracy, we saw that reducing the sampling data and increasing the noise can affect the model performance directly and make them less accurate. Future directions on this work will be focused on testing our methods with experimental data to formulate the models and proposing more efficient techniques to train hybrid models and improving their robustness towards data sparsity and noise.

## 5. ACKNOWLEDGEMENTS

The authors acknowledge support from the National Science Foundation (NSF-1944678)

## REFERENCES

- Bradley, W., et al., Perspectives on the integration between first-principles and data-driven modeling. *Computers & Chemical Engineering*, 2022. **166**: p. 107898.
- van de Berg, D., et al., Data-driven optimization for process systems engineering applications. *Chemical Engineering Science*, 2022. **248**: p. 117135-117135.
- Quade, M., et al., Prediction of dynamical systems by symbolic regression. *Physical Review E*, 2016. **94**(1): p. 12214-12214.
- Koza, J., On the programming of computers by means of natural selection. *Genetic programming*, 1992.
- Brunton, S., J. Proctor, and N. Kutz. Sparse identification of nonlinear dynamics (sindy).
- Brunton, S.L., J.L. Proctor, and J.N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 2016. **113**(15): p. 3932-3937.
- Champion, K., et al., Data-driven discovery of coordinates and governing equations. *Proceedings of the National Academy of Sciences*, 2019. **116**(45): p. 22445-22451.
- Chen, Z., Y. Liu, and H. Sun, Physics-informed learning of governing equations from scarce data. *Nature Communications*, 2021. **12**(1): p. 6136-6136.
- Sun, F., Y. Liu, and H. Sun, Physics-informed spline learning for nonlinear dynamics discovery. *arXiv preprint arXiv:2105.02368*, 2021.
- Lejarza, F. and M. Baldea, Data-driven discovery of the governing equations of dynamical systems via moving horizon optimization. *Scientific Reports*, 2022. **12**(1): p. 11836.
- Von Stosch, M., et al., Hybrid semi-parametric modeling in process systems engineering: Past, present and future. *Computers & Chemical Engineering*, 2014. **60**: p. 86-101.
- Bradley, W. and F. Boukouvala, Two-Stage Approach to Parameter Estimation of Differential Equations Using Neural ODEs. *Industrial & Engineering Chemistry Research*, 2021. **60**(45): p. 16330-16344.
- Roesch, E., C. Rackauckas, and M.P.H. Stumpf, Collocation based training of neural ordinary differential equations. *Statistical Applications in Genetics and Molecular Biology*, 2021. **20**(2): p. 37-49.
- Chen, R.T.Q., et al., Neural ordinary differential equations. *Advances in neural information processing systems*, 2018. **31**.
- Lee, D., A. Jayaraman, and J.S. Kwon, Development of a hybrid model for a partially known intracellular signaling pathway through correction term estimation and neural network modeling. *PLoS Computational Biology*, 2020. **16**(12): p. e1008472.
- Lai, Z., et al., Structural identification with physics-informed neural ordinary differential equations. *Journal of Sound and Vibration*, 2021. **508**: p. 116196.
- Ravutla, S., J. Zhai, and F. Boukouvala, Hybrid Modeling and Multi-Fidelity Approaches for Data-Driven Branch-and-Bound Optimization, in *Computer Aided Chemical Engineering*. 2023, Elsevier. p. 1313-1318.
- Meng, X. and G.E. Karniadakis, A composite neural network that learns from multi-fidelity data: Application to function approximation and inverse PDE problems. *Journal of Computational Physics*, 2020. **401**: p. 109020-109020.
- Lagaris, I.E., A. Likas, and D.I. Fotiadis, Artificial neural networks for solving ordinary and partial differential equations. *IEEE transactions on neural networks*, 1998. **9**(5): p. 987-1000.
- Guo, M., et al., Multi-fidelity regression using artificial neural networks: efficient approximation of parameter-dependent output quantities. *Computer methods in applied mechanics and engineering*, 2022. **389**: p. 114378-114378.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Technoeconomic and Sustainability Analysis of Batch and Continuous Crystallization for Pharmaceutical Manufacturing

Jungsoo Rhim<sup>a,b</sup> and Zoltan Nagy<sup>b\*</sup>

<sup>a</sup> Purdue University, Department of Aeronautics and Astronautics Engineering, West Lafayette, IN, USA

<sup>b</sup> Purdue University, Department of Chemical Engineering, West Lafayette, IN, USA

\* Corresponding Author: znagy@purdue.edu

---

## ABSTRACT

Continuous manufacturing in pharmaceutical industries has shown great promise to achieve process intensification. To better understand and justify such changes to the current status quo, a technoeconomic analysis of a continuous production must be conducted to serve as a predictive decision-making tool for manufacturers. This paper uses PharmaPy, a custom-made Python-based library developed for pharmaceutical flowsheet analysis, to simulate an annual production cycle for a given active pharmaceutical ingredient (API) of varying production volumes for a batch crystallization system and a continuous mixed suspension, mixed product removal (MSMPR) crystallizer. After each system is optimized, the generalized cost drivers, categorized as capital expenses (CAPEX) or operational expenses (OPEX), are compared. Then, a technoeconomic and sustainability cost analysis is done with the process mass intensity (PMI) as a green metric. The results indicate that while the batch system does have an overall lower cost and better PMI metric at smaller manufacturing scales in comparison with the continuous system, the latter system showed more potential for scaling-up for larger production volumes.

---

**Keywords:** Technoeconomic Analysis, Industry 4.0, Process Design, Modelling and Simulations, Optimization

## INTRODUCTION

As technology develops and industries advance into the “Industry 4.0” era, the sector of chemical and pharmaceutical manufacturing is no exception. As such, the pharmaceutical industry has been working tirelessly to discover and apply innovations to the field [1]. In particular, the concept of Quality-by-design (QbD), which was first adopted by the FDA as a means to ensure quality in the development, manufacturing, and regulation of drugs [2], has been augmented by the Quality-by-Control (QbC) framework, which employs real-time process observation and control [3]. The paradigm of QbC has notably been applied to concept of continuous crystallizers [3,4]. Continuous manufacturing has been accepted as a promising technology to achieve process intensification in pharmaceutical manufacturing [4,5]. Such methods are important as it promises flexibility and efficiency for both high volume products as well as personalized medicine

[4-6]. However, before the entire industry can adopt a new method of production, technoeconomic cost analyses of continuous production are necessary as a predictive decision-making tool for manufacturers. This is because an intimate understanding of the cost drivers and performance of continuous crystallizers is necessary for manufacturers to adopt change in an already batch system dominated industry [7]. But beyond the importance of technoeconomic analyses in industry, the issue of sustainable processes has also become more pressing. The importance of coupling technoeconomic models with sustainable metrics for a technoeconomic sustainability analysis, not just chemical processes, but for process design in general has been noted [8]. Thus, the application of sustainability metrics, either in forms of life cycle assessment or simple quantitative standards is important for chemical manufacturing [9].

In this paper, a preliminary investigation on the comparison of conventional batch crystallizers and mixed

suspension, mixed product removal (MSMPR) continuous crystallizers is conducted. First, simulated models of the annual performance of both types of systems are constructed using PharmaPy, a custom-made library for pharmaceutical flowsheet analysis [10]. For both layouts, the common active pharmaceutical ingredient (API) of paracetamol (PCM), a common analgesic drug, has been selected. Then, given three different fixed annual production volumes, operational parameters, and desired critical quality attributes (CQAs), each system is optimized for both overall cost as well as sustainability, using the process mass intensity (PMI) as a quantitative metric. For each case, a derivative-free optimizer was used.

## METHODOLOGY

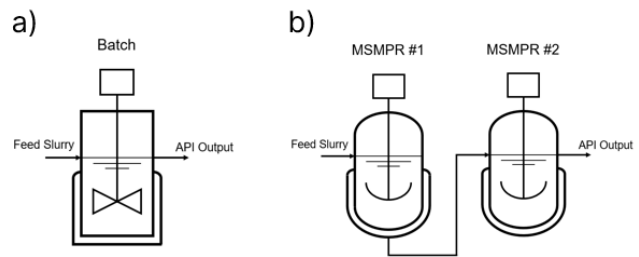
### Modeled Flowsheets

For this study, the analyses were conducted on the crystallizer unit operation. For both batch and continuous crystallizer layouts, the selected API for simulation was paracetamol. The kinetic parameters for the API have been adapted from Szilagyi et al. [11]. Additionally, an arbitrary number of annual workdays were selected wherein the system was set to produce three different annual production volumes. For the batch cooling crystallizer setup, a single unit is set as the default. However, as part of the decision variables, up to three parallel units are considered, thus allowing for the batch crystallizer to be optimized for numbering-up as well as scaling-up. In comparison, the continuous crystallizer setup is comprised of two chained MSMPR crystallizer units. The reason for this setup is because while a single batch cooling crystallizer has the flexibility to operate with a dynamic cooling profile, each MSMPR unit can only be operated under a static temperature value. Thus, by setting up two MSMPR units, the optimizer can affect the process with two different crystallizer dimensions, residence times, and operating temperatures, which enable to manipulate both throughput and critical quality attributes (CQAs), thus allowing for sufficient control and complexity for comparison with the batch process.

For the batch layout, the API is produced in batch, wherein the number of total batches are calculated from the optimal cycle time, which factors in a static one-hour ramp-down/cleanup time. However, for the continuous crystallizer, API is produced continuously in a singular campaign duration. This optimal campaign duration is calculated by multiplying the residence time of the secondary MSMPR,  $\tau_{CR02}$ , by the optimal steady state horizon multiplier,  $H_{ss}$ . This determines the duration at which the MSMPR system operates at steady state and continuously produces API. Additionally, the optimal value for  $\tau_{CR02}$  was obtained from a preliminary optimization sequence prior. Like the batch system, the MSMPR system also has a ramp-down/cleanup time, but it is only initiated

once after the campaign. However, the MSMPR also has ramp-up time wherein it must reach steady state, which was estimated for the temperature decision variables of the MSMPR units.

Finally, for each setup, an inlet of feed slurry is defined. This feed slurry serves to represent an input from a reactor unit. However, as the focus of this work is to analyze the performance of the crystallizer unit, the reactor unit has been omitted. Then, the output of each crystallizer is run through a filtration process step. However, like the reactor, the filtration unit operation has been omitted. A schematic summary of the two setups can be seen in **Figure 1**.



**Figure 1.** Schematic summary of the two different crystallization unit operations, a) batch crystallization and b) continuous crystallization.

Both crystallization configurations were simulated using PharmaPy. As previously mentioned, PharmaPy is a custom-made Python-based library for the analysis of pharmaceutical flowsheets. Using this tool, the defined unit operations can have operational parameters, such as inlet flows, initial conditions, and pharmacokinetic parameters assigned to in an object-oriented software structure [10]. Then, once the unit operations are defined as such, a simulation object of the flowsheet is created. This coupled with a callback function, decision variables can be set as inputs and PharmaPy operations as outputs allows for an optimization framework to be established.

### Optimization Formulation

For this study, the optimization problem is expressed as a non-linear constrained design problem wherein the objective is to minimize either the total cost of the manufacturing the API or the sustainability metric, PMI. The proper definition can be seen in Equation (1) [12].

$$\min_x J(\mathbf{x}, \mathbf{y}, \mathbf{z}) \quad (1)$$

$$\text{s. t. } \frac{d\mathbf{y}}{dt} = \mathbf{f}_1(t, \mathbf{y}, \mathbf{z}, \mathbf{u}(t)),$$

$$\mathbf{f}_1(t, \mathbf{y}, \mathbf{z}, \mathbf{u}(t)) = \mathbf{0}, \quad (2)$$

$$\mathbf{y}(t = 0) = \mathbf{y}_0, \mathbf{z}(t = 0) = \mathbf{z}_0,$$

and

$$g_i(x, y, z, u) \leq 0, \quad \forall i \in I, \quad (3)$$

$$\mathbf{x}_{lb} \leq \mathbf{x} \leq \mathbf{x}_{ub}$$

The equations in Equation (2), adapted from Casas-Orozco *et al.* [12], correspond to the process model of a differential-algebraic equation (DAE) system, with  $\mathbf{y} \in \{y_1, \dots, y_j, \dots, y_{n_y}\}$  being the set of differential states and  $\mathbf{z} \in \{z_1, \dots, z_{n_z}\}$  being the set of algebraic states, and  $\mathbf{y}_0$  and  $\mathbf{z}_0$  being their respective initial values. Additionally, the model inputs are represented by the variables,  $u(t) \in \{u_1(t), \dots, u_{n_u}(t)\}$ . The entirety of the DAE system is represented by the PharmaPy simulation. Finally, Equation (3) shows the nonlinear constraints and decision variable bounds for the problem.

For this study, the decision variables are dependent on the type of crystallizer that is being simulated. Consequently, the lower and upper bounds represented in Equation (3) vary by the system. Thus, the decision variables that are considered as well as their bounds are listed in **Table 1**.

**Table 1:** Description of decision variables considered in the optimization problem along with their bounds.

Variable	System	Description	Bounds
$V_{CR}$	Batch	Cryst. Volume	0.1 ~ 7.5 [m <sup>3</sup> ]
$t_{CR}$	Batch	Cycle Time	10 ~ 720 [min]
$n_{CR}$	Batch	No. of parallel process lines	1 ~ 3 [lines]
$T_{CR,i}$	Batch	Cryst. <i>i</i> th Temp. Point	273 ~ 330 [K]
$V_{CR01}, V_{CR02}$	Cont.	Cryst. Volume	0.1 ~ 7.5 [m <sup>3</sup> ]
$T_{CR01}, T_{CR02}$	Cont.	Cryst. Temp.	273 ~ 330 [K]
$H_{SS}$	Cont.	Steady state multiplier	1 ~ 100,000

Finally, the function  $g_i(x, y, z, u)$  represent the nonlinear constraints that are being applied to the problem. These constraints are representative of CQAs or standards that would be an important metric for determining the success of the system. The first constraint is that the API produced in a system must be at least a certain diameter. This an important CQA for a crystallization unit as the mean size of the crystals, and to an extent the crystal size distribution (CSD) can determine the flowability and filterability of the API produced [13]. This then has large implications for how easily the API is handled, or even how effective the drug is. The second constraint is the production volume. This is simply to ensure that the optimal results always at least produce enough API product to meet the fixed annual production volume. The third constraint is the yield constraint, which is in place to make sure that the optimal results would produce a certain percentage of the theoretical maximum yield, thus ensuring a certain level of efficiency. The fourth constraint is an operational constraint to make sure that the

temperature profile for the batch and the temperatures in the consecutive MSMPRs are monotonically decreasing. Finally, the fifth and final constraint are implemented to make sure that the total calculated time for manufacturing does not exceed the allotted annual workdays. This is in place because while the total time for the batch system is divided into distinct batches, the continuous system simply has a single campaign to continuously create API. These constraints are summarized in **Table 2**. In addition, **Table 2** also lists the weights for each constraint. These weights were then applied to a penalty function for the constraints to ensure that the optimal solution wasn't trivial or impractical.

**Table 2:** Description of the constraints considered in the optimization problem and their respective weights.

Variable	Description	Weight	Constraint
$g_1$	Mean Crystal Size	$w_1 = 10^2$	$40 [\mu\text{m}] < \bar{L}$
$g_2$	Production Volume	$w_2 = 10^3$	$PV_{target} < PV_{actual}$
$g_3$	Overall Yield	$w_3 = 10^2$	$0.9Y_{max} < Y_{actual}$
$g_4$	Decreasing Temp.	$w_4 = 10^0$	$T_{i+1} \leq T_i$
$g_5$	Total Time	$w_5 = 10^1$	$260 [\text{days}] > t_{total}$

For this simulation, the defined optimization problem was then solved with the adaptive Nelder-Mead algorithm included in the SciPy library. Thus, to translate the problem to an unconstrained optimization problem for the derivative-free Nelder-Mead algorithm, the nonlinear constraints were reformulated into an augmented objective function. It should be noted that given the non-convexity of the problem and due to the challenges of employing a gradient-based method in a simulation-optimization approach to the problem, a derivative-free algorithm was preferred. Furthermore, to improve optimizer performance, the objective function and constraint values were normalized. Thus, the objective function was transformed to Equation (4) and the constraints were transformed to Equations (5).

$$J_{norm}(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \left[ \frac{J(\mathbf{x}, \mathbf{y}, \mathbf{z}) - J_{min}(\mathbf{x}, \mathbf{y}, \mathbf{z})}{J_{min}(\mathbf{x}, \mathbf{y}, \mathbf{z})} \right]^2 \quad (4)$$

$$g_{norm,i} = \begin{cases} i \in [1,4], & 1 - \frac{g_{i,target}}{g_i} \\ i = 5, & \frac{g_{i,target}}{g_i} - 1 \end{cases} \quad (5)$$

where  $J_{min}(\mathbf{x}, \mathbf{y}, \mathbf{z})$  is the overall lowest function evaluation the optimizer found and  $g_{i,target}$  is the constraint value for  $g_i$ . Thus, the final augmented and normalized objective function is shown in Equation (6).

$$\min_{\mathbf{x}} J_{norm}(\mathbf{x}, \mathbf{y}, \mathbf{z}) + \sum_{i \in [1,5]} [\max(0, w_i \cdot g_{norm,i})]^2 \quad (6)$$

## Cost Calculation

As previously mentioned, one of the objective functions used to evaluate the simulation is the overall cost of the system. The cost of the system can be broken down into two categories, capital expense (CAPEX) and operational expense (OPEX).

### CAPEX Calculation

CAPEX involves all the terms that are related to the purchase of equipment. However, while the volume of the crystallizers as a decision variable are not discrete values, in reality, equipment are usually made and sold at discrete capacities. Thus, to account for this, a cost-capacity correlation equation is from Diab *et al.* [6] is used:

$$C_B = f C_A \left( \frac{S_B}{S_A} \right)^n \quad (7)$$

Wherein, the  $C_i$  is the cost of the equipment and  $S_i$  is the capacity of the given equipment. Next,  $f$  are equipment-dependent coefficients to account for indirect costs that may be involved for certain equipment. Finally,  $n$  is a value is the cost exponent to represent the exponential increase in cost of equipment as capacity increases. The index of  $A$  in Equation (7) represent existing equipment while the index of  $B$  refer to the equipment selected for the simulation. The specific values for base equipment values in Equation (7) are based on the Chemical Engineering Plant Cost Indices (CEPCIs) [14]. However, for this study, the values are identical as the ones used by Diab *et al.* [6].

Furthermore, to provide a more realistic estimation for the technoeconomic cost model, rather than taking the flat equipment cost, a battery limit installed cost (BLIC) is calculated. Thus, the additional indirect costs associated with installing the equipment are considered. To calculate the BLIC, the Chilton method is employed [6]. In summary, the BLIC is a factor of the total physical plant cost (TPPC), which is the sum of the installed equipment cost (IEC) and the process piping and instrumentation (PPI) cost. The PPI is a percentage of the IEC while the IEC is a factor of the previously calculated equipment cost. The exact coefficients and factors for these calculations are the same as the ones used by Diab *et al.* [6].

Finally, once the BLIC was calculated, rather than apply the flat equipment cost, an equivalent uniform annual cost (EUAC) was calculated. This is to reflect the fact that while it is not unheard of for a company to outright purchase all the required equipment for a new setup, it is more customary for the annual equipment cost of a production line to be expressed as an annuity [15]. The calculation can be seen in Equation (8). An interest rate,  $i_{rate}$  of 5% and a project timeline,  $t_{PL}$  of 20 years is taken from literature [6].

$$EUAC = BLIC \left( \frac{i_{rate}(1+i_{rate})^{t_{PL}}}{(1+i_{rate})^{t_{PL}} - 1} \right) \quad (8)$$

In addition to the equipment cost, the CAPEX for this

simulation also takes into consideration for the working capital (WC) and the contingency costs (CC) of the plant. The WC and CC are set to be 3.5% of the annual material costs and 20% of the total BLIC, respectively. These values can be found in literature [16].

Summary of the total CAPEX calculation is summarized in Equation (9).

$$CAPEX_{total} = EUAC + WC + CC \quad (9)$$

### OPEX Calculation

As opposed to CAPEX, which represented expenses that are investments that need to be made prior to establishing a setup, OPEX represents all expenses that are incurred during the hours of plant operation. The major cost drivers here are the cost of the materials that serve as inputs for the pharmaceutical process and the cost involved with dealing with the waste material of a process. The material cost,  $C_{material}$ , can be straightforwardly calculated as the product of the total mass of each chemical species  $m_i$  and their respective cost per mass,  $C_i$ . Finally, the waste cost,  $C_{waste}$ , can be calculated as a certain percentage of the total cost of solvents,  $C_{solv}$ , involved in the process. The exact percentage, while it may vary on a plant-by-plant basis, was set as 35%. The summary of these calculations as well as the final total OPEX calculation can be seen in Equations (10-12).

$$C_{material} = \sum_i C_i m_i \quad (10)$$

$$C_{waste} = 0.35 \sum_i C_{solv,i} \quad (11)$$

$$OPEX_{total} = C_{material} + C_{waste} \quad (12)$$

Thus, the objective function when calculating for minimal costs can be seen in Equation (13).

$$\min_x J_{cost}(x, y, z) = CAPEX_{total} + OPEX_{total} \quad (13)$$

### Sustainability Calculation

In the previous section, the method of calculating the cost of a crystallization unit operation has been outlined. However, while cost is an important metric for a technoeconomic model, sustainability, or "green processing", is an ever-growing concern for the future of the sustainable pharmaceutical industry [17]. Thus, optimizing the manufacturing system in regard to a sustainability metric in addition to the cost is a necessary perspective to take. Over time, many different metrics for sustainability have emerged. Notably, the E factor, seen in Equation (14), has often been used in studies as a metric for efficiency of pharmaceutical manufacturing [17].

$$E \text{ factor} = \frac{\text{total mass of waste from a process [kg]}}{\text{total mass of product [kg]}} \quad (14)$$

However, as the E factor serves as a ratio of the API produced and the waste material produced from the entire process, the E factor can be a misleading metric as it only represents the waste material, thus not guaranteeing the efficiency or lack thereof in regard to the other materials involved in the process [18]. Thus, as a correction, the process mass intensity (PMI) metric has been introduced. The PMI, shown in Equation (15), serves a ratio of the total API produced with the mass of all chemical species that was involved in the process [18].

$$PMI = \frac{\text{total mass from a process [kg]}}{\text{total mass of product [kg]}} \quad (15)$$

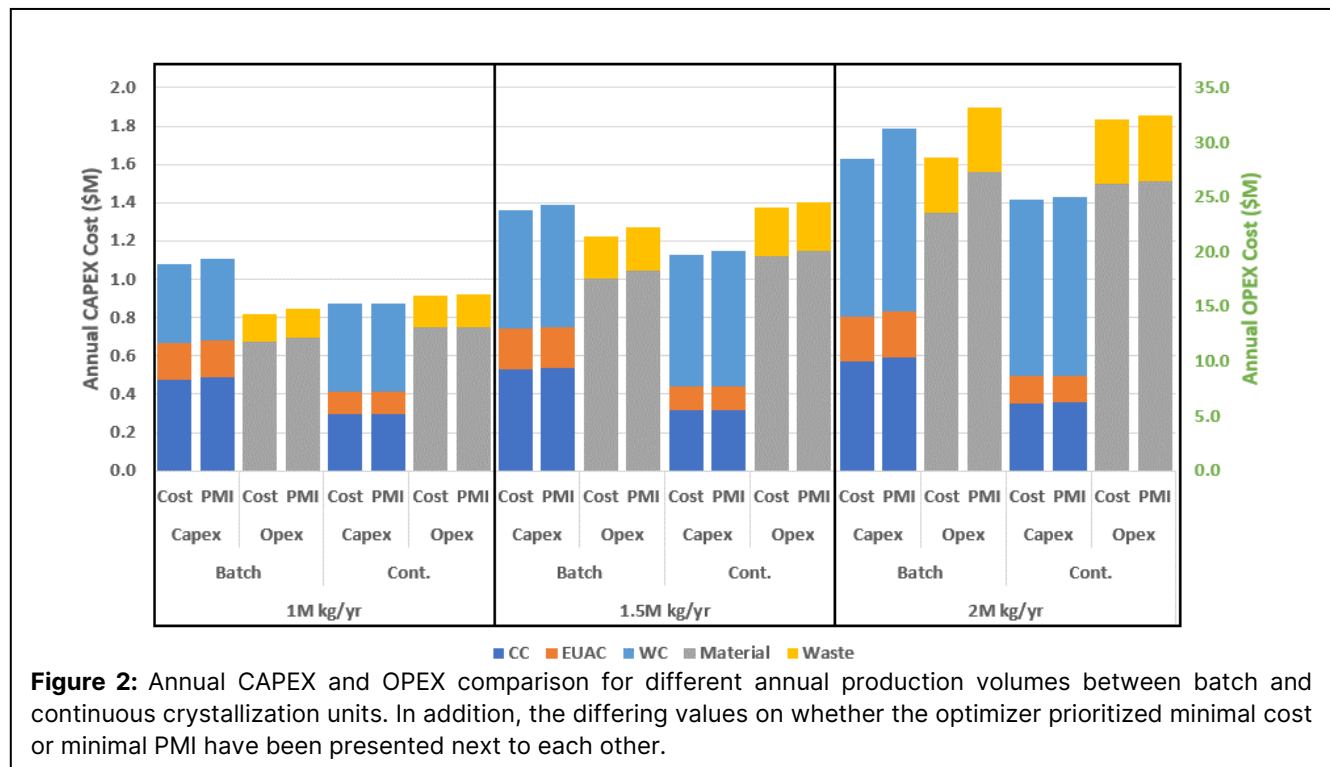
Thus, with PMI as a secondary objective function, the manufacturing system can also be investigated regarding sustainability as well as minimized costs.

## RESULTS AND DISCUSSION

With the optimization problem defined in the previous section as well as the layout of the batch and continuous crystallization manufacturing units, we can then compare the differences in performance and cost for the systems. For both batch and continuous systems, the simulation was optimized for minimal total cost, which was the sum of the CAPEX and OPEX, and for minimal PMI. The results of the optimization for both systems with the two different objective functions and at three different target annual production volumes can be seen in **Figure 2**. Additionally, it is important to note that the CAPEX and

OPEX have been graphed separately on different axes. This is because while the techno-economic cost model was created to provide a holistic view of the unit operations, the inclusion of CAPEX is not always relevant for some industries, for example where the equipment is already in place. Also, the numerical values of the simulations, optimal decision variables as well as some additional performance metrics for the batch system and the continuous system can be seen in **Table 3** and **Table 4**, respectively.

When first observing the results of the simulation, it is important to note that from **Figure 2**, we can observe that the overall OPEX for the continuous system, regardless of annual production volume, is higher than that for the batch system. This is a logical outcome as one of the drawbacks of MSMR crystallizers is that a constant feed of slurry needs to be input. However, it should also be noted that the CAPEX values for continuous systems are always lower than their batch counterparts across the board. This is even with the consideration that a single production line of continuous crystallization requires two MSMR units in cascade. This is reflective of the result that, due to the higher throughput of the continuous systems, a smaller crystallizer unit is sufficient to meet the annual production targets. However, this result comes with the caveat that the inclusion of CAPEX may not be significant for manufacturers who are not looking to create a new manufacturing line from scratch, thus making the difference in CAPEX irrelevant.





**Table 4:** Numerical results of the continuous crystallization setup simulation.

	2M kg / yr		1.5M kg / yr		1M kg / yr	
	Cost Obj.	PMI Obj.	Cost Obj.	PMI Obj.	Cost Obj.	PMI Obj.
<b>Material</b>	\$ 26,259,268.76	\$ 26,505,728.06	\$ 19,627,749.43	\$ 20,080,882.20	\$ 13,092,875.99	\$ 13,144,073.54
<b>Waste</b>	\$ 5,872,371.26	\$ 5,927,758.35	\$ 4,390,333.75	\$ 4,491,690.05	\$ 2,928,844.40	\$ 2,940,464.34
<b>EUAC</b>	\$ 142,092.38	\$ 142,770.88	\$ 126,473.28	\$ 126,473.28	\$ 119,259.38	\$ 118,712.96
<b>WC</b>	\$ 919,074.41	\$ 927,700.48	\$ 686,971.23	\$ 702,830.88	\$ 458,250.66	\$ 460,042.57
<b>CC</b>	\$ 354,157.04	\$ 355,848.15	\$ 315,227.32	\$ 315,227.32	\$ 297,247.11	\$ 295,885.18
<b>Total</b>	\$ 33,546,963.84	\$ 33,859,805.92	\$ 25,146,755.02	\$ 25,717,103.73	\$ 16,896,477.54	\$ 16,959,178.60
<b>API Made</b>	2000003.58 kg	2020834.54 kg	1500045.79 kg	1534693.85 kg	1001400.91 kg	1006269.14 kg
<b>Solvent Used</b>	8389101.80 kg	8468226.21 kg	6271905.36 kg	6416700.08 kg	4184063.43 kg	4200663.35 kg
<b>API Used</b>	3160355.06 kg	3189758.55 kg	2361312.90 kg	2415827.35 kg	1574916.38 kg	1580915.61 kg
<b>Total Time</b>	210.29 days	258.41 days	226.41 days	231.63 days	163.70 days	174.56 days
<b>Throughput</b>	396.285 kg/h	325.840 kg/h	276.058 kg/h	276.062 kg/h	254.888 kg/h	240.195 kg/h
<b>Availability</b>	99.93%	99.94%	99.93%	99.93%	99.91%	99.91%
<b>Cost/Kg</b>	\$16.77 /kg	\$16.76 /kg	\$16.76 /kg	\$16.76 /kg	\$16.87 /kg	\$16.85 /kg
<b>PMI</b>	5.774718E+00	5.768896E+00	5.755303E+00	5.755237E+00	5.750923E+00	5.745559E+00
<b>Cost Change</b>		0.933%		2.268%		0.371%
<b>PMI Change</b>		-0.101%		-0.001%		-0.093%
<b>Optimal Decision Variables</b>						
$V_{CR01}$	2.020 m <sup>3</sup>	1.653 m <sup>3</sup>	1.396 m <sup>3</sup>	1.531 m <sup>3</sup>	1.288 m <sup>3</sup>	1.213 m <sup>3</sup>
$V_{CR02}$	4.123 m <sup>3</sup>	3.833 m <sup>3</sup>	3.068 m <sup>3</sup>	3.332 m <sup>3</sup>	2.752 m <sup>3</sup>	2.821 m <sup>3</sup>
$T_{CR01}$	273.00 K	273.00 K	273.00 K	273.00 K	273.00 K	273.00 K
$T_{CR02}$	273.00 K	273.00 K	273.00 K	273.00 K	273.00 K	273.00 K
$H_{ss}$	9078.73	11156.65	9774.24	8940.25	7065.28	7534.09

In addition to this observation, we can see from **Table 3** and **Table 4** that despite the differences in CAPEX, the overall cost for the batch system is lower than their continuous counterparts. However, that comparison does not provide a full view of the comparison. An important aspect of continuous crystallization is the efficiency in terms of throughput and availability. This is observed in the simulation results in **Table 4**. The results show that for all annual production volumes, the continuous system resulted in a more consistent throughput rate while **Table 3** indicates that the batch system has a steeper drop in throughput. This is also reflected in the fact that for all three annual production volumes, the batch system needed to use all 260 days for all cases, even though the batch cases have already optimized for 2 parallel lines.

Additionally, we can see that for the continuous system, the overall optimized crystallizer volume is lower than for the batch scenario. In conjunction with this, we can see that the total availability, which is the percentage of time in which the system is actually running and not ramping up, ramping down or cleaning, the continuous system predictably has a higher percentage. With all these observations, we can see that the MSMR setup, while for the selected parameters may have an overall higher cost, shows a better potential for scaling up and provides a more agile manufacturing alternative.

Finally, the previously made observations can also be seen when considering the PMI. From the same tables, we when we see the results for optimizing for PMI rather than cost, we can compare the sustainable nature of both setups. As expected, the continuous system has a higher PMI value due to the necessity for a continuous input of slurry. However, when comparisons with cost in mind, we can see that while optimizing for PMI decreases the resultant PMI by less than 1%, the cost in continuous systems increases less significantly than the batch systems. Thus, while the significance of decreasing PMI would be different on a case-by-case basis, we can observe that the continuous system has a lower cost necessary for improving the overall sustainability of the process.

## CONCLUSION

This study was an example to show the capabilities of the technoeconomic cost model simulation to serve as a decision-making tool for manufacturers. By employing a simulation-optimization strategy with the annual production of paracetamol as a generic representative API and applying CAPEX and OPEX calculations that have been standardized in literature, a good first simulation result was achieved. Furthermore, with the simulation, other than directly comparing cost, the potential

capabilities of continuous production methods in the pharmaceutical industry could be explored. While the overall costs of continuous systems may be higher than the existing batch production setup, the potential for continuous systems to scale up and maintain efficiency in both availability and throughput shows promise. Furthermore, in the light of sustainability for pharmaceutical processes, we could see that the trade-off in cost for improving PMI metrics would be much less than that for batch processes, thus additionally showing how continuous systems could be more easily adapted to be more sustainable and embody the idea of “green chemistry”.

## ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation (NSF) under Grant No. 2229250.

## REFERENCES

- Rantanen, J., & Khinast, J. (2015). The future of pharmaceutical manufacturing sciences. *Journal of pharmaceutical sciences*, 104(11), 3612-3638.
- Yu, L. X. (2014). Gregory A, Khan MA, Hoag, SW, Polli, J, Raju, GK, Woodcock, J. *Understanding pharmaceutical quality by design AAPS J*, 16(4), 771-783.
- Su, Q., Ganesh, S., Moreno, M., Bommireddy, Y., Gonzalez, M., Reklaitis, G. V., & Nagy, Z. K. (2019). A perspective on Quality-by-Control (QbC) in pharmaceutical continuous manufacturing. *Computers & Chemical Engineering*, 125, 216-231.
- Yang, Y., Song, L., & Nagy, Z. K. (2015). Automated direct nucleation control in continuous mixed suspension mixed product removal cooling crystallization. *Crystal Growth & Design*, 15(12), 5839-5848.
- Sarkis, M., Bernardi, A., Shah, N., & Papathanasiou, M. M. (2021). Emerging challenges and opportunities in pharmaceutical manufacturing and distribution. *Processes*, 9(3), 457.
- Diab, S., & Gerogiorgis, D. I. (2020). No more than three: techno-economic mixed integer nonlinear programming optimization of mixed suspension, mixed product removal crystallizer cascades for melitracen, an antidepressant API. *Industrial & Engineering Chemistry Research*, 59(49), 21458-21475.
- Vetter, T., Burcham, C. L., & Doherty, M. F. (2014). Regions of attainable particle sizes in continuous and batch crystallization processes. *Chemical Engineering Science*, 106, 167-180.
- Mahmud, R., Moni, S. M., High, K., & Carbajales-Dale, M. (2021). Integration of techno-economic analysis and life cycle assessment for sustainable process design—A review. *Journal of Cleaner Production*, 317, 128247.
- Nikolopoulou, Amalia, and Marianthi G. Ierapetritou. "Optimal design of sustainable chemical processes and supply chains: A review." *Computers & Chemical Engineering* 44 (2012): 94-103.
- Casas-Orozco, D., Laky, D., Wang, V., Abdi, M., Feng, X., Wood, E., Laird, C., Reklaitis, G. V., & Nagy, Z. K. (2021). PharmaPy: An object-oriented tool for the development of hybrid pharmaceutical flowsheets. *Computers & Chemical Engineering*, 153, 107408.
- Szilagyi, B., Eren, A., Quon, J. L., Papageorgiou, C. D., & Nagy, Z. K. (2020). Application of model-free and model-based quality-by-control (QbC) for the efficient design of pharmaceutical crystallization processes. *Crystal Growth & Design*, 20(6), 3979-3996.
- Casas-Orozco, D., Laky, D. J., Wang, V., Abdi, M., Feng, X., Wood, E., Reklaitis, G. V., & Nagy, Z. K. (2023). Techno-economic analysis of dynamic, end-to-end optimal pharmaceutical campaign manufacturing using PharmaPy. *AIChE Journal*, e18142.
- Gyulai, O., Kovács, A., Sovány, T., Csóka, I., & Aigner, Z. (2018). Optimization of the critical parameters of the spherical agglomeration crystallization method by the application of the quality by design approach. *Materials*, 11(4), 635.
- Turton, R., Bailie, R. C., Whiting, W. B., & Shaeiwitz, J. A. (2008). *Analysis, synthesis and design of chemical processes*. Pearson Education.
- Gurnani, C. (1983). Economic analysis of inventory systems. *The International Journal of Production Research*, 21(2), 261-277.
- Schaber, S. D., Gerogiorgis, D. I., Ramachandran, R., Evans, J. M., Barton, P. I., & Trout, B. L. (2011). Economic analysis of integrated continuous and batch pharmaceutical manufacturing: a case study. *Industrial & Engineering Chemistry Research*, 50(17), 10083-10092.
- Tucker, J. L. (2006). Green chemistry, a pharmaceutical perspective. *Organic process research & development*, 10(2), 315-319.
- Jimenez-Gonzalez, C., Ponder, C. S., Broxterman, Q. B., & Manley, J. B. (2011). Using the right green yardstick: why process mass intensity is used in the pharmaceutical industry to drive more sustainable processes. *Organic Process Research & Development*, 15(4), 912-917.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Enhancing Polymer Reaction Engineering Through the Power of Machine Learning

Habibollah Safari<sup>a</sup> and Mona Bavarian<sup>a\*</sup>

<sup>a</sup> Department of Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, NE, 68588

\* Corresponding Author: [mona.bavarian@unl.edu](mailto:mona.bavarian@unl.edu).

## ABSTRACT

Copolymers are commonplace in various industries. Nevertheless, fine-tuning their properties bears significant cost and effort. Hence, an ability to predict polymer properties a priori can significantly reduce costs and shorten the need for extensive experimentation. Given that the physical and chemical characteristics of copolymers are correlated with molecular arrangement and chain topology, understanding the reactivity ratios of monomers—which determine the copolymer composition and sequence distribution of monomers in a chain—is important in accelerating research and cutting R&D costs. In this study, the prediction accuracy of two Artificial Neural Network (ANN) approaches, namely, Multi-layer Perceptron (MLP) and Graph Attention Network (GAT), are compared. The results highlight the potency and accuracy of the intrinsically interpretable ML approaches in predicting the molecular structures of copolymers. Our data indicates that even a well-regularized MLP cannot predict the reactivity ratio of copolymers as accurately as GAT. This is attributed to the compatibility of GAT with the data structure of molecules, which are graph-representative.

**Keywords:** Reaction Engineering, Polymerization, Artificial Neural Network, Multilayer Perceptron, Graph Attention Network,

## INTRODUCTION

Copolymers are widely used and have a variety of applications such as coatings, in electronic devices, the packaging industry, or pharmaceutical manufacturing[1-3]. Copolymers are often designed with the aim of creating materials that possess the characteristics of their constituent monomers.

The properties of a copolymer are basically determined by the paired monomers' sequence distribution in the constructed copolymers. This distribution is commonly defined by the reactivity ratios, presenting the ratio of each monomer's propensity to react with itself over the inclination to react with another monomer[4]:

$$r_1 = \frac{k_{p,11}}{k_{p,12}}$$

$$r_2 = \frac{k_{p,22}}{k_{p,21}}$$

Here, the  $k_{p,xy}$  is the rate coefficient of propagation of radical  $x$  with monomer species  $y$ . For instance, poly

(ethylene-co-vinyl acetate) or EVA is a commercial polymer in various industries and is constructed from ethylene and vinyl acetate. The reactivity ratios for these monomers are  $r_1=0.88\approx 1$  and  $r_2=1.03\approx 1$ [5]. In a system where both  $r_1$  and  $r_2$  are close to one, the copolymerization tends to produce a random copolymer, and no specific sequence is expected. On the other hand, when  $r_1\ll 1$  and  $r_2\ll 1$ , an alternating copolymer can be expected Like Styrene-Maleic Anhydride copolymer or SMA in which  $r_1=0.02$  and  $r_2=0.003$ [5]. In this condition, each monomer has a preference to react with the other monomer instead of itself. As evident, varying values of reactivity ratios ( $r_1$  and  $r_2$ ) in copolymerization lead to different arrangements in the copolymer structure, subsequently influencing the final properties of the copolymer. Thus, having a reasonable prediction for the sequence of monomers in a copolymer chain facilitates the process of producing fit-for-purpose macromolecules. Traditionally, estimation of the reactivity ratio for a new polymer heavily depends on experimental work and the repeatability of the experiments, which are laborious, sluggish, and costly.

Numerous computational methods, such as Density Functional Theory (DFT), are used for predicting the reactivity ratio in copolymers. However, these methods generally incur significant computational costs, rendering them impractical for certain engineering applications[6].

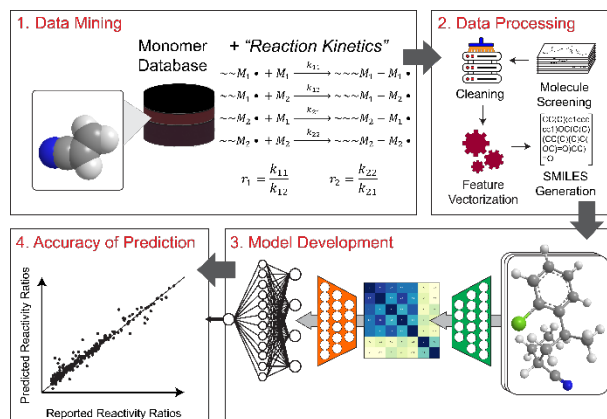
To this end, using of Machine Learning (ML) approaches can be a pragmatic solution for addressing the reaction engineering problems specifically for copolymer synthesis. From a process design perspective, ML models serve as a pivotal tool in predicting the chain topology of copolymers derived from novel monomer pairs. This application, termed the **Forward Design Problem**, utilizes predictive analytics to determine the reactivity ratios of two monomers, thereby identifying the final chain topology of the resultant copolymers. Such predictive capability enables researchers to envision the structural and functional attributes of new materials before their synthesis, streamlining the development process and enhancing the efficiency of material discovery. Conversely, the **Backward Design Problem** represents an equally vital application of ML in polymer science. By integrating ML models with optimization frameworks, it is possible to reverse-engineer the design process to identify monomer structures that yield specific reactivity ratios[7]. This is particularly crucial when aiming for copolymers with precise structural configurations, such as those with alternating monomer sequences achieved when the reactivity ratios of two monomers are significantly less than one. Through optimization, the model identifies 'monomer fingerprints' that are most likely to result in the desired topology, thus guiding the synthesis towards copolymers with predetermined properties and applications. This dual approach—predicting copolymer topologies through forward design and deducing optimal monomer structures for targeted outcomes via backward design illustrates the notable potential of ML in polymer design. This not only accelerates the material development cycle but also opens new realm for the tailored synthesis of copolymers, marking a significant advancement in the field of polymer reaction engineering. The organization of the rest of this paper is as follows. The methodology section presents the data preprocessing and implementation of the Multilayer Perceptron (MLP) and Graph Attention Network (GAT). The result and discussion describe the models' performance in predicting the monomers' reactivity ratios. Eventually, we present some concluding remarks.

## METHODOLOGY

### Data Curation and Preprocessing

In the large view, the development of an ML model constitutes 4 phases including Data Mining, Data Cleaning, Model Construction, and Performance Assessment. Figure 1 presents a general overview of the

development of an ML approach for reactivity ratio prediction in copolymers. In the development of machine learning models, our investigation is positioned within the realm of supervised learning, where each data point is labelled. Our dataset comprises pairs of monomer names, with the associated label being their reactivity ratio during the copolymerization process. Given that the output variable, the reactivity ratio, is a continuous value, our problem is identified as a regression task within the supervised learning framework. A significant challenge in this context is the effective introduction of monomer pairs to the machine learning model (particularly in the development of Multi-Layer Perceptron).



**Figure 1.** Schematic Workflow for Presentation of an ML Approach for Reactivity Ratio Prediction.

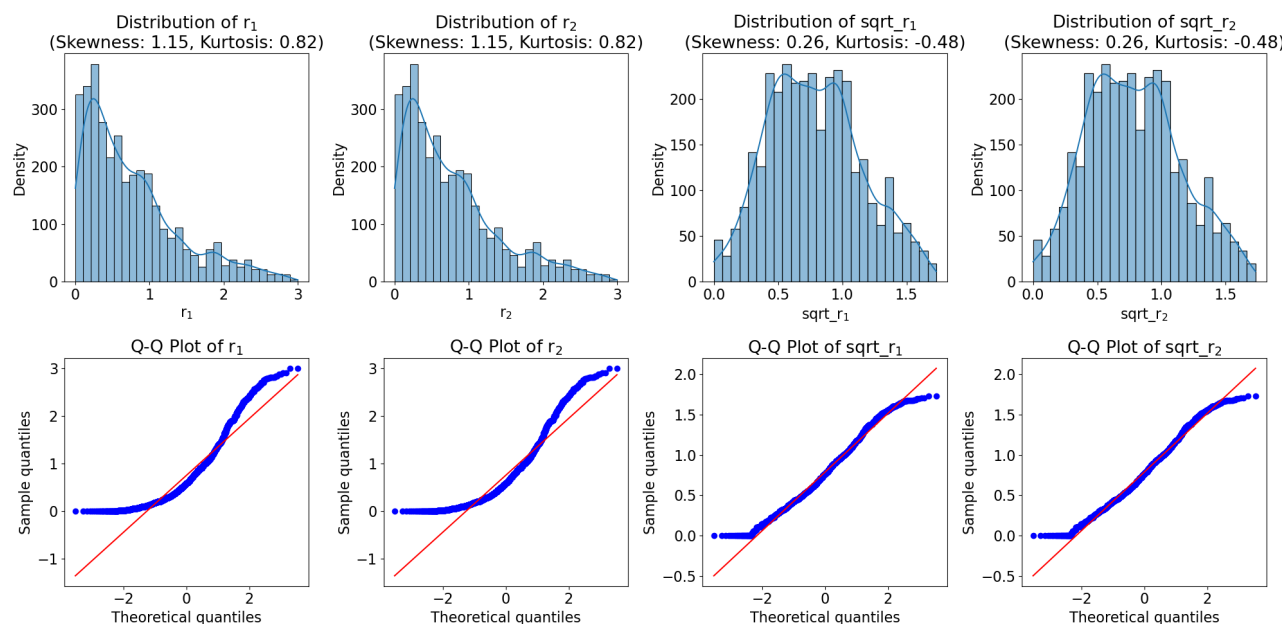
After the collection of reactivity ratios of monomers from different references, we evaluated the data in terms of value and distribution. It is generally observed that extracting patterns from a dataset with a distribution close to normal is more effective. To ameliorate the training process of the ML model, we applied a square root transformation to the raw data, aiming to achieve a distribution that is more closely aligned with a normal distribution. Figure 2. represents the original reactivity ratio ( $r_1$  and  $r_2$ ) distribution and also schemes the data distribution when we impose a square root transformation on our data ( $\sqrt{r_1}$  and  $\sqrt{r_2}$ ).

In the next phase, we convert the monomers into the Simplified Molecular-Input Line-Entry System (SMILES) format using open-source cheminformatics toolkits. This allows us to transform the monomers into a machine-readable language, suitable for machine learning development.

### Multi-Layer Perceptron

Here, a Multilayer Perceptron (MLP) for the prediction of the reactivity ratio in copolymers is considered. For converting the SMILES to a numerical vector that is ANN-compatible, a Morgan fingerprint with 2048 bits for each monomer was employed. In the





**Figure 2.** Distribution of original data ( $r_1$  and  $r_2$ ) and distribution of data with applied square root transformation ( $\sqrt{r_1}$  and  $\sqrt{r_2}$ ). As shown in the figure, the skewness of data reduces significantly and make data close to normal distribution. In the quantile-quantile (Q-Q) plot, the degree to which our data aligns with the reference line ( $y = ax+b$ ) provides an indication of how closely the data conform to a normal distribution. A close alignment suggests that the data are approximately normally distributed.

following, the fingerprint vectors of the two monomers were concatenated, and the concatenated vector was considered as the input of the MLP. So, the input of the model is a one-dimensional vector ( $X \in \mathbb{R}^{4096}$ ), and the outputs of the model are  $\sqrt{r_1}$  and  $\sqrt{r_2}$  ( $y \in \mathbb{R}^2$ ).

In defining the structure of the model, a proposed MLP with 4096 inputs, 80 neurons in the first hidden layer, 40 neurons in the second hidden layer, and 2 outputs is used[8]. Regarding the full connectivity of the neurons in all layers, the MLP performance was improved using the dropout technique. **Figure 3** represents a schematic overview of the multilayer perceptron used for reactivity ratio estimation. For a better training process, it was found that the implementation of a regularization technique can improve the Mean Square Error (MSE) significantly. In this regard, a Grid Search Optimization method was implemented to find  $L_1$  Regularization,  $L_2$  Regularization, and Dropout Rate. It was found that using Dropout Rate = 0.65 can significantly ameliorate the MLP performance. Using the optimum dropout rate in the MLP could significantly improve the reported MSE in the modified model (0.1 in the test dataset and 0.08 in the training dataset) compared to the reference model[8].

For the training of our optimized Multilayer Perceptron (MLP), the dataset was allocated as follows: 10% was reserved for testing, while the remaining 90% was utilized for both training and validation. Specifically, of the data allocated for model training and validation, 90%

was used for actual training purposes, and the remaining 10% served as validation data. This approach of incorporating a validation subset within the training data allows for regular assessment of the model's performance against overfitting. By doing so, we ensure that the model not only learns from the training data but also generalizes well to unseen data, thereby enhancing its reliability and applicability in real-world scenarios.

## Graph Attention Network

In the realm of reaction engineering, the structural intricacies of molecular data present unique challenges and opportunities for computational analysis. A promising approach for addressing these challenges is the adoption of graph-based machine learning approaches in which molecules are represented as graphs, atoms as nodes, and chemical bonds as edges. This approach appears to be an encouraging solution for various problems, including reactivity ratio prediction in polymer science. Graph Attention Networks (GAT) are a specific type of Graph Neural Network (GNN) that incorporate attention mechanisms to specify different weights to different nodes in a graph[9,10].

Herein, In the second part, a Graph attention Network with an Attentive Fingerprint was utilized. This approach employs a Recurrent Neural Network (RNN) and an Attention Mechanism for the extraction of the

most important features from the input in a molecule structure. The attention mechanism in the GAT operates based on the three which are mechanism-alignment, weighting, and context operation[11]:

The alignment equation is represented as follows:

$$e_{vu} = (W \cdot [h_v, h_u]) \quad (1)$$

For weighting, the softmax function is applied:

$$a_{vu} = \text{softmax}(e_{vu}) = \frac{\exp(e_{vu})}{\sum_{u \in N(v)} \exp(e_{vu})} \quad (2)$$

The context vector is then calculated using the Exponential Linear Unit (ELU) function:

$$C_v = \text{elu}(\sum_{u \in N(v)} a_{vu} \cdot W \cdot h_u) \quad (3)$$

Here, ' $v$ ' denotes a specific target node or atom, ' $N(v)$ ' its neighboring nodes, ' $h_v$ ' the states vector of node ' $v$ ', ' $h_u$ ' the states vector of a neighboring atom or node ' $u$ ', and ' $W$ ' the learnable weight matrix indicating relationships between the target node and its neighbors. The alignment scores are calculated using Equation 1 and then normalized (Equation 2). Subsequently, the context vector is formulated using the ELU function, allowing a non-zero gradient for negative inputs[11]. Features close to a score of 1 exert more influence on the output, while those closer to 0 are deemed less significant, and negative values result in feature exclusion.

The second stage involves constructing a viable model incorporating a Gate Recurrent Unit (GRU), which enhances the RNN by adding reset and update gates. This allows the GRU to selectively retain and disregard information, thereby augmenting RNN's memory capabilities. The GRU functions in two phases within the model: messaging and readout, described mathematically in Equations (4) and (5).

$$C_v^{k-1} = \sum_{u \in N(v)} M^{k-1}(h_v^{k-1}, h_u^{k-1}) \quad (4)$$

$$h_v^k = \text{GRU}^{k-1}(C_v^{k-1}, h_v^{k-1}) \quad (5)$$

In the messaging stage of the GAT model, the message function, denoted as  $M^{k-1}$  plays a crucial role. This function operates at the  $k - 1$  iteration, where it aids in assimilating the learned features of the nodes. During this phase, the representation of the nodes within the molecules is compiled. The message function aggregates details from neighboring nodes on the graph for each target node. This process is pivotal as the graph attention mechanism focuses on collating data from all adjacent nodes in the messaging phase to effectively update their state in the subsequent read-out phase. In this context, the GRU acts by integrating inputs from the previous state vector  $h_v^{k-1}$  of the target node and the attention context  $C_v^{k-1}$  from its neighboring nodes. In the readout phase, the GRU updates the current hidden state of the target node by employing information obtained from the messaging phase and the node's prior hidden state.

The representations of the target nodes, once learned, are then employed in the read-out phase to predict molecular properties. Detailed explanations of the functioning and application of these processes and equations (1-5) within the GAT framework can be found in the cited references[11]. The model further refines its accuracy by using features such as atom symbols, neighboring atoms, atom masks, bond types, and neighboring bonds to effectively differentiate each target node from its neighbors. In our study, the Multi-Input-Multi-Output Graph Attention Network (MIMO GAT) was used as an advanced version of the Graph Attention Network. This new model includes a special multimodal fusion block, making it different from the Attentive FP model. This network was employed to predict the reactivity ratios of monomers, using SMILES notations of monomers and copolymers. Leveraging multi-task learning, we first converted molecular structures into graph representations using RDKit for feature extraction. These features are then encoded and processed through individual Graph Attention Modules within MIMO GAT. Each module incorporated attentive-layer embedding for both atom and full molecule levels. After that, the outputs from these modules were concatenated and fed to fully connected layers for final prediction. The details of this approach can be found in the reference[10].

## RESULT AND DISCUSSION

Traditional Artificial Neural Networks (ANNs) often employ a black-box methodology for problem-solving, where their primary goal is to identify patterns or relationships within raw input data, without an explicit focus on underlying physical laws or domain-specific knowledge. While these approaches prove effective in numerous scenarios, they may fail to capture crucial

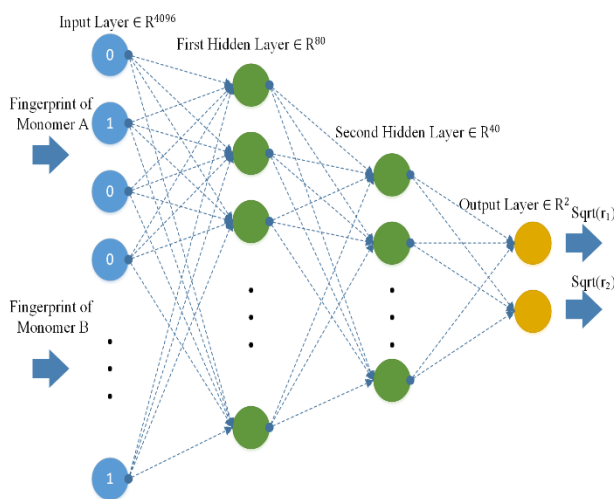


Figure 3. Representation of MLP Architecture

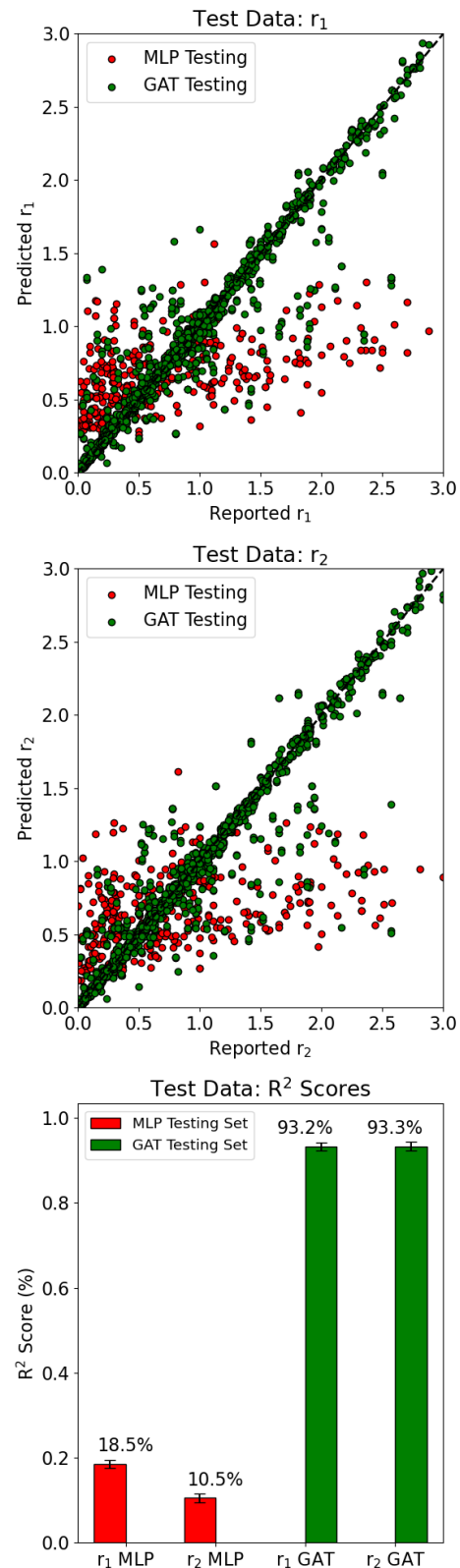
insights tied to the intrinsic properties or governing principles of the system in question.

In contrast, there is an emerging interest in the development of specialized ANNs tailored for particular data types and tasks. For example, Graph Attention Networks (GATs) have shown superior performance compared to Multi-Layer Perceptron (MLPs) in handling graph-structured data. This advancement is largely due to GATs' ability to more accurately represent and process-relational and structural information inherent in such data. Given that chemical science data fundamentally relates to molecular structures, which are naturally representable as graphs, Graph Neural Networks, including GATs, emerge as particularly promising for tackling challenges in polymer science. Their capacity to intuitively map and interpret the complex, interconnected nature of molecular structures positions them as a fitting choice for this field.

In this study, an MLP model was developed for the prediction of the reactivity ratio of paired monomers in copolymerization. It was observed that a regularized MLP predicted the reactivity ratios with improved accuracy, yet the Mean Square Error (MSE) could not be reduced to less than 0.1 in the evaluation of the test dataset. Although this MSE is substantially lower than those reported in other references[8], the model does not seem to be entirely appropriate for accurate prediction in this context.

Figure 4. compares the  $R^2$  score on the test dataset when MLP and GAT are used as the reactivity ratio estimator. As shown in the figure, GAT significantly outperforms in comparison to MLP in the prediction of  $r_1$  and  $r_2$ . While the  $R^2$  score could hardly be more than 10% in using MLP, the GAT could achieve  $85\pm 5\%$ [10].

It is evident that there is a substantial enhancement in the prediction of reactivity ratios for copolymerization when using GATs. This suggests that employing a graph-based machine-learning approach can outperform conventional ANNs. The key advantage lies in the graph-based model's ability to incorporate the natural intrinsic structure of the data into the training process, enabling it to capture patterns more effectively than other machine learning models, which may treat each model completely as a black box without considering the system's inherent nature.



**Figure 4.** Comparing the  $R^2$  score in test data for Graph Attention Network (GAT) and Multilayer Perceptron (MLP)

## CONCLUSION

Here, this study clearly demonstrates the superiority of Graph Attention Networks (GATs) over traditional Multilayer Perceptron (MLPs) in the context of predicting reactivity ratios for copolymers. The substantial improvement in prediction accuracy, as evidenced by the  $R^2$  scores, underscores the effectiveness of GATs in handling complex, graph-structured data inherent in chemical science. By integrating the intrinsic structural information of molecular data into their learning process, GATs not only outperform conventional ANNs but also pave the way for more nuanced and accurate models in polymer science. This comparison highlights the potential of graph-based machine learning approaches to revolutionize data analysis in fields where understanding the interconnected nature of data is critical.

## ACKNOWLEDGMENT

The authors acknowledge the support of the National Science Foundation (NSF) under the Award Number 2238147. The authors would like to extend their thanks to the Holland Computing Center at the University of Nebraska-Lincoln for the essential computational resources and support provided, which played a crucial role in the conduct of this research.

## REFERENCES

1. Z Dang, F Guo, Z Wu, K Jin, J Hao: Interface Engineering and Device Applications of 2D Ultrathin Film/Ferroelectric Copolymer P(VDF-TrFE). *Adv Phys Res*, 2:1 (2023), doi: 10.1002/apxr.202200038.
2. N Politakos: Block Copolymers in 3D/4D Printing: Advances and Applications as Biomaterials. *Polymers*, 15:2 (2023), doi: 10.3390/polym15020322.
3. M Gigli, N Lotti, M Gazzano, V Siracusa, L Finelli, A Munari, M Dalla Rosa: Biodegradable aliphatic copolyesters containing PEG-like sequences for sustainable food packaging applications. *Polym Degrad Stab*, 105:1, 96-106 (2014), doi: 10.1016/j.polymdegradstab.2014.04.006.
4. FR Mayo, C Walling: *Chem Rev*, 1950, 46, 191-287.
5. KI Takahashi, H Mamitsuka, M Tosaka, N Zhu, S Yamago: CoPoIDB: a copolymerization database for radical polymerization. *Polym Chem*, 15:10, 965-971 (2024), doi: 10.1039/d3py01372c.
6. M Dossi, D Moscatelli: A QM Approach to the Calculation of Reactivity Ratios in Free-Radical Copolymerization. *Macromol React Eng*, 6:2-3, 74-84 (2012), doi: 10.1002/mren.201100065.
7. T McDonald, C Tsay, AM Schweidtmann, N Yorke-

Smith: Mixed-integer optimization of graph neural networks for computer-aided molecular design. *Comput Chem Eng*, 185 (2024), doi: 10.1016/j.compchemeng.2024.108660.

8. K Farajzadehahary, X Telleria-Allika, JM Asua, N Ballard: An artificial neural network to predict reactivity ratios in radical copolymerization. *Polym Chem*, 14:23, 2779-2787 (2023), doi: 10.1039/d3py00246b.
9. W Sha, Y Li, S Tang, J Tian, Y Zhao, Y Guo, W Zhang, X Zhang, S Lu, Y-C Cao, S Cheng: Machine learning in polymer informatics. *InfoMat*, 3:4, 353-361 (2021), doi: 10.1002/inf2.12167.
10. T Nguyen, M Bavarian: Machine learning approach to polymer reaction engineering: Determining monomers reactivity ratios. *Polymer*, 275 (2023), doi: 10.1016/j.polymer.2023.125866.
11. P Veličković, G Cucurull, A Casanova, A Romero, P Liò, Y Bengio: Graph Attention Networks. [Online]. Available: <http://arxiv.org/abs/1710.10903> (2017).

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Integrated Design, Control, and Techno-Ecological Synergy: Application to a Chloralkali Process

Utkarsh Shah<sup>a,b</sup>, Akshay Kudva<sup>b</sup>, Kevin B. Donnelly<sup>b</sup>, Wei-Ting Tang<sup>b</sup>, Bhavik R. Bakshi<sup>b,c\*</sup>, and Joel A. Paulson<sup>b\*</sup>

<sup>a</sup> Google Research, Seattle, WA, USA

<sup>b</sup> The Ohio State University, Department of Chemical and Biomolecular Engineering, Columbus, OH, USA

<sup>c</sup> Arizona State University, School for Engineering of Matter, Transport and Energy, Tempe, AZ, USA

\* Corresponding Authors: [bhavik.bakshi@asu.edu](mailto:bhavik.bakshi@asu.edu), [paulson.82@osu.edu](mailto:paulson.82@osu.edu)

## ABSTRACT

The integrated design and control (IDC) framework is becoming increasingly important for systematic design of flexible manufacturing and energy systems. Recent advances in computing and derivative-free optimization have enabled more tractable solution methods for complex IDC problems that involve, e.g., multi-period dynamics, the presence of high-variance and non-stationarity probabilistic uncertainties, and mixed-integer control/scheduling decisions. Parallely, developments in techno-ecological synergy (TES) have allowed co-design of industrial and environmental systems that have been shown to lead to win-win solutions in terms of the economy, ecological, and societal benefits. In this work, we propose to combine the IDC and TES frameworks to more accurately capture the real-time interactions between process systems and the surrounding natural resources (e.g., forests, watersheds). Specifically, we take advantage of (multi-scale) model predictive control to close the loop on a realistic high-fidelity simulation of the overall TES system. Since this closed-loop simulation is computationally expensive, we propose to solve the resulting design problem using a data-efficient constrained Bayesian optimization method. We demonstrate that the new perspective offered by the proposed TES-IDC framework leads to robust win-win solutions that can more effectively handle uncertainty in future disturbances compared to technology-only solutions on a chloralkali manufacturing unit built in an urban forest.

**Keywords:** Sustainable design, Model predictive control, Bayesian optimization, Uncertain systems

## INTRODUCTION

Within the discipline of process systems engineering (PSE), the systematic and simultaneous consideration of design and time-varying operation parameters has become paramount for the creation of next-generation manufacturing and energy systems. Over the last four decades, there have been several contributions in PSE that seek to optimize process design with the consideration of future uncertainties and operational decisions. One of the earliest works that provided the quantitative framework for assessing various designs and their ability to handle uncertainty for steady state systems was the concept of flexibility analysis [1, 2], which was later extended to dynamical systems by inclusion of time varying

constraints [3]. More recently, flexibility analysis of black-box functions has also been proposed [4–6]. One of the major shortcomings of the flexibility approach was the fact that this paradigm assumed analysis of open-loop systems, which entails that an attainable operating point according to flexibility analysis, would in fact be intractable under closed-loop conditions. During the development of flexibility analysis, another concept was parallely developed, which assess controllability of process design, namely, the dynamic resilience [7] which was extended to account for other inherent characteristics of the system, such as constraints on manipulated variables, plant-model mismatch, etc. These traditional engineering design approaches, however, considered design and control of processes and systems as separate



problems. The design problem establishes the flow of components within a process, while the control problem addresses the maintenance of process conditions under disturbances and constraints. Addressing these problems separately can lead to decisions that result in sub-optimal process performance operationally, economically, or sustainably. Thus, the advent of concurrent design and control methodologies for process systems gained momentum in the early 1990s [8].

To improve the flexibility and resilience of process systems, the integrated design and control (IDC) framework relies on the fact that achievable plant dynamics are completely tied to the selected process design [9]. Early IDC literature [10–12] framed the problem to consist of an economic objective function constrained by the operating capabilities of classical control strategies such as PI controllers. Incorporation of more advanced optimization-based controller strategies such as model predictive controllers (MPC) were proposed in [13, 14], with assumptions that simplified the overall optimization problem. One can apply coarse approximations to the problem formulation; however, this will fundamentally degrade the quality of the problem/solution – if we miss out on these realistic phenomena then it is not clear that IDC provides any added value over traditional two-stage “design then control” strategies. The requirement of high-fidelity simulated solutions is buttressed by advancements in computational resources that allow implementation of powerful optimization methods, and detailed simulations which can span multiple spatiotemporal scales. The standard formulation of an IDC problem is a challenging stochastic optimization problem; although the design variables are often finite dimensional, the set of all possible feedback control policies is infinite dimensional. Recent work [15, 16] has shown that a combination of efficient derivative-free optimization (DFO) solvers and low-dimensional decision rule approximations can be used to tackle the challenging optimization problem.

These works, however, only consider technocentric solutions and don’t incorporate the interactions between process operations with nature. The coupling of technocentric and nature-based solutions is important because multi-scale dynamics results in complex, unpredictable systems. Historically, engineering has developed a tendency to modify and dominate ecosystems to improve predictability. Building large dams, canals and levees are a consequence of engineering’s tendency to convert the homeorhetic nature of ecosystems into homeostatic behavior. As another example, for air quality regulation, conventional engineering systems would recommend fast-growing genetically modified trees that provide constant air quality regulation services throughout the year. In the short run, such highly engineered systems would be able to meet human needs, by transferring variability to other parts of watershed or airsheds. Unfortunately,

this improvement in “efficiency” comes at the cost of resiliency. Due to poor biodiversity, genetically engineered trees would be more susceptible to pests or diseases as compared to natural forests. Similarly, floods and hurricanes can damage dams and levees leaving them inoperable. To analyse the interaction between nature and technology, we refer to techno-ecological synergy (TES) [17], which is a systems-based framework to account for ecosystems capacity and benefit from its synergy with technological systems. TES offers an opportunity towards resilient systems when compared to technological only systems. However, the resilience comes at a cost of homeorhetic behavior. For gaining maximum benefits from TES, engineering design and operation needs to evolve beyond hemostatic behavior towards a system that adapts nature’s intermittency. In this work, we address the open question - *how to systematically solve the IDC problem with embedded real-time interactions between nature and technology?*

## BACKGROUND AND PRELIMINARIES

In this section we will briefly discuss two important components of the proposed framework, namely the IDC formulation, and Bayesian optimization, which is our selected method for tackling the IDC problem.

### Integrated Design and Control under Uncertainty

The integrated design and control problem can be formulated in terms of the following stochastic optimization problem:

$$\min_{(d,z) \in \mathcal{D}, \mathcal{Z}} C(d) + \mathbb{E}_{\omega} \{O(d, z, \omega)\}, \quad (1)$$

where  $d \in \mathbb{R}^{n_d}$  are the design variables that must be selected from a compact set  $\mathcal{D} \in \mathbb{R}^{n_d}$ ;  $z \in \mathbb{R}^{n_z}$  are the control decisions restricted to compact set  $\mathcal{Z} \in \mathbb{R}^{n_z}$ , which parameterize recourse/control actions in response to uncertainties (only impact operating cost of the process);  $\omega \in \mathbb{R}^{n_{\omega}}$  is a sequence of random variables that describe the complete set of exogenous disturbances/uncertainties in the process. The objective function (1) is the composition of two functions, namely, the capital expenditure  $C(d): \mathcal{D} \rightarrow \mathbb{R}$ , and the expected value  $\mathbb{E}_{\omega}$  of the operating expenditure  $O(d, z(d), \omega)$  with respect to  $\omega$ . In this work, we are interested in computing realistic estimates of operating costs  $\mathbb{E}_{\omega} \{O(d, z, \omega)\}$ , meaning we do not want to make coarse approximations (e.g., steady-state operation). To achieve this, we assume that the system dynamics can be represented by the following generalized time-varying nonlinear system in the presence of uncertainty:

$$\begin{aligned} x_{t+1}(d, z, \omega) &= h_t(x_t(d, z, \omega), u_t(x_t(d, z, \omega), w_t, d), \\ x_0(d, z, \omega) &= b_0(d), \end{aligned} \quad (2)$$

where  $t \in \mathcal{T} := \{0, \dots, T-1\}$  is the discrete time index that ranges over finite number of  $T$  time stages;  $\omega$  is a sequence of random disturbances  $\omega \in \mathcal{W} \subset \mathbb{R}^{n_\omega}$ , which is assumed to be measurable at corresponding  $t \in \mathcal{T}$ ;  $u_t(d, z, \omega) \in \mathcal{U} \subset \mathbb{R}^{n_u}$  are control inputs parameterized by  $z$ ; and  $x_t(d, z, \omega) \in \mathcal{X} \subset \mathbb{R}^{n_x}$  are state variables with initial conditions  $b_0(d)$  that depend on design variables. Our only assumption is that we can simulate the process (2) to get future state sequences  $\{x_0(d, z, \omega), \dots, x_T(d, z, \omega)\}$  for specific realizations of the design  $d$ , control parameters  $z$ , and disturbances  $\omega$ . We do not need closed-form expressions for any of these components – we only need to be able to generate sample trajectories, which could be from some high-fidelity black-box process simulator.

The control inputs are assumed to be set by a parametrized *decision rule* (DR) with the following structure:

$$u_t(d, z, \omega) = \kappa_t(x_t(d, z, \omega), w_t, z), \quad (3)$$

where  $\kappa_t: \mathcal{X} \times \mathcal{W} \times \mathcal{Z} \rightarrow \mathcal{U}$  is a known function with free parameters  $z$ . The value of (3) is twofold: (i) it enables us to work in a reduced-dimensional space through proper selection of  $z$  and (ii) it enforces causality by design such that only past-revealed information can be exploited in the selection of the current control action. We typically select (3) to be some form of advanced control such as MPC due to its nice combination of performance and flexibility.

Given this dynamic model, the operating cost will typically be represented as a sum of stage costs over the entire system lifetime such as:

$$O(d, z, \omega) = \sum_{t=0}^{T-1} \ell_t(x_t(d, z, \omega), u_t(d, z, \omega), w_t, d) + \ell_T(x_T(d, z, \omega, \gamma), d), \quad (4)$$

where  $\ell_t: \mathcal{X} \times \mathcal{U} \times \mathcal{W} \times \mathcal{D} \rightarrow \mathbb{R}$  is the contribution to the operating cost at each time step  $t \in \mathcal{T}$ , and  $\ell_T: \mathcal{X} \times \mathcal{D} \rightarrow \mathbb{R}$  is contribution to the operating cost at the final time step (e.g., recoverable cost from equipment or product).

## The Bayesian Optimization Framework

Bayesian optimization (BO) is a powerful tool for solving black-box optimization problems in the presence of zeroth-order data and expensive to evaluate functions. The objective of BO is to globally minimize a cost function  $f$  such that  $f^* = \min_{\theta \in \Theta} f(\theta)$ , where domain  $\Theta$  is a closed set. The optimization framework usually begins with "warm-starting" a surrogate model with small number of randomly selected function evaluations over the domain of decision variables. In low-data regime, Gaussian process (GP) regression has been most widely used surrogate model since they are both *probabilistic* and *nonparametric*. A GP prior is placed on the objective function  $f(\cdot) \sim \mathcal{GP}(m(\cdot), k(\cdot, \cdot))$ , which can be fully specified by its mean function  $m(\cdot)$ , and covariance kernel  $k(\cdot, \cdot)$ . The mean function and covariance kernels for a pair of inputs  $\theta, \theta' \in \mathbb{R}^{n_\theta}$  is defined as follows:

$$m(\theta) = \mathbb{E}_f\{f(\theta)\}, \quad (5a)$$

$$k(\theta, \theta') = \mathbb{E}_f\{(f(\theta) - m(\theta))(f(\theta') - m(\theta'))\}. \quad (5b)$$

Given a set of inputs  $\theta = \{\theta_1, \dots, \theta_n\}$ , we can update the GP model prior with available data using Gaussian conditioning formula and obtain the following analytical form for mean and variance for a test point  $\theta$ , and noisy measurements  $y_i = f(\theta_i) + \epsilon_i$  for all  $i = 1, \dots, n$  [18]:

$$\mu_n(\theta) = m(\theta) + k(\theta, \theta)(k(\theta, \theta) + \sigma_\epsilon^2 I_n)^{-1}(y_{1:n} - m(\theta)), \quad (6a)$$

$$\sigma_n^2(\theta) = k(\theta, \theta) - k(\theta, \theta)(k(\theta, \theta) + \sigma_\epsilon^2 I_n)^{-1}k(\theta, \theta), \quad (6b)$$

Given a predictive surrogate model, one must now select a so-called "acquisition function"  $\alpha_n(\theta)$  whose value at any  $\theta \in \Theta$  provides a good quantitative measure of the (expected) benefit in querying  $f$  at this point in the future. By definition, the next best sample can then be found by solving an optimization problem:

$$\theta_{n+1} = \max_{\theta \in \Theta} \alpha_n(\theta). \quad (7)$$

Note that this problem is significantly easier to solve the original problem since it is only defined in terms of the GP surrogate model, which can be cheaply evaluated (in the low data limit) and whose derivatives can be computed using established automatic differentiation schemes. A variety of acquisition functions have been proposed for single-level optimization problems [19, 20]. The same concept can also be generalized to other types of problems with additional structure such as optimization with black-box constraints [21] [22], multi-objective optimization [23], robust optimization [24–26], just to name a few.

In this work, we use a high-fidelity, offline simulator to compute cost and operating expenditures for a given  $(d, z)$ . Similar to previous work [15], the decision variable of BO framework is given by the concatenation of design variables and control decisions of the IDC problem, i.e.,  $\theta := (d, z)$ . We leverage steps (5-7) to sequentially search over  $\theta \in \Theta := \mathcal{D} \times \mathcal{Z}$ , and find the optimal solution of (1) in a data-efficient manner.

## CASE STUDY: CHLORALKALI SYSTEM DESIGN

We consider a chloralkali manufacturing process located in Freeport, Texas, which was originally studied by Shah and Bakshi [27]. The energy-intensive process derives its energy from a coal-fired generator, which results in air emissions such as nitrogen dioxide (NO<sub>2</sub>), sulfur dioxide (SO<sub>2</sub>), carbon dioxide (CO<sub>2</sub>) and particulate matter. The NO<sub>2</sub> emissions acts as a precursor to ground level ozone O<sub>3</sub>. Our goal is to design a minimal cost system that can provide effective air quality regulation (such that it is potentially a net benefit to society), even under unknown variability in weather conditions. We specifically

want to leverage the natural ability of trees to uptake air emissions, while accounting for dynamic uncertainty in solar irradiance, wind velocity, and leaf surface area.

## Model Description

For the decision rule, we formulate an economic MPC (eMPC), whose objective is to minimize the total annualized cost,  $Z^{TAC}$  of the chloralkali production unit with hourly sampling, which corresponds to  $T = 8760$  time-steps. Two design variables ( $n_d = 2$ ) can be varied at the BO level; first, is the technocentric choice of the size of selective catalytic reactor (SCR),  $S^{SCR}$  to treat  $\text{NO}_2$ , whereas the second is the eco-centric option of reforestation area,  $A^r$ , in the vicinity of the facility. There are seven uncertain variables in the air quality regulation problem - namely  $\text{NO}_2$  concentration in ambient air,  $\text{O}_3$  concentration in ambient air,  $\text{NO}_2$  dry deposition rate, atmospheric mixing height, photo-chemical destruction rate of  $\text{NO}_2$ , and photo-chemical production rate of  $\text{O}_3$ . Presence of uncertain variables leads to high computational demand because the eMPC not just has to predict the future behavior of the uncertain variable, but also be robust against future changes or deviations. We select a control horizon of one hour ( $N_u = 1$ ) and we take advantage of the fixed total time-steps  $T$  of the closed loop simulation and consider a prediction horizon of  $N_p = T - t_0 + 1$ , where  $t_0$  is the initial time step of simulation. For example, at first hour,  $t_0 = 1$  and  $N_p = 8760$ ; after implementation of control action recommended by the DR, at the next time step (i.e., second hour),  $t_0 = 2$ , prediction horizon decreases to  $N_p = 8759$ . Although the proposed shrinking horizon approach improves computational tractability of the IDC simulation, there would still exist the computational burden for long prediction horizons. To alleviate this drawback, we divide the prediction horizon into two parts; the first 24 hours after  $t_0$ , for which we build a multivariate time series forecasting model using a Gated Recurrent Unit (GRU) neural network using past data of 168 hours; the predictions for time steps after 24 hours to  $N_p$  are allotted values using time series clustering of historical data. The number of time series clusters ( $|\mathcal{D}|$ ) is set to be one of the control decisions ( $n_z = 1$ ) at the BO level. The DR in this setting is given by the minimizer of the below optimization problem:

$$\min_{\{F_{Cl_2}(t+k|t)\}_{k=1}^{N_u}} Z^{TAC}, \quad (8a)$$

subject to chlorine production model and power requirement,

$$P_t = f_1(F_t^{Cl_2}), \forall t \in 1, \dots, N_p, \quad (8b)$$

supply-demand accounting for  $\text{NO}_2$  and  $\text{O}_3$ ,

$$D_{i,t}^E = f_2(P_t, S^{SCR}), \forall t \in 1, \dots, N_p, i = \{\text{NO}_2, \text{O}_3\}, \quad (8c)$$

$$S_{i,t}^E = f_3(A^r), \forall t \in 1, \dots, N_p, i = \{\text{NO}_2, \text{O}_3\}, \quad (8d)$$

air quality constraints,

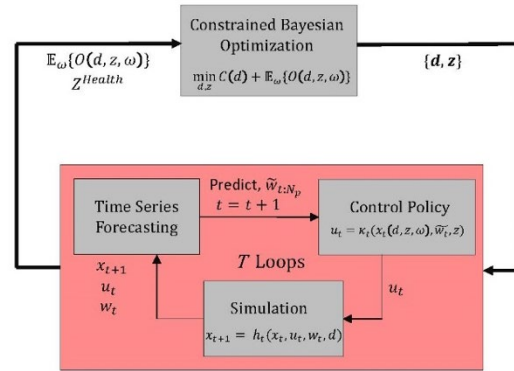
$$C_{i,t}^f = f_4(C_{i,t}^0, D_{i,t}^E, S_{i,t}^E), \forall t \in 1, \dots, N_p, i = \{\text{NO}_2, \text{O}_3\}, \quad (8e)$$

and economic and health impact cost calculation,

$$Z^{TAC} = \sum_{t=1}^{N_p} f_5(F_t^{Cl_2}, A^r, S^{SCR}), \quad (8f)$$

$$Z^{Health} = \sum_{t=1}^{N_p} f_6(C_{i,t}^f, C_{i,t}^0), \quad (8g)$$

where  $F_{Cl_2}$  is the production rate of chlorine. For details about the constraints in the control policy, the reader is referred to Shah and Bakshi [28] (see supporting information). We refer to the same study to define societal health impact cost,  $Z^{Health}$ , which reflects the monetary valuation of the chloralkali facility's operation on the health of neighbourhood. We use iTree Eco v6 to calculate hourly dry deposition velocity of  $\text{NO}_2$  and  $\text{O}_3$ , and atmospheric mixing height. The deposition velocities for year 2006–2015 are obtained using local meteorological conditions obtained from the National Climate Data Center (NCDC) for Clover Field Airport, which is the nearest good quality data source to our study site. The case study assumes the vicinity of the manufacturing facility to be barren i.e. having zero supply capacity and availability of  $15 \text{ km}^2$  for restoration to increase capacity. The cost of reforestation is set at  $\$75/\text{km}^2$  [29]. Native species of White Ash is considered to approximate the capacity of the restored forest ecosystem. Constrained BO [21, 30] is used to make design decisions on  $d = \{S^{SCR}, A^r\}$ , and  $z = \{|\mathcal{D}|\}$ , with constraints that enforce  $Z^{Health} \leq 0$ . The workflow is summarized in Figure 1.

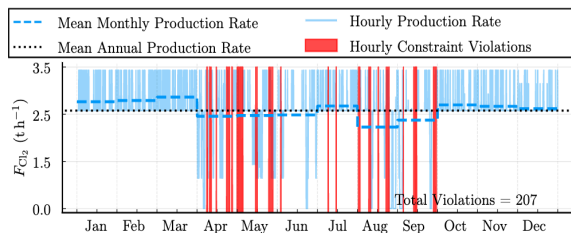


**Figure 1:** Illustration of the proposed efficient TES-IDC framework combining MPC and BO.

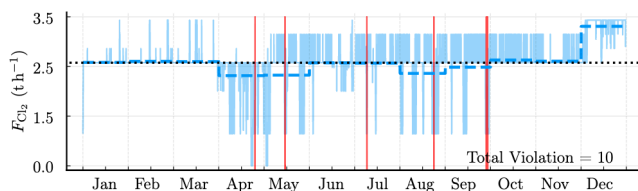
## Results and Discussion

The BO algorithm applied to the TES-IDC problem converged to a solution of reforestation area of  $11.798 \text{ km}^2$ , 17 representative cluster days, and no SCR requirement. The optimal total annualized cost of production has a mean value of  $\$920\text{k}$  and standard estimation error of  $\$4.53\text{k}$ . The uncertainty in objective value results from consideration of 10 different meteorological years data used in the simulation. The optimal solution obtains a

societal health impact cost with mean \$-194k (i.e., a net-positive impact on society) and standard estimation error of \$57.72k. The operational solution needs to satisfy hourly air quality constraints. Due to forecasting error and prediction mismatch, it is not possible to satisfy constraints at each and every time step. Thus, the magnitude of constraint violation (8 b-g) is minimized by adding a corresponding weighted exact penalty function to the objective function in (8a), which leads enforcement of soft constraints in the optimization problem (8). The optimal operational solution is expected to have a mean of 10.3 constraints violations with standard measurement error of 5 instances. This is compared with a retrospective scheduling approach, which uses past data to set an operating schedule for the future plant operation [27, 28, 31]. The optimal operational schedule for meteorological year 2010 obtained from TES-IDC framework is depicted in Figure 2b. It should be noted that the operational schedule was obtained in closed-loop manner, where the optimal control actions were updated with each realization of uncertainty. Figure 2a shows retrospective operational schedule of year 2009 applied to year 2010. The retrospective solution (207 violations) has an order of magnitude more hourly constraint violations as compared to TES-IDC solution (10 violations). Therefore, the TES-IDC optimal configuration is robust to changes in meteorological conditions. Assuming meteorological conditions remain similar to that of the last 10 years, the optimal solution (design conditions and controller configuration) found by BO can maintain optimality and constraint satisfaction of the chloralkali unit.



**(a)** Simulation of operations in year 2010 using an optimal retrospective solution obtained for meteorological year 2009 using the previous method by Shah and Bakshi [27].



**(b)** Simulation of operations in year 2010 using solution obtained from proposed IDC-TES. The number of instances of air quality constraint violation decreases significantly as compared to previous retrospective approach.

**Figure 2:** Operational schedule using (a) retrospective and (b) TES-IDC approach. TES-IDC approach allows for better constraint satisfaction as compared to previous approach.

## CONCLUSIONS

In this work, we consider real-time interactions between nature and technology in the IDC simulation, referred to as TES-IDC, which has not been previously addressed in the literature. Furthermore, the proposed solution does not make any assumptions about continuity, linearity or convexity of the underlying design and control problem. Application of TES-IDC framework on air quality regulation problem demonstrates the ability of technological systems to account for intermittency and variability of ecosystems in real-time operations. The operation using TES scheme can ensure long-term net-positive manufacturing goals and societal gain. Explicitly accounting for uncertainty through scenarios, allows us to report expected values of objective along with a measure of uncertainty. This framework motivates application in design and operation of flexible renewable systems. The case study considered in this work has assumed constant spatial capacity of ecosystem. Future work is to formalize the TES-IDC framework and provide the PSE community with access to open-source codes for benchmark. Another interesting avenue for further investigation is the inclusion of spatial dimension to the IDC problem to obtain spatio-temporally explicit work design and operation of TES-IDC problems.

## REFERENCES

- Grossmann, I.E., Calfa, B.A., Garcia-herreros, P.: Evolution of Concepts and Models for Quantifying Resiliency and Flexibility of Chemical Processes. 1–34 (2013)
- Swaney, R.E., Grossmann, I.E.: An index for operational flexibility in chemical process design. Part I: Formulation and theory. *AIChE Journal*. 31, 621–630 (1985). <https://doi.org/10.1002/aic.690310412>
- Dimitriadis, V.D., Pistikopoulos, E.N.: Flexibility Analysis of Dynamic Systems. *Ind. Eng. Chem. Res.* 34, 4451–4462 (1995)
- Rogers, A., Ierapetritou, M.: Feasibility and flexibility analysis of black-box processes part 2: Surrogate-based flexibility analysis. *Chem Eng Sci.* 137, 1005–1013 (2015). <https://doi.org/10.1016/j.ces.2015.06.026>
- Kudva, A., Tang, W.-T., Paulson, J.A.: Robust Bayesian optimization for flexibility analysis of expensive simulation-based models with rigorous uncertainty bounds. *Comput Chem Eng.* 181, 108515 (2024). <https://doi.org/10.1016/j.compchemeng.2023.108515>
- Boukouvala, F., Ierapetritou, M.G.: Feasibility analysis of black-box processes using an adaptive



- sampling Kriging-based method. *Comput Chem Eng.* 36, 358–368 (2012).  
<https://doi.org/10.1016/j.compchemeng.2011.06.005>
7. Morari, M.: Design of resilient processing plants-III A general framework for the assessment of dynamic resilience. *Chem Eng Sci.* 38, 1891 (1983)
  8. Burnak, B., Diangelakis, N.A., Pistikopoulos, E.N.: Towards the grand unification of process design, scheduling, and control-utopia or reality? *Processes.* 7, (2019).  
<https://doi.org/10.3390/pr7070461>
  9. Vega, P., Lamanna de Rocco, R., Revollar, S., Francisco, M.: Integrated design and control of chemical processes - Part I: Revision and classification. *Comput Chem Eng.* 71, 602–617 (2014).  
<https://doi.org/10.1016/j.compchemeng.2014.05.010>
  10. Narraway, L.T., Perkins, J.D., Barton, G.W.: Interaction between process design and process control: economic analysis of process dynamics. (1991)
  11. Shah, N., Pantelides, C.C., Sargent, R.W.H.: The Design and Scheduling of Multipurpose Batch Plants. *IFAC Proceedings Volumes.* 25, 203–208 (1992)
  12. Luyben, M.L., Floudas, C.A.: Analyzing the interaction of design and control-I. A multiobjective framework and application to binary distillation synthesis. *Comput Chem Eng.* 18, 933–969 (1994)
  13. Bemporad, A., Morari, M., Dua, V., Pistikopoulos, E.N.: The explicit linear quadratic regulator for constrained systems. *Automatica.* 38, 3–20 (2002)
  14. Sakizlis, V., Perkins, J.D., Pistikopoulos, E.N.: Parametric controllers in simultaneous process and control design optimization. *Ind Eng Chem Res.* 42, 4545–4563 (2003).  
<https://doi.org/10.1021/ie0209273>
  15. Sorourifar, F., Choksi, N., Paulson, J.A.: Computationally efficient integrated design and predictive control of flexible energy systems using multi-fidelity simulation-based Bayesian optimization. *Optim Control Appl Methods.* (2021).  
<https://doi.org/10.1002/oca.2817>
  16. Choksi, N., Paulson, J.A.: Simulation-based Integrated Design and Control with Embedded Mixed-Integer MPC using Constrained Bayesian Optimization. In: *American Control Conference. IEEE* (2021)
  17. Bakshi, B.R., Ziv, G., Lepech, M.D.: Techno-ecological synergy: A framework for sustainable engineering. *Environ Sci Technol.* 49, 1752–1760 (2015). <https://doi.org/10.1021/es5041442>
  18. Rasmussen, C.E., Williams, C.K.I.: *Gaussian Processes for Machine Learning.* The MIT Press (2005)
  19. Jones, D.R., Schonlau, M., Welch, W.J.: Efficient Global Optimization of Expensive Black-Box Functions. *Journal of Global Optimization.* 13, 455–492 (1998)
  20. Srinivas, N., Krause, A., Kakade, S.M., Seeger, M.W.: Information-theoretic regret bounds for Gaussian process optimization in the bandit setting. *IEEE Trans Inf Theory.* 58, 3250–3265 (2012). <https://doi.org/10.1109/TIT.2011.2182033>
  21. Gardner, J.R., Kusner, M.J., Xu, Z.E., Weinberger, K.Q., Cunningham, J.P.: Bayesian Optimization with Inequality Constraints. In: *Proceedings of Machine Learning Research* (2014)
  22. Lu, C., Paulson, J.A.: No-Regret Bayesian Optimization with Unknown Equality and Inequality Constraints using Exact Penalty Functions. *Dynamics and Control of Process Systems, including Biosystems.* (2022)
  23. Daulton, S., Eriksson, D., Balandat, M., Bakshy, E.: Multi-Objective Bayesian Optimization over High-Dimensional Search Spaces. In: *Proceedings of Machine Learning Research* (2022)
  24. Bogunovic, I., Jegelka, S., Scarlett, J.: Adversarially Robust Optimization with Gaussian Processes. In: *Advances in Neural Information Processing Systems* (2018)
  25. Paulson, J.A., Makrygiorgos, G., Mesbah, A.: Adversarially robust Bayesian optimization for efficient auto-tuning of generic control structures under uncertainty. *AIChE Journal.* 1–21 (2022).  
<https://doi.org/10.1002/aic.17591>
  26. Kudva, A., Sorourifar, F., Paulson, J.A.: Constrained robust Bayesian optimization of expensive noisy black-box functions with guaranteed regret bounds. *AIChE Journal.* (2022).  
<https://doi.org/10.1002/aic.17857>
  27. Shah, U., Bakshi, B.R.: Accounting for nature's intermittency and growth while mitigating NO<sub>2</sub> emissions by technoecological synergistic design—Application to a chloralkali process. *J Adv Manuf Process.* 1, (2019).  
<https://doi.org/10.1002/amp2.10013>
  28. Shah, U., Bakshi, B.R.: Improving Air Quality by Adapting Manufacturing to Pollution Uptake by Vegetation. Under submission. (2024)
  29. Sofie, A., Nielsen, E., Plantinga, A.J., Alig, R.J.: New Cost Estimates for Carbon Sequestration Through Afforestation in the United States. (2014)
  30. Balandat, M., Karrer, B., Jiang, D.R., Daulton, S., Letham, B., Wilson, A.G., Bakshy, E.: BoTorch: A Framework for Efficient Monte-Carlo Bayesian Optimization. (2019)



31. Palys, M.J., Kuznetsov, A., Tallaksen, J., Reese, M., Daoutidis, P.: A Novel System for Ammonia-Based Sustainable Energy and Agriculture: Concept and Design Optimization. *Chemical Engineering and Processing - Process Intensification*. 11–21 (2019)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Use of Discrete Element Method to Troubleshoot Aesthetic Defects in Pharmaceutical Tablets

Jerrin Job Sibychan<sup>a\*</sup>, Nicola Sorace<sup>a</sup>, Jason Melnick<sup>b</sup>, Salvador Garcia Muñoz<sup>b</sup>, David Mota-Aguilar<sup>b</sup>, Eduardo Hernandez-Torres<sup>b</sup>, and David Boush<sup>a</sup>

<sup>a</sup> Altair Engineering Inc, Troy, Michigan, USA

<sup>b</sup> Eli Lilly and Company, Indianapolis, Indiana, USA

\* Corresponding Author: [jerrinjobs@altair.com](mailto:jerrinjobs@altair.com)

## ABSTRACT

Pharmaceutically elegant tablets are an expectation from pharmacists, health care providers and consumers for solid oral dosage forms. The presence of non-aesthetically pleasing defects in solid oral dosage forms can result in complaints back to the manufacturer and potentially non-compliance with medicines. The purpose of this study was to simulate and analyze the design of a tablet core and the aqueous film-coating process, to gain a better understanding of tablet defect generation, and to help eliminate the defects from the finished product. This evaluation employs Discrete Element Method (DEM) using the software product Altair® EDEM™ to understand the potential mechanisms that are causing the defects, based on the forces tablets experience in the coating operation, along with the number of tablet-to-tablet interactions that occur during the duration of the process. Defects observed during the scale up of the coating process to a commercial production scale confirmed the DEM results where physical damage was observed more on the edges of the tablets than the face of the tablets. Also based on the number of tablet-to-tablet interactions, operating the coating process under thermodynamically wetter processing conditions can result in elevated levels of picking and sticking defects being observed based on the specific tablet design evaluated. The results of these efforts allowed the manufacturing and development team to evaluate improvement opportunities not only in tablet design but also to re-evaluate the thermodynamic design space of the coating operation and the mechanical set up of the coating equipment.

**Keywords:** Defects, Discrete Element Method, EDEM, Pharmaceuticals, Round Concave Tablet, Solid Oral Dosage Forms, Tablet Coating

## INTRODUCTION

Pharmaceutically elegant tablets are an expectation from pharmacists, health care providers and consumers. One approach to ensure aesthetically pleasing dosage forms is to apply a film coating [1]. Not only does the coating aid in the appearance and identification of the tablet, but it also helps patients by protecting the product from light, aids in swallowability of the tablets and also protects the caregivers from accidental exposure to the drug substance contained within the film-coated tablet.

Periodically, non-aesthetically pleasing defects can be present in the finished product. These defects have no impact on the safety or efficacy of the tablet but may

garner complaints from the marketplace and impact the company's brand and reputation and patient compliance to the medicine. This is of extreme importance for the Japan market. If the defect level present in each batch of drug product does not meet the current standards, the batch can be considered unfit for sale and then must be sorted or in a worse case discarded. Destruction of a batch of product not only interrupts the manufacturer supply chain because the batch will need to be replaced but can also impact the patients depending on the inventory levels. Batches can be sorted by employing an inspection/sorting step. This step can be manual, mechanical, or electronic based on the nature of the defect. The sorting technique could result in a considerable time

delay in releasing the batch.

Many variables come into effect when applying aqueous film coatings to tablet cores that can impact the elegance of the coated tablet. These include the physical design of the tablet, mechanical strength of the tablet core, mechanical design of the coating equipment, the thermodynamics of the coating operation and handling of coated tablets once the coating process is complete. Each one of these variables will impact the type, severity and number of defects created. In some cases, it takes a combination of the variables to result in the generation of a specific defect.

The purpose of this study was to simulate and analyze the design of a specific tablet core shape with the commercial production coating equipment design i.e., pan size, number of baffles and baffle design and coating pan load. This effort would allow for a better appreciation for the interaction between the tablet shape and the coating equipment and the operation parameters. The defects being observed were a result of a transfer and scale up of a manufacturing process from pilot to commercial scale. The two main defects being observed were edge damage, Figure 1 and picking and sticking, Figure 2. These defects were observed during the end of the batch statistical inspection of the bulk finished product. Understanding of the interactions and the defects being observed can then be applied back to the coating operation, allowing for the appropriate actions to be taken to aid in the reduction or elimination of the defect in future batches of finished product, resulting in increased quality and production yields.



**Figure 1.** Tablet edge damage defects



**Figure 2.** Picking and sticking defects on tablet faces

## LITERATURE REVIEW

Since the introduction of discrete element method (DEM) simulation techniques [2], there has been an increasing interest in utilizing DEM to improve understanding of the intricate behavior exhibited by particulate systems. Furthermore, there has been a recent expansion in its application across various industries [3], and more recently in the pharmaceutical industry [4]. The coating operation has been one of the very common applications where the coating uniformity and breakage are the main concerns in the process. Although there are many computational studies on the coating of the tablets [5][6][7], there are few computational studies on the breakage of the tablet core itself.

Computational analyses of breakage have been done for general cases wherein the impact of agglomerate particles across various impact velocities [8] and impact angles [9] were examined. However, there has been little work done on pharmaceutical tablets as opposed to general spherical agglomerates. Bharadwaj et al [10] used DEM to compare the characteristic forces exerted on tablets in a friability tester with those experienced during the operation of lab-, pilot-, and commercial-scale film coating pans. But this study also lacks the prediction of the likelihood of tablet breakage during the coating process. Ketterhagen et al [11] overcomes this limitation by providing a probabilistic model to predict tablet fracture in the film coating process. Ketterhagen uses the impact velocity of the tablets during contact to predict probability of breakage of a tablet using a population balance model proposed by Vogel & Peukert [12][13].

All of the above studies mentioned above have been performed for spherical agglomerates using a “glued-sphere” approach. This present study extends from the previous studies and takes into account the physical shape of a pharmaceutical polyhedral tablet and provides a surface distribution of forces on regions of the tablet to predict localized breakage.

## METHODOLOGY

### Discrete Element Method

Discrete Element Method (DEM) is a numerical method which is used to simulate bulk material. This method was first developed by Cundall and Strack [2] in which all the individual particles in a bulk system are modelled as discrete elements. A Lagrangian-based approach is used in DEM to track these individual particles in the system at every timestep. The contacts between particles are detected at each timestep, and the contact forces are calculated based on this contact. Newton’s second law of motion is applied to each particle to

calculate the accelerations of each particle. Explicit numerical integration is then done twice to calculate the position of each particle in the subsequent timestep. Contact detection is carried out again for these updated positions, and then this cycle repeats for every timestep. DEM is the highest fidelity numerical modelling approach for particulate solids available because the discrete nature of these particulate solids is not ignored, although it comes with very high computational expense compared to other methods.

DEM was traditionally developed for spherical particles. For non-spherical arbitrarily shaped particles in a bulk, the particles are approximated by idealized numerical elements such as clumped spheres or multi-spheres, and this multi-sphere approach has been widely used and validated for particle distributions which do not have a fixed shape or size [14]. This is due to the simplicity of contact detection algorithms for spheres. However, in this study, a polyhedral particle shape, standard round concave shape [15], was used since tablets in pharmaceutical industries generally include sharp edges that are difficult to capture with spheres. Consequently, the contact detection algorithm needs to be modified to accommodate such polyhedral shapes.

## Governing Equations

This study employs DEM using the software product Altair® EDEM™. The default contact model in EDEM is based on the Hertzian theory of contact mechanics [16]. However, for polyhedron-shaped particles, the force equations are modified to account for non-spherical shapes, according to Nassaeur and Kuna [17]. Additionally, a spinning friction model was also introduced to account for the friction that would occur if a particle face were rotating against another particle or geometry.

Contact forces between two particles in the normal direction are given by the following equation:

$$F_n = \left( \frac{0.62}{0.752} \right) \left( \frac{4}{3\sqrt{\pi}} \right) E^* \sqrt{V d_n}$$

where

$V$  = overlap volume between two particles during contact

$d_n$  = indentation depth which is the extension of the overlapping region in the direction of force

$E^*$  = equivalent Young's modulus, defined as

$$\frac{1}{E^*} = \frac{1 - \nu_1^2}{E_1} + \frac{1 - \nu_2^2}{E_2}$$

Dissipative effects were also taken into account by modelling a damping force in the normal direction. The equation for the damping force is given by the following:

$$F_{n,d} = 2 \sqrt{\frac{5}{6} \left( \frac{-\log e}{\log^2 e + \pi^2} \right)} \sqrt{2E^* m^*} \sqrt{\frac{V}{\pi d_n}} v^{rel}_n$$

where

$e$  = coefficient of restitution between two particles

$v^{rel}_n$  = magnitude of relative velocity between both particles in the direction of normal force

$m^*$  = equivalent mass, defined as

$$\frac{1}{m^*} = \frac{1}{m_1} + \frac{1}{m_2}$$

The tangential force between two particles during contact is given by the following equation which is also limited by the Coulomb friction model  $\mu F_n$  where  $\mu$  is the coefficient of friction.

$$F_t = -8G^* \sqrt{\frac{V}{\pi d_n}} d_t$$

where

$d_t$  = extension of the overlapping region in the direction normal to the force

$G^*$  = equivalent shear modulus, defined as

$$\frac{1}{G^*} = \frac{2 - \nu_1}{G_1} + \frac{2 - \nu_2}{G_2}$$

Dissipative effects in the tangential direction were also considered, and are given by the following equation:

$$F_{t,d} = -2 \sqrt{\frac{5}{6} \left( \frac{-\log e}{\log^2 e + \pi^2} \right)} \sqrt{8G^* m^*} \sqrt{\frac{V}{\pi d_n}} v^{rel}_t$$

where

$v^{rel}_t$  = magnitude of relative velocity in the tangential direction

The terms  $E$ ,  $\nu$ ,  $m$ , and  $G$  with subscripts 1 and 2 correspond to the Young's modulus, Poisson's ratio, mass, and Shear modulus of both particles in contact, respectively.

The introduction of spinning friction [18] was essential in this model to account for the friction between two particles whose faces are rotating against each other. This involves calculating the contact area and then applying the torque in the opposite direction to the normal component of the relative angular velocity. The equation for torque on the particle is given as follows:

$$M = \frac{2}{3} \mu F_n \sqrt{\frac{A}{\pi}}$$

where

$\mu$  = coefficient of friction

$A$  = normal area of the overlap region

This torque is limited to avoid numerical instability and any oscillating behavior when the angular velocity is

small. This was done by calculating the torque required to damp the angular velocity in one timestep.

$$M = \frac{0.125\omega_n^{rel} \min(I_1, I_2)}{\Delta t}$$

where

$\omega_n^{rel}$  = normal component of the relative angular velocity

$\Delta t$  = timestep used in the model

$\min(I_1, I_2)$  = minimum value of the moment of inertia of the two contacting particles

## Material model calibration

Material model calibration is a fundamental component of the Discrete Element Modelling methodology. Material model calibration typically consists of replicating a rheological measurement in the model, so that there is confidence in the model behaving in a physically accurate manner. Bulk density is a rheological measurement that needs to be replicated for any DEM model to capture the inertial effects. The other test done for this study to ensure the tablet-tablet interactions are modelled correctly, is the angle of repose test. Since a coating process is modelled for this study, both the bulk density and the angle of repose measurements were done for both core and coated tablets. The average value for both measurements were computed, and the average value was replicated within the DEM model.

For the bulk density measurement, tablets were filled in a known cylindrical volume, and was weighed. The mass of the bulk of the tablets was then divided by the known volume to calculate the bulk density. This process was done for both core and coated tablets. The average value of 601.9 kg/m<sup>3</sup> was replicated within a similar virtual cylinder within the DEM model. This involved a trial-and-error method with running multiple simulations with different parameters on the same setup, until the target bulk density was achieved.

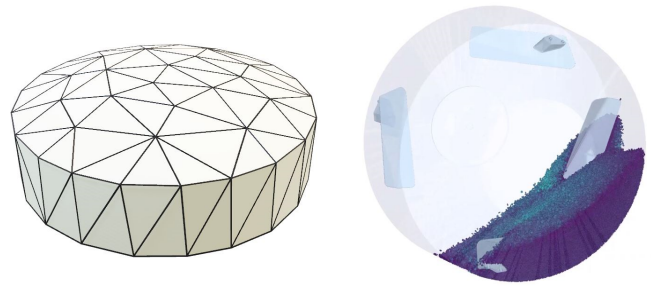
Similarly, for the angle of repose measurement, the tablets were allowed to form a pile on the horizontal after lifting a hollow cylinder filled with these tablets. The angle that this pile makes with the horizontal was measured, for both core and coated tablets respectively. Again, the average value of 35 degrees was replicated within a similar setup in the DEM model. The angle of repose is sensitive to the coefficient of friction values along with other parameters like the individual particle density and the coefficient of restitution, and therefore, these parameters were changed using a trial-and-error method until the DEM model reproduced the same pile with the same angle in the virtual setup.

## Process Modelling

The calibrated interaction parameters are used for the modelling of the industrial scale tablet coater operation. As mentioned earlier, the tablets are polyhedron-

shaped with 78 faces and approximately 10 mm in size, which is a very close estimation of the physical particle shape. The particle shape is shown in Figure 3.

A singular batch in the coater consists of approximately 200,000 tablets, therefore, the simulation was also done for the same load of tablet cores. The simulation was split into two stages – the filling stage and the process modelling stage. To introduce realistic settling of the tablets within the coater before the operation, the tablets were introduced into the coating pan from an opening on the side and allowed to settle under gravity. The coating pan was then rotated at 7 rpm, for 30 seconds of operation. The model was run on an NVIDIA Tesla V100 GPU card which took an approximate of 225 hours of clock time for the whole 30 seconds of operation.



**Figure 3:** Tablet shape used in DEM (left); EDEM simulation of the industrial tablet coating process (right).

All the forces during each contact of the tablet core, with another tablet core or with the coating pan itself, were calculated using the governing equations. However, data was saved at every 0.1 seconds interval due to practical hardware limitations. This data includes position and orientation of the tablets, their velocities, their contact forces and the location of the contact on both interacting bodies. In total, about 98 million contacts were recorded, for these 200,000 tablets at these 0.1 seconds interval during the whole 30 second operation, and analyzed. The force values on each vertex of the polyhedron during each of these contacts are then compared to determine the region of the polyhedron-shaped tablet with maximum propensity for breakage or defect formation.

## RESULTS

In a DEM simulation, for every collision, the impact force increases until it reaches a peak value, and then it decreases with the particles separating from each other. The value of the forces acting on each of the tablets, in this case, and their location of impact for each contact, or collision, was recorded at every 0.1 seconds. Due to this frequency of data recording, it is possible that the peak forces may not be recorded if the peak occurs between data save periods. However, since steady state is achieved relatively early in the process, at about



5 seconds, the simulation consists of random sampling of contacts during twenty-five seconds of steady state operation. This study assumes that the highest recorded forces are indicative of the impact events that lead to tablet defects. Hence, the magnitude of the force itself is insignificant, but the relative distribution of forces is bound to give more insight into the results.

### Force Analysis

In this coating pan simulation, there is a huge number of contacts, and forces are evaluated at each of these contacts. However, a large number of these contacts are weak, primarily in the small movements of the tablets in the avalanche region. There are a few contacts which are stronger, and it is assumed that these stronger contacts are the reason for tablet defects. Considering this, the ten contacts with highest force values were analyzed in Figure 4.

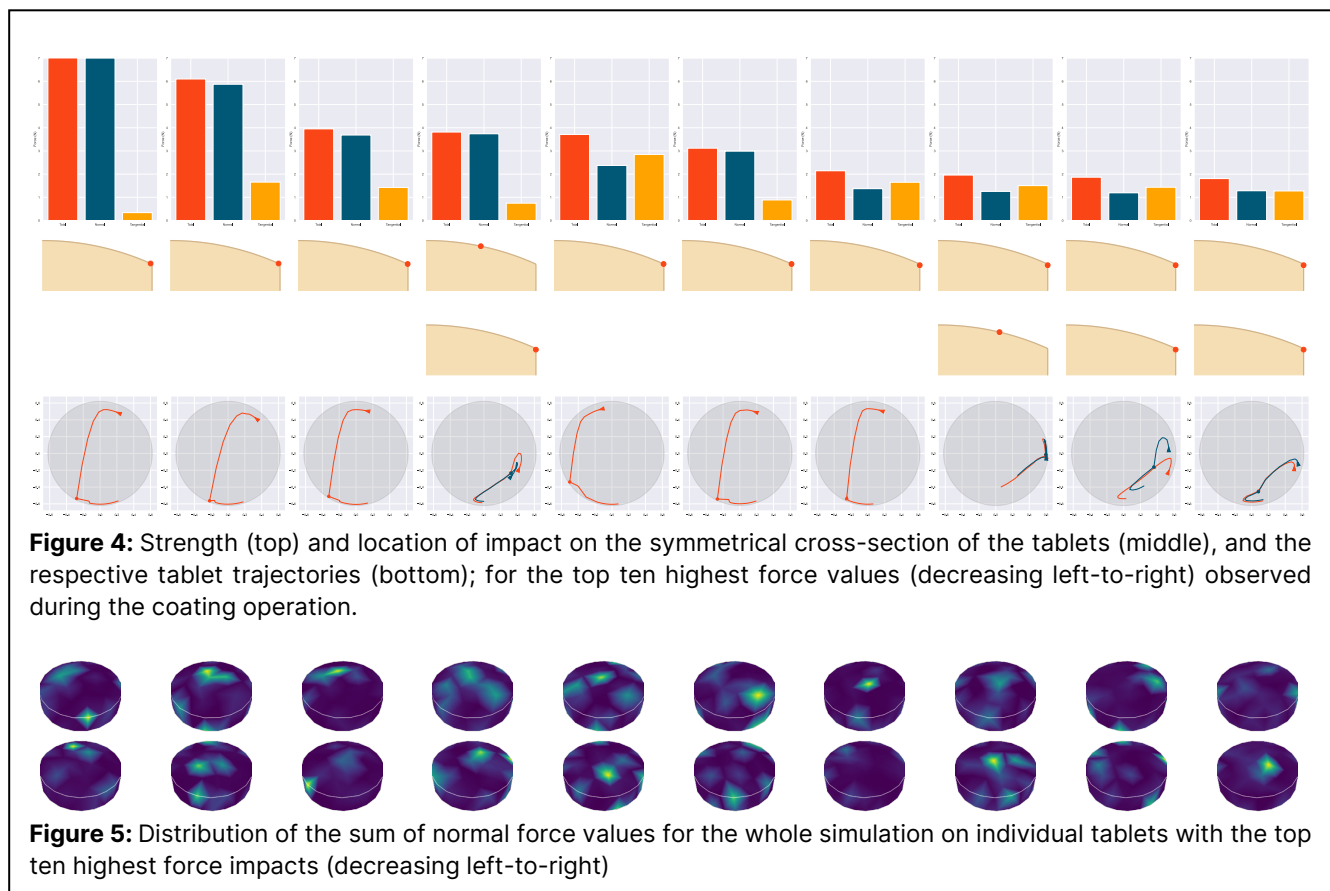
On the top, the orange bar represents the total force magnitude, and the blue and yellow bars represent the normal and tangential force components respectively. This shows the strength of the contact relative to the other contacts in the system. In the first column, since the tangential force is very insignificant, it is almost entirely a normal contact, with an impact on the edge of the tablet. The streamline plot at the bottom also shows that this was caused due to the lifting the tablet by the coater

baffle and dropping it on the coater itself. Figure 4 reveals that although there are some impacts from tablets colliding with other tablets, it is less significant than the tablets impacting the coater pan itself.

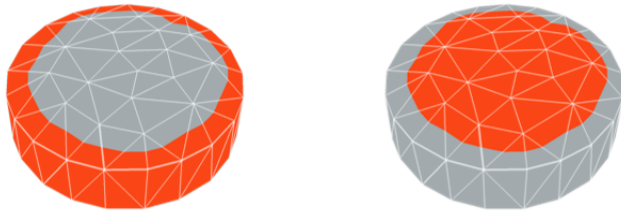
The force values at each vertex were summed for all the contacts in the simulation, and a heat map of the force distribution was plotted in Figure 5. This figure gives a pictorial representation of the locations on the tablet that have the highest propensity for defect formation.

### Grouped Statistics

For practical purposes, the defects were classified into two groups – edge defects and face defects. The force distribution on every individual tablet was used to group them. The force value at each vertex were summed up for each of the 200,000 tablets in the system, for every contact that was experienced during the twenty-five seconds of steady-state operation. Out of these summed up force values at every vertex, if the highest force value on an individual tablet was present near the edge of the tablet, that tablet was classified to incur an edge defect, or chipping defect. Consequently, if the highest force value was on the face of the tablet, it was classified as a face defect, or picking defect. Figure 6 shows the regions on the tablet based on which the groups were classified.



The simulation found that 72.12% of the particles are prone to have a chipping defect or an edge defect, while only 27.88% of the particles are prone to have a face defect. Although the highest force values on an individual contact was from lifting of the coater baffles, the highest sum of forces across all contacts is caused in the avalanche region of the coater.



**Figure 6.** Regions on the tablet classified for edge defects (left) and face defects (right)

## CONCLUSIONS

The simulation estimated ~70/30 split between edge defects and face defects. The end-of-batch statistical inspection for visual defects confirmed that physical damage to the tablet edges constituted a far greater percentage of tablet defects than those observed on the face of the tablets. Electronic inspection of the batches also confirmed the simulation results, again showing that the majority of the physically damaged tablets experienced edge damage, with little physical damage being observed on the face of the tablets. During process scale up, when coating conditions fell on the wetter side of the thermodynamic design space, picking defects were observed mainly on the face of the tablets, Figure 2. This would not be unexpected due to the number of tablet-to-tablet interactions that occur during the coating process while the tablets are avalanching in the spray zone of the coating pan. The result of this DEM analysis was used by the manufacturing and development teams to investigate potential modifications. The tablet core robustness was modified to prevent damage to the tablet edges. To prevent picking damage, the coating operation was moved towards the drier thermodynamic space, since reducing the number of tablet-to-tablet interactions would have required altering the production batch size, which was not an option.

The use of the Discrete Element Method to look at the forces that the tablets were experiencing during the coating operation was one part of an investigation that was critically reviewing the drug product formulation and process to establish causal factors associated with tablet edge and face defects. The evaluation looked at the drug production formulation, raw material attribute and variability within a given attribute for the excipients and drug substance. The processing evaluation included the production of the tablet cores, the tablet core robustness to

downstream processing conditions and environment, review of the coating process and the thermodynamic space for which processing was being performed. Additionally, the evaluation also included identifying the location within processing time of the coating operation that the tablet damage was occurring, i.e. beginning middle or end of the coating operation. Since the edge and face defects did not result in exposed cores, the investigation concentrated to the loading of core tablet to the coating pan, the start up of the coating process though the middle of the coating operation. No review of the post coating operation was evaluated for the investigation for causal factors again supported by the defects present did not show any signs of exposed cores.

This present study can also inform population balance models to predict breakage of the tablet cores. This DEM model can provide micromechanical data which can be used in conjunction with a breakage criterion resulting in an analytical breakage model. An approach similar to Ketterhagen et al [11] can be taken to utilize the material strength parameter and the impact energy to characterize a stochastic breakage model using the Vogel and Peukert [12][13] model to quantitatively predict breakage.

## REFERENCES

1. Zaid AN. A Comprehensive Review on Pharmaceutical Film Coating: Past, Present, and Future. *Drug Design, Development and Therapy* 14: 4613-4623 (2020)
2. Cundall PA, Strack ODL. A discrete numerical model for granular assemblies. *Géotechnique* 29 (1): 47-65 (1979)
3. Zhu HP, Zhou ZY, Yang RY, Yu AB. Discrete particle simulation of particulate systems: a review of major applications and findings. *Chemical Engineering Science* 63 (23): 5728-5770 (2008)
4. Yeom SB, Ha ES, Kim MS, Jeong SH, Hwang SJ, Choi DH. Application of the discrete element method for manufacturing process simulation in the pharmaceutical industry. *Pharmaceutics* 11 (8): 414 (2019)
5. Boehling P, Toschkoff G, Just S, Knop K, Kleinebudde P, Funke A, Rehbaum H, Rajniak P, Khinast JG. Simulation of a tablet coating process at different scales using DEM. *European Journal of Pharmaceutical Sciences* 93: 74-83 (2016)
6. Madlmeir S, Forgber T, Trogrlic M, Jajcevic D, Kape A, Contreras L, Carmody A, Liu P, Davies C, Sarkar A, Khinast JG. Modeling the coating layer thickness in a pharmaceutical coating process. *European Journal of Pharmaceutical Sciences* 161: 105770 (2021)
7. Chen J, Yang Q, Liu J, Jin M, He S, Zhou X, Zhou H,

- Dong J, Yang G, Zhu J. Understanding the correlations between tablet flow dynamics and coating uniformity in a pan coater. *Chemical Engineering Journal* 471: 144392 (2023)
8. Thornton C, Ciocomos MT, Adams MJ. Numerical simulations of agglomerate impact breakage. *Powder Technology* 105 (1-3): 74-82 (1999)
  9. Moreno R, Ghadiri M, Antony SJ. Effect of the impact angle on the breakage of agglomerates: a numerical study using DEM. *Powder Technology* 130 (1-3): 132-137 (2003)
  10. Bharadwaj R, Ketterhagen WR, Hancock BC. Characterization of tablet attritional forces in a friability tester and film coating pan using DEM. *AIChE Annual Meeting* (2010)
  11. Ketterhagen WR, Larson J, Spence K, Baird JA. Predictive approach to understand and eliminate tablet breakage during film coating. *AAPS PharmSciTech* 22(5): 178 (2021)
  12. Vogel L, Peukert W. Breakage behavior of different materials – construction of a mastercurve for the breakage probability. *Powder Technology* 129 (1-3): 101-110 (2003)
  13. Vogel L, Peukert W. From single particle impact behavior to modelling of impact mills. *Chemical Engineering Science* 60 (18): 5164-5176 (2005)
  14. Lu G, Third JR, Müller CR. Discrete element models for non-spherical particle systems: From theoretical developments to applications. *Chemical Engineering Science* 127: 425-465 (2015)
  15. American Pharmacists Association. Tablet Design. In: Tableting Specification Manual. 7th Edition. American Pharmacists Association (2006)
  16. Hertz H. On the contact of elastic solids. *Journal für die Reine und Angewandte Mathematik* 92: 156-171 (1882)
  17. Nassauer B, Kuna M. Contact forces of polyhedral particles in discrete element method. *Granular Matter* 15: 349-355 (2013)
  18. Altair® EDEM™.  
[https://help.altair.com/edem/Creator/Physics/Spinning\\_Friction/Spinning.htm](https://help.altair.com/edem/Creator/Physics/Spinning_Friction/Spinning.htm)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Reinforcement Learning-Driven Process Design: A Hydrodealkylation Example

Yuhe Tian<sup>a\*</sup>, Ayooluwa Akintola<sup>a</sup>, Yazhou Jiang<sup>a</sup>, Dewei Wang<sup>b</sup>, Jie Bao<sup>b</sup>, Miguel A. Zamarripa<sup>c</sup>, Brandon Paul<sup>c</sup>, Yunxiang Chen<sup>b</sup>, Peiyuan Gao<sup>b</sup>, Alexander Noring<sup>c</sup>, Arun Iyengar<sup>c</sup>, Andrew Liu<sup>b</sup>, Olga Marina<sup>b</sup>, Brian Koepfel<sup>b</sup>, Zhijie Xu<sup>b</sup>

<sup>a</sup> West Virginia University, Department of Chemical and Biomedical Engineering, Morgantown, WV, US

<sup>b</sup> Pacific Northwest National Laboratory, US

<sup>c</sup> KeyLogic Systems LLC, Morgantown, WV, US

\* Corresponding Author: [yuhe.tian@mail.wvu.edu](mailto:yuhe.tian@mail.wvu.edu).

---

## ABSTRACT

In this work, we present a follow-up work of reinforcement learning (RL)-driven process design using the Institute for Design of Advanced Energy Systems Process Systems Engineering (IDAES-PSE) Framework. Herein, process designs are generated as stream inlet-outlet matrices and optimized using the IDAES platform, the objective function value of which is the reward to RL agent. Deep Q-Network is employed as the RL agent including a series of convolutional neural network layers and fully connected layers to compute the actions of adding or removing any stream connections, thus creating a new process design. The process design is then informed back to the RL agent to refine its learning. The iteration continues until the maximum number of steps is reached with feasible process designs generated. To further expedite the RL search of the design space which can comprise the selection of any candidate unit(s) with arbitrary stream connections, we investigate the role of RL reward function and their impacts on exploring more complicated versus intensified process configurations. A sub-space search strategy is also developed to branch the combinatorial design space to accelerate the discovery of feasible process design solutions particularly when a large pool of candidate process units is selected by the user. The potential of the enhanced RL-assisted process design strategy is showcased via a hydrodealkylation example.

**Keywords:** Process Design, Process Synthesis, Machine Learning, Reinforcement Learning, Optimization

## 1. INTRODUCTION

Process synthesis aims to determine the optimal selection of unit operations and their flowsheet interconnections at the optimal operating conditions [1]. However, it is not a trivial task to select the optimal process design considering the plethora of plausible unit operations and flowsheet configurations which have been investigated in chemical process industry.

Optimization-based approaches [2-3] offer a systematic strategy utilizing mathematical programming to synthesize the overall flowsheet based on a superstructure which contains all the possible flowsheet structural alternatives of interest. The representation of chemical processes also plays a key role to ensure fit-for-purpose design creativity and sufficient modeling accuracy with

tractable computational formulations (e.g., state-task-network [4], unit-port-conditioning streams representation [5], phenomena-based representation [6-7]). Although the efficacy of optimization-based process synthesis has been well demonstrated in many chemical and energy systems, major challenges remain on: (i) the request of a user-specified superstructure which heavily relies on engineering expertise to ensure the quality of optimal design solutions, and (ii) algorithmic complexity to solve the resulting large-scale mathematical optimization problems, etc.

To address these challenges, recent research efforts have been made to drive process synthesis using reinforcement learning (RL) in place of optimization algorithms [8-9]. RL-driven process design typically starts from a maximum pool of unit operations without

requesting any superstructure pre-specification. The intelligent RL agent will select among the available unit operations and generate arbitrary process design structures. The objective function value of the derived process design serves as the reward to keep training RL agent to generate better process designs. Despite the advantages, key open research questions lie in the search efficiency and algorithmic scalability as the number of available unit operations increases.

In this paper, we present a follow-up work to the RL-driven process design approach introduced by Wang et al. [8] integrated with the Institute for Design of Advanced Energy Systems Process Systems Engineering (IDAES-PSE) Framework [10]. To enhance the efficiency of RL-driven design, we will discuss the role of reward functions and the implementation of a sub-space branch and search strategy. Section 2 provides a brief overview of the methodology workflow and highlights the interactions between IDAES-PSE and reinforcement learning. Section 3 introduces the hydrodealkylation case study and the process design analyses. Section 4 presents concluding remarks and ongoing work.

## 2. RL-DRIVEN PROCESS DESIGN WITH IDAES-PSE FRAMEWORK

### 2.1. Overview of the Methodology

The methodology workflow is summarized in Fig. 1 with stepwise procedure detailed in what follows [8].

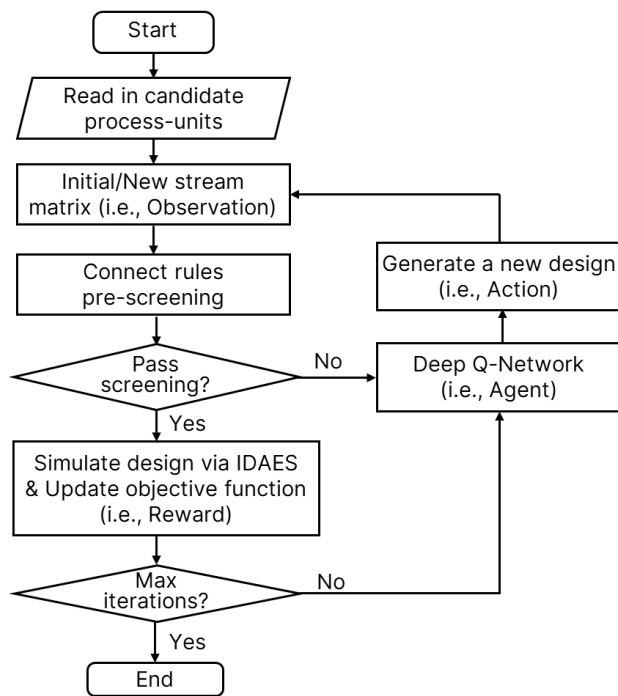


Figure 1. The methodology workflow.

#### 2.1.1. Step 1 – Candidate process-units pool

The design process starts with the users selecting a maximum set of candidate process units which can be used for process design in the next steps. An indicative list of the process unit models, which are supported in the IDAES model library, includes feed (coupled with mixer), product (coupled with flash), flash drum, mixer/splitter, heat/cooler, heat exchanger, stoichiometric reactor, etc. Users are required to specify the types and the maximum number of units available (e.g., maximum of 3 flash drums), while no pre-postulation of flowsheet connections or superstructures is needed.

#### 2.1.2. Step 2 – Flowsheet representation

After defining the candidate process units, an initial process design will be generated via random initializations in the form of inlet-outlet stream matrix. Table. 1 presents an example of the stream matrix representation if the current process flowsheet comprises a heater followed by a stoichiometric reactor. Lists of unit inlets and outlets are first generated to describe the flowsheet connections. The stream matrix then maps the inlet-outlet relationships using 0-1 variables, which serves as the observations to the reinforcement learning algorithm.

Table 1: Stream matrix representation.

	Feed. outlet	Heater. outlet	Reactor. outlet	Flash. outlet
Product.inlet	0	0	1	0
Heater.inlet	1	0	0	0
Reactor.inlet	0	1	0	0
Flash.inlet	0	0	0	0

#### 2.1.3. Step 3 – Connect rules pre-screening

As the flowsheet structure is generated via random initialization (at the first iteration) or via RL (at the successive iterations), it is essential to pre-screen if the structure satisfies general connectivity rules before proceeding with rigorous simulation. Some examples of such rules include:

- At least one unit operation must be selected.
- The outlet of a unit operation cannot be connected with its own inlet.
- Liquid outlet streams cannot connect to compressors or expanders.
- Heaters cannot directly connect to coolers, and compressors cannot directly connect to expanders (turbines).
- There must be a reactor between an outlet product or flash and an inlet feed.

#### 2.1.4. Step 4 – Simulate and optimize process design using IDAES platform

If the structural screening is passed, the flowsheet



is to be simulated and optimized using the IDAES platform with its built-in process-units modeling library, property packages, and mathematical solvers [10]. It is worth noting that the RL reward function is defined as a function of the key process design objectives (e.g., product flow and purity). In this way, the reinforcement learning is continually rewarded to identify better design solutions.

### 2.1.5. Step 5 – RL agent to generate new design alternatives

This step aims to intelligently learn from the observations (step 2) and reward (step 4), thus making a decision on the next action (i.e., generate a new design). Herein, the RL agent adapts the Deep Q-Network (DQN) as illustrated in Fig. 2. DQN utilizes a series of convolutional neural network layers to extract key features from the 2D stream matrices followed by successive fully connected layers to compute the  $Q$  value as per Eq. 1. The  $Q$  value will dictate the value of each action, i.e. adding or removing any unit operation to or from a certain location of the current flowsheet structure. The  $\max(Q_{next})$  term selects the best action in the future  $Q$  values.

$$Q = reward + \gamma \cdot \max(Q_{next}) \quad (1)$$

where *reward* is calculated in step 4 as a function of the objective function,  $\gamma$  is the decay factor,  $Q_{next}$  is the set of future  $Q$  values by taking each possible action.

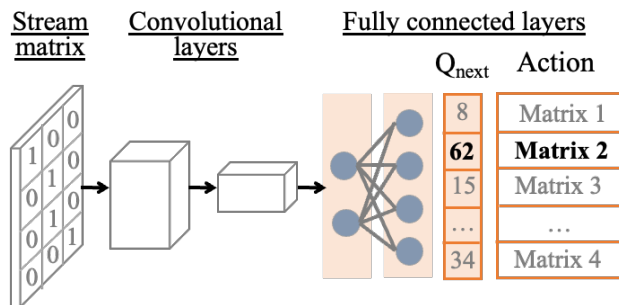


Figure 2. RL agent using Deep Q-Network.

Therefore, a new process design will be generated by the DQN agent which will be sent to Step 3 for connectivity rule pre-screening and IDAES optimization, etc. until a maximum number of iterations is reached. The outcome of this RL-driven process design strategy is a number of feasible process designs with improved quality against the objective function.

This RL-driven process design method can be applied for grassroots chemical process design, as will be showcased in Section 3. It can also significantly contribute to: (i) Retrofitting to improve existing flowsheet designs, and (ii) Integrating new process technologies (e.g., fuel cell, novel intensified reactors) to existing flowsheet infrastructure. In the latter cases, this method may demonstrate more superior computational efficiency as the design space is smaller.

## 2.2. Integration of RL and IDAES

This integrated tool can be accessed via the open-source IDAES platform. The information flow between IDAES and RL interactions is summarized in Fig. 3 from software perspective. An interactive user interface has also been developed as a Jupyter Notebook (i.e., Notebook Interface.ipynb) which allows the users to select candidate process units, adjust RL algorithm parameters (e.g., learning rate), execute RL-driven design, and report the results. The RL algorithm provides new process designs to the IDAES platform for optimization, and the IDAES platform supplies RL with the rewards based on optimization results. The flowsheets are represented respectively using unit list arrays in IDAES and stream matrices in RL, the conversion of which is performed by the obs2list file to enable mutual communication.

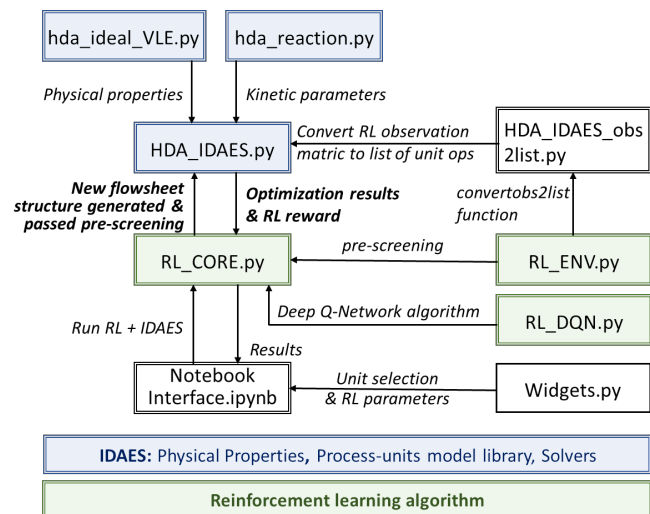


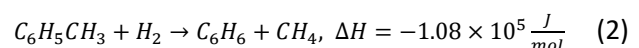
Figure 3. Integrated IDAES and RL information flow.

## 3. CASE STUDY: HYDRODEALKYLATION PROCESS DESIGN

In this section, we apply the afore-introduced methodology to design a hydrodealkylation (HDA) process. Based on the scenario analyses, we discuss several implementations to further improve the RL-driven design efficacy by refining reward function and performing sub-space branching.

### 3.1. Problem Statement

An HDA process is to be designed to produce benzene ( $C_6H_6$ ) from toluene ( $C_6H_5CH_3$ ) and hydrogen ( $H_2$ ) at elevated temperatures. The reaction scheme is shown in Eq. 2, which takes place in the vapor phase. The undesired side reactions are not considered in this case study (e.g., diphenyl production). Ideal vapor-liquid equilibrium can be used to describe the mixture separation behavior.



The feed streams are given in Table 2, consisting of a vapor stream of hydrogen and methane and a liquid stream of toluene. Available process units include feed (coupled with mixer), product (coupled with flash operation), heater, cooler, stoichiometric reactor, mixer, flash, splitter, and compressor. The operating constraints of each type of units are also given, such as heat outlet temperature between 500–600 K. Product specification is set to obtain a vapor product with benzene purity higher than 0.55. The design objective is to determine process solution(s) with optimal vapor product benzene flowrate.

**Table 2:** HDA feed conditions.

	Vapor feed	Liquid feed
Temperature (K)	303.2	303.2
Pressure (kPa)	350	350
Flowrate (mol/s)		
Benzene	0	0
Toluene	0	0.30
Hydrogen	0.30	0
Methane	0.02	0

### 3.2. Scenario Analysis

We first investigate a series of scenarios with different candidate process units selected by user, i.e. different maximum unit sets that can be used for HDA design.

#### 3.2.1. Scenario 1

In this scenario, at most 1 heater and at most 1 reactor can be used to design the HDA process (Table 3). Feed and product are regarded as defaults to be included in process design. Applying the RL-driven approach, the first design solution obtained is shown in Fig. 4. This design solution is referred as Design 1-1, which represents Scenario 1, Solution 1. It is identified at the 63<sup>rd</sup> episode.

The flowsheet comprises a heater with an outlet temperature of 536.9 K followed by a reactor with an outlet temperature of 897.8 K. The flash coupled with outlet product only produces a vapor product stream at 705.7K around 1 atm. The objective value is 0.225 mol/s benzene flowrate, with benzene purity at 0.75 and revenue at  $\$7.65 \times 10^5$ /year.

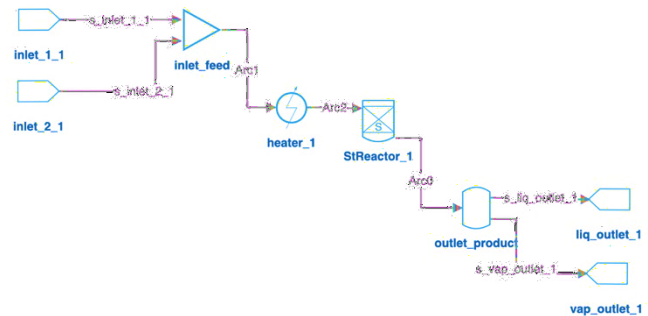
To run a total of 300,500 episodes for this scenario, the total computational time is 2.9 minutes. A substantial portion of the time is spent on RL learning to generate process structures which can pass the pre-screening rules. Only around 4 seconds are consumed by IDAES simulation and optimization. From the aspect of conceptual design, Design 1-1 is actually the best solution of all plausible designs with the maximum benzene flowrate and revenue as well as a most simplified flowsheet.

**Table 3:** User specified candidate units – Scenario 1.

Heater	Cooler	Reactor	Mixer
1	0	1	0
Flash	Splitter	Compressor	
0	0	0	

**Table 4:** User specified candidate units – Scenario 2.

Heater	Cooler	Reactor	Mixer
1	1	1	1
Flash	Splitter	Compressor	
1	1	1	



**Figure 4:** Design 1-1 (and Design 3-1).

#### 3.2.2. Scenario 2

The pool of candidate process units is increased in this scenario, allowing the use of at most 1 unit of each type as defined in Table 4. The first flowsheet design solution (Design 2-1, representing Design Scenario 2, Solution 1) is identified at the 151,507<sup>th</sup> episode, as depicted in Fig. 5. The objective value is 0.180 mol/s benzene flowrate, with benzene purity at 0.75 and revenue at  $\$2.49 \times 10^5$ /year. The inlet feed streams are first heated to 537.6K before entering the reactor. The reactor effluent is cooled down from 898.4K to 325.0K giving a liquid-vapor mixture. The outlet product flash operation then reheats the product temperature to 595 K and reduces the pressure to around 1 atm. A total of 0.18 mol/s benzene flowrate is obtained from the outlet product vapor stream, while no liquid stream is obtained. Another 0.045 mol/s benzene flowrate exists in the outlet exhaust (or purge) stream. While this flowsheet is feasible, it is not a good design. The cooler decreases the energy efficiency and results in the loss of benzene product in the vapor phase via the outlet exhaust.

Another flowsheet design solution (Design 2-2) is generated at the 284,253<sup>rd</sup> episode, as shown in Fig. 6. The objective value is 0.225 mol/s benzene flowrate, with benzene purity at 0.75 and revenue at  $\$5.08 \times 10^5$ /year. To run a total of 300,500 episodes for this scenario, the total computational time is 12.0 minutes.

Several observations can be made for Design 2-2: (i) Similar to Design 2-1, the cooler decreases the energy efficiency, (ii) Compressor is redundant as pressure change is not an essential requirement, (iii) Despite the

fact that the splitter is not completely connected to mixer (one of the pre-screening constraints), the loop of mixer, compressor, and splitter is redundant.

Flash	Splitter	Compressor
2	2	2

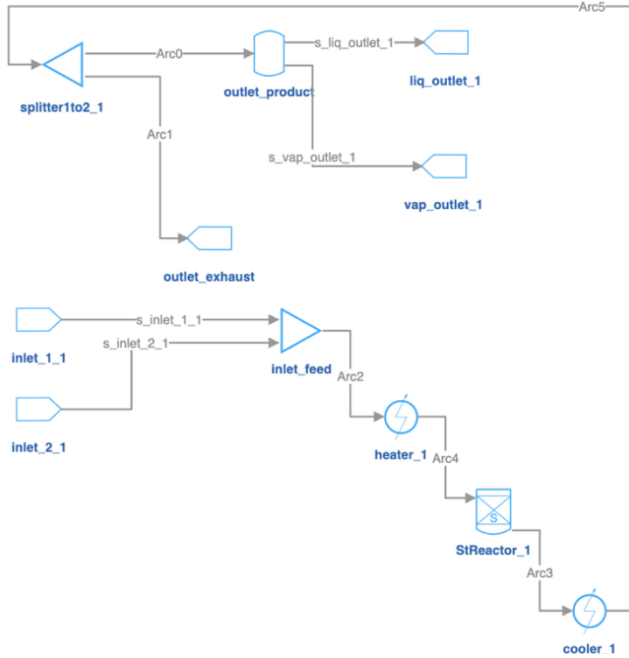


Figure 5: Design 2-1.

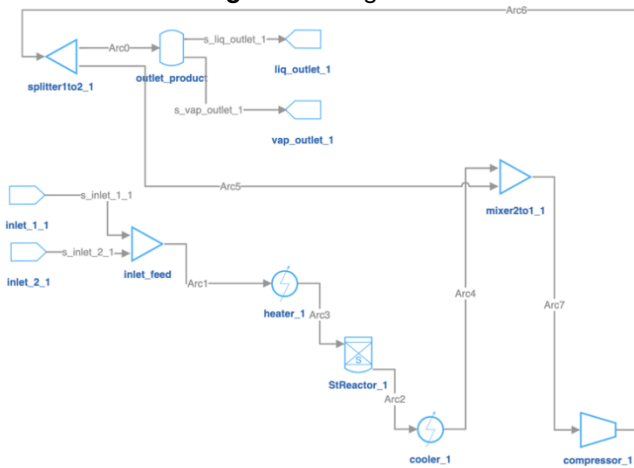


Figure 6: Design 2-2.

### 3.2.3. Scenario 3

The pool of candidate process units is further increased to allow the use of at most 2 units of each type (Table 5). To run 300,500 episodes for this scenario, the total computational time is 24.6 minutes. A flowsheet design solution (Design 3-1) is found at the 788,161<sup>st</sup> episode, which is identical with Design 1-1. No more design solutions are found up to 4,150,000 episodes.

Table 5: User specified candidate units – Scenario 3.

Heater	Cooler	Reactor	Mixer
2	2	2	2

### 3.2.4. Summary of Observations

As the design space increases (i.e., user selects more unit operations to use), the following can be observed based on the above scenario analyses:

- More computational efforts are required to identify any feasible flowsheet design solution. For example, to generate the first feasible design, Scenario 1 at 63<sup>rd</sup> episode, Scenario 2 at 151,507<sup>th</sup> episode, Scenario 3 at 788,161<sup>st</sup> episode.
- The RL algorithm tends to explore the use of more unit operations, despite the possible use of redundant units and/or stream loops (e.g., cooler, compressor, recycle loop).
- The solutions generated from larger combinatorial design space may not recover that from smaller design space. For example, Scenario 2 cannot recover Design 1-1, which is the best design solution.

### 3.3. Impact of Reward Function

To drive the RL algorithm automatically toward a more refined design space with a minimum essential number of unit operations, changing the reward function can be of potential help. The current reward function is given in Eq. 1 which accounts for benzene purity and flowrate. This renders the reward values independent of the number of units. Or in other words, RL does not obtain higher rewards by generating more simplified flowsheet designs. In this context, we propose to incorporate the number of unit operations as part of the reward consideration as defined in Eq. 2.

$$Reward = 1000 + \frac{flow-0.10}{0.05} \times \delta_{scoreA} + \frac{purity-0.55}{0.10} \times \delta_{scoreB} \quad (1)$$

$$Reward = 1000 + \frac{flow-0.10}{0.05} \times \delta_{scoreA} + \frac{purity-0.55}{0.10} \times \delta_{scoreB} + 50 \times (max\_number\_units - actual\_number\_units) \quad (2)$$

Testing again on Scenario 2, the modified reward function leads to the identification of a second flowsheet design at the 242,654<sup>th</sup> episode, which is identical to Design 1-1. This is a faster identification compared to the previous 284,253<sup>th</sup> episode using Eq. 1 while featuring a reduced number of unit operations to identify the best design solution. This reward function is also tested on Scenario 3, in which the user selects maximum 2 units of each type. The first feasible flowsheet design is still

reported at the 788,161<sup>st</sup> episode. Compared to the reward function of Eq. 1 which cannot generate a second flowsheet design within 4,150,000 episodes, the modified function of Eq. 2 reports a second design at the

3,069,993<sup>rd</sup> episode. The flowsheet design comprises a heater, a reactor, a splitter, the outlet exhaust, and the outlet product.

The modified reward function has been

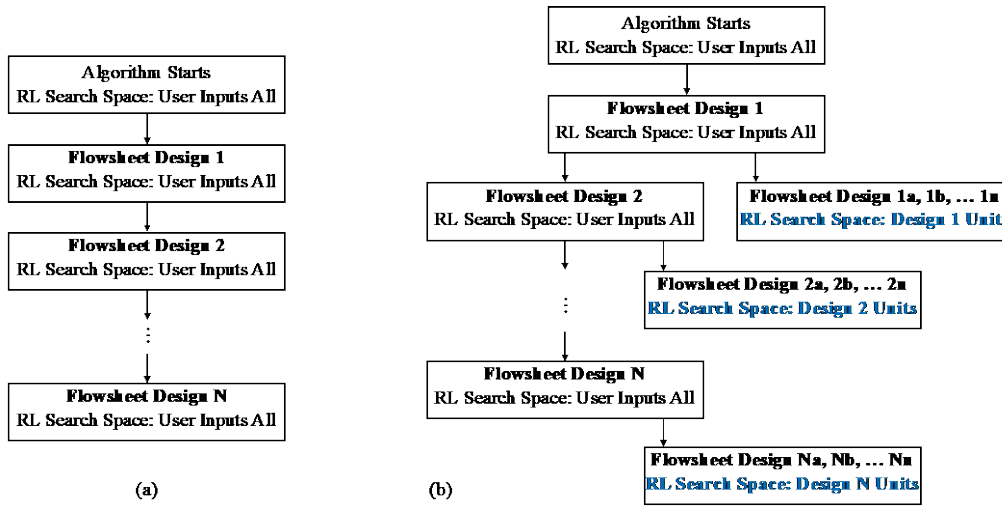


Figure 7: Search strategy for RL-driven design.

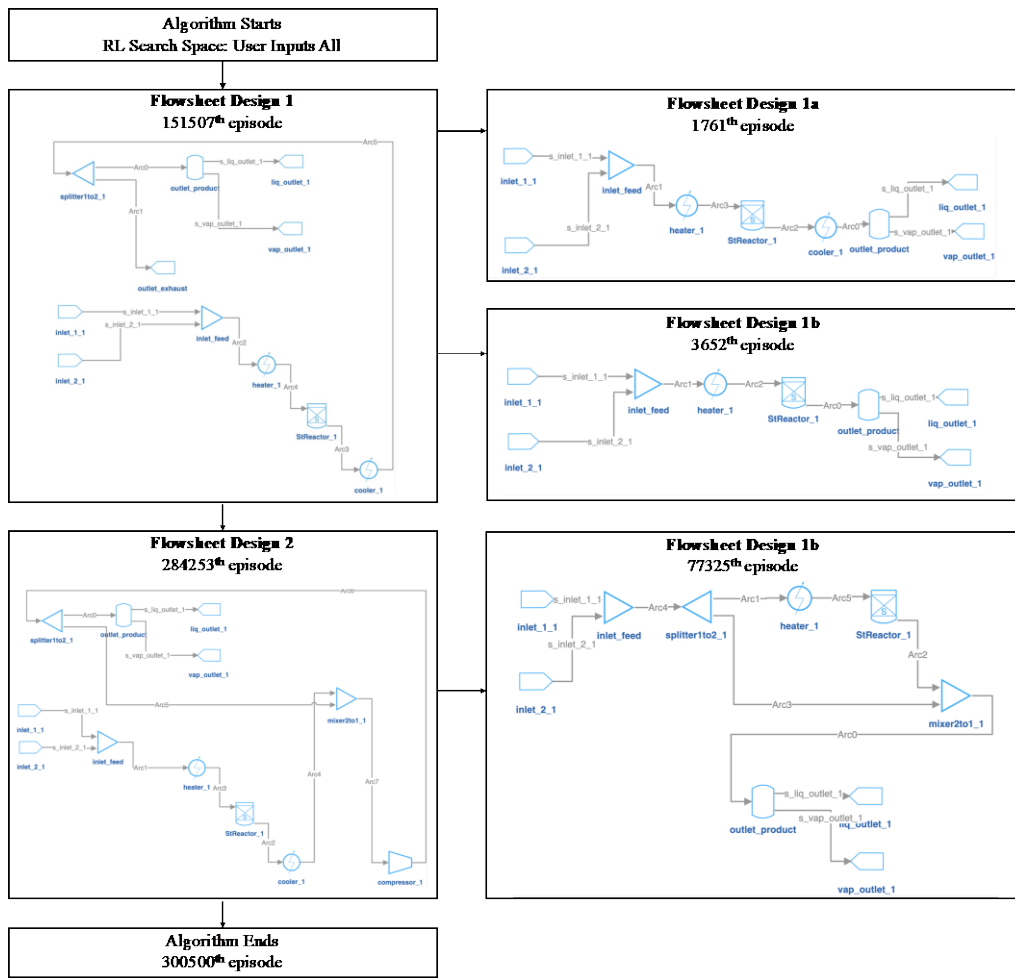


Figure 8: Sub-space search illustration for Scenario 2.

demonstrated to drive the RL search to prioritize the use of fewer unit operations and, thus, to accelerate the search of flowsheet designs after the first design is identified. However, this formulation cannot contribute to accelerate or improve the identification of the first feasible flowsheet design.

### 3.4. Sub-Design Space Branch and Search

As shown in Fig. 7a, the current RL-driven design is set to always explore the entire design space of all candidate process units specified by the user. Arguably, the infeasible flowsheet design space for the arbitrary selections and combinations of these units can be significantly larger than the feasible space. In this way, with a larger number of available unit operations, the RL-driven design may spend the most time learning the infeasible space while directed away from the feasible space. As such, we propose a parallel sub-space search as illustrated in Fig. 7b. Namely, for each of the feasible flowsheet design generated from learning the entire design space, a quick RL search is performed using the list of units in this design as candidate process units (ignoring the specific inlet and outlet connections in this design). Thus, the sub-space search aims to identify, using a reduced pool of unit operations that can construct a feasible flowsheet design, if better design solutions can be generated. This branch and search strategy is implemented for Scenario 2 and the reported results are given in Fig. 8. Better design options are efficiently identified in the sub-space search.

## 4. CONCLUDING REMARKS

In this paper, we have introduced a reinforcement learning-driven process design approach with application to HDA production. Two improvements are implemented to enhance the RL efficiency to search the combinatorial design space, respectively by refining the reward function and branching the search space. Ongoing work is applying superstructure optimization to expedite the sub-design space search. In other words, RL will identify a minimum essential set of unit operations based on which superstructure optimization will screen this sub-design space to find the optimal design solution. The integrated strategy aims to augment RL exploration ability in large combinatorial design space and superstructure optimization rapid screening in smaller design space when mixed-integer nonlinear optimization becomes tractable.

## 5. ACKNOWLEDGEMENTS

Financial support from NETL and ARPA-E DIFFERENTIATE project is gratefully acknowledged.

## REFERENCES

1. Nishida N, Stephanopoulos G, Westerberg AW. A review of process synthesis. *AIChE J.*, 27(3), 321-351 (1981)
2. Mencarelli L, Chen Q, Pagot A, Grossmann IE. A review on superstructure optimization approaches in process system engineering. *Comput. Chem. Eng.*, 136, 106808 (2020)
3. Pistikopoulos EN, Tian Y. Advanced modeling and optimization strategies for process synthesis. *Annu. Rev. Chem. Biomol. Eng.* (2024)
4. Yeomans H, Grossmann IE. A systematic modeling framework of superstructure optimization in process synthesis. *Comput. Chem. Eng.*, 23(6), 709-731 (1999)
5. Wu W, Henao CA, Maravelias CT. A superstructure representation, generation, and modeling framework for chemical process synthesis. *AIChE J.*, 62(9), 3199-3214 (2016)
6. Tian Y, Meduri V, Bindlish R, Pistikopoulos EN. A Process Intensification synthesis framework for the design of dividing wall column systems. *Comput. Chem. Eng.*, 160, 107679 (2022)
7. Demirel SE, Li J, El-Halwagi M, Hasan MF. Sustainable process intensification using building blocks. *ACS Sustain. Chem. Eng.*, 8(48), 17664-17679 (2020)
8. Wang D, Bao J, Zamarripa-Perez M, Paul B, Chen Y, Gao P, Ma T, Noring A, Iyengar A, Schwartz D, Eggleton E, He Q, Liu A, Marina O, Koeppel B, Xu Z. (2023). A coupled reinforcement learning and IDAES process modeling framework for automated conceptual design of energy and chemical systems. *Energy Advances*, 2(10), 1735-1751.
9. Stops L, Leenhouts R, Gao Q, Schweidtmann AM. Flowsheet generation through hierarchical reinforcement learning and graph neural networks. *AIChE J.*, 69(1), e17938 (2023)
10. Lee A, Ghouse JH, Eslick JC, Laird CD, Sirola JD, Zamarripa, MA, Gunter D, Shinn JH, Dowling A, Bhattacharyya D, Biegler LT, Burgard AP, Miller DC. The IDAES process modeling framework and model library – Flexibility for process simulation and optimization. *J. Adv. Manuf. Process.*, 3(3), e10095 (2021)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>







# Part 1

## Peer-Reviewed Articles

Section 4: Design and Energy Transitions



# Equation-Oriented Modeling of Water-Gas Shift Membrane Reactor for Blue Hydrogen Production

Damian T. Agi<sup>a</sup>, Hani A. E. Hawa<sup>b</sup>, and Alexander W. Dowling<sup>a\*</sup>

<sup>a</sup> University of Notre Dame, Department of Chemical and Biomolecular Engineering, South Bend, IN, United States

<sup>b</sup> Precision Combustion, Inc., North Haven, CT, United States

\* Corresponding Author: [adowling@nd.edu](mailto:adowling@nd.edu).

## ABSTRACT

Water-gas shift membrane reactors (WGS-MRs) offer a pathway to affordable blue H<sub>2</sub> generation/purification from gasified feedstock or reformed fuels. To exploit their cost benefits for blue hydrogen production, WGS-MRs' performance needs to be optimized, which includes navigating the multidimensional design space (e.g., temperature, feed pressures, space velocity, membrane permeance and selectivity, catalytic performance). This work describes an equation-oriented modeling framework for WGS-MRs in the Pyomo ecosystem, with an emphasis on model scaling and multi-start initialization strategies to facilitate reliable convergence with nonlinear optimization solvers. We demonstrate, through sensitivity analysis, that our model converges rapidly (< 1 CPU second on a laptop computer) under a wide range of operating parameters (e.g., feed pressures of 1-3 MPa, reactor temperatures of 624-824 K, sweep-to-feed ratios of 0-0.5, and steam/carbon ratios of 1-5). Ongoing work includes (1) validation and calibration of the WGS-MR model using benchtop laboratory data and (2) design, intensification, and optimization of blue H<sub>2</sub> processes using the WGS-MR model.

**Keywords:** Modelling, Membranes, Hydrogen, Process Design, Water-Gas Shift, Model Initialization

## INTRODUCTION

The water-gas shift (WGS) reaction is essential for converting CO into CO<sub>2</sub> and producing additional hydrogen from syngas generated from reforming or gasification [1-2]. WGS is especially critical in blue H<sub>2</sub> production technologies, where the produced CO<sub>2</sub> is captured and sequestered downstream [3].

WGS is an established industrial reaction that has been broadly studied [4]. Ebrahimi et al. [5] provide a comprehensive overview of the WGS reaction, including the CO conversion, H<sub>2</sub> selectivity, and structural properties of transition and noble metal catalysts on oxide and carbon-based supports. The state-of-the-art design for WGS reactors includes a two-staged packed bed reactor: a high-temperature shift (643 K to 673 K) stage followed by inter-stage cooling and then a low-temperature shift (450 K to 553 K) [6]. One prevailing challenge with WGS is that it is a reversible and exothermic reaction ( $\Delta H_{298}^0 = -41$  kJ/mol), which imposes thermodynamic limitations on the attainable conversion.

Incorporating a membrane within a water gas shift reactor enhances H<sub>2</sub> production [2, 6-7]. H<sub>2</sub>-selective membranes (e.g., dense metallic membranes, polymer membranes, and other inorganic membranes) have been previously used for this purpose [7-8]. Particularly, Pd-based membranes are promising for H<sub>2</sub> separation due to their ultra-high H<sub>2</sub>-selectivity [2, 9]. One major challenge with Pd-based membranes is inadequate thermal stability/chemical tolerance in these harsh operating conditions, which can be mitigated by alloying Pd with other elements [2, 10]. Pd-based membranes are used for H<sub>2</sub> production in the form of membrane separators [7, 10] or membrane reactors, i.e., membrane reformers or water-gas shift membrane reactors (WGS-MR) [9, 11].

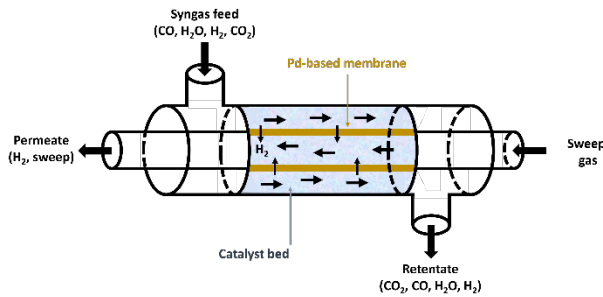
WGS-MRs offer the added benefit of combining separation and reaction in one unit, providing process intensification opportunities. The continuous separation of H<sub>2</sub> from the reactor through Pd-based membranes provides three distinct advantages:

1. The thermodynamic equilibrium continuously shifts in favor of the forward reaction, leading to

- improved CO conversion [2, 12, 13].
- The continuous production of H<sub>2</sub> from the reaction boosts the H<sub>2</sub> partial pressure on the retentate side, enhancing the recovery of H<sub>2</sub> through separation.
  - The retentate is a high-pressure CO<sub>2</sub>-concentrated stream, making carbon capture less energy intensive [14].

Mathematical modeling and numerical simulation are needed to elucidate the design of WGS-MRs for maximizing technical and economic benefits [7, 9, 12]. Membrane modules have been widely modeled in flow sheet simulators such as Aspen Plus to support process-scale optimization and techno-economic analysis [13, 15]. Our approach, based on equation-oriented (EO) design, supports the simultaneous solution of the various model equations, making it easier to embed them directly into large-scale optimization models [16-18]. However, EO models require careful equation and variable scaling and initialization to ensure reliable solver convergence [19]. Additionally, the literature on EO modeling of WGS-MR is notably sparse, with significant contributions from Gosieswki et al. [14].

This brief paper introduces an EO modeling framework of WGS-MRs. We focus on model scaling analysis and multi-start initialization strategies to promote fast and reliable convergence in the solution of the nonlinear model. Finally, we show the model's capabilities for the rapid study of the WGS-MR system through sensitivity analysis.



**Figure 1.** Schematic of a WGS-MR module. The syngas feed flows on the shell side, which is packed with a catalyst for the WGS reaction; the tube is made of a Pd-based membrane that is selectively permeable to H<sub>2</sub>; permeated H<sub>2</sub> and an optional sweep gas flow on the tube side in a counter-current flow direction. Based on similar schematics in Brunetti et al. [9].

## METHODS

### WGS-MR model

We consider a cylindrical, tube-in-shell reactor module with the reaction occurring on the shell side in a packed catalyst bed, as shown in **Figure 1**. The tube

comprises a Pd-based membrane that selectively supports H<sub>2</sub> permeation, which flows on the tube side in a counter-current direction.

### Mass Balance:

The retentate side mass balance for this reactor configuration is given as:

$$\frac{dF_{ret,i}}{dz} = -J_i \frac{A_{mem}}{l} + v_i(-r_{CO}) \frac{V_{rxn}}{l}. \quad (1)$$

Similarly, the permeate side mass balance is given as:

$$\frac{dF_{perm,i}}{dz} = -J_i \frac{A_{mem}}{l}, \quad (2)$$

where  $F_{ret,i}$  (mol s<sup>-1</sup>) and  $F_{perm,i}$  (mol s<sup>-1</sup>) are the retentate and permeate side flowrates of species  $i$  respectively;  $z$  (m) is the axial distance along the WGS-MR module;  $A_{mem}$  (m<sup>2</sup>) is the membrane area;  $l$  (m) is the total length of the module;  $v_i$  (unitless) is the stoichiometric coefficient of species  $i$  in the WGS reaction;  $r_{CO}$  (mol m<sup>-3</sup> s<sup>-1</sup>) is the reaction rate of CO in the WGS reaction;  $V_{rxn}$  (m<sup>3</sup>) is the retentate side volume of the WGS-MR available for the reaction.  $J_i$  (mol m<sup>-2</sup> s<sup>-1</sup>) is the flux of species  $i$  through the Pd-based membrane.

The flux,  $J_i$  is defined by the Sievert-type expression [9, 20]:

$$J_i = \begin{cases} Q(T) [(P_{ret} y_{ret,i} + \varepsilon)^n - (P_{perm} y_{perm,i} + \varepsilon)^n], & i = H_2 \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

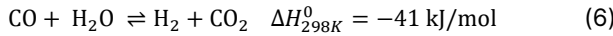
Here,  $Q(T)$  (mol m<sup>-2</sup> Pa<sup>-0.5</sup> s<sup>-1</sup>) is the permeance of the membrane, which correlates with the reactor temperature as given in **Eq. (4)**.  $P_{ret}$  (Pa) and  $P_{perm}$  (Pa) are the retentate and permeate side pressures, respectively.  $y_{ret,i}$  and  $y_{perm,i}$  are the retentate and permeate side compositions of gas species  $i$ , respectively, as defined in **Eq. (5)**. The Sievert's law pressure exponent,  $n$ , ranges from 0 to 1 and takes a value of 0.5 for the ideal Sievert behavior where the diffusion of H atoms in the bulk Pd metal forms the rate-limiting step for H<sub>2</sub> permeation in the Pd-based membrane, and the Pd-H system is infinitely diluted [20].  $\varepsilon$  is a small number (e.g., 10<sup>-8</sup>) that prevents computing the  $n < 1$  exponent of near zero when the partial pressure is very small.

$$Q(T) = Q_0 \exp\left(-\frac{E_a}{RT}\right), \quad (4)$$

$$y_{ret,i} = \frac{F_{ret,i}}{\sum_j F_{ret,j}}, \quad y_{perm,i} = \frac{F_{perm,i}}{\sum_j F_{perm,j}}. \quad (5)$$

In **Eq. (4)**,  $Q_0$  (mol m<sup>-2</sup> Pa<sup>-0.5</sup> s<sup>-1</sup>) is the pre-exponential factor;  $E_a$  (J mol<sup>-1</sup>) is the activation energy of permeation; and  $R$  is the universal gas constant (8.314 J mol<sup>-1</sup> K<sup>-1</sup>).

WGS is a reversible exothermic reaction given in **Eq. (6)** [9]. In this work, we use the reaction rate expression, **Eq. (7)**, proposed by Amadeo and Laborde [21] for WGS catalyzed by copper/zinc oxide/alumina.



$$-r_{\text{CO}} = \frac{0.92 e^{\left(\frac{-454.3}{T}\right)} P_{\text{CO}} P_{\text{H}_2\text{O}} \left(1 - \frac{P_{\text{CO}_2} P_{\text{H}_2}}{P_{\text{CO}} P_{\text{H}_2\text{O}} K_{\text{eq}}}\right) \left(\frac{1}{K_{\text{eq}}}\right)}{\left(1 + 2.2 e^{\left(\frac{101.5}{T}\right)} P_{\text{CO}} + 0.4 e^{\left(\frac{158.3}{T}\right)} P_{\text{H}_2\text{O}} + 0.0047 e^{\left(\frac{2737.9}{T}\right)} P_{\text{CO}_2} + 0.05 e^{\left(\frac{1596.1}{T}\right)} P_{\text{H}_2}\right)^2} \times 16.6667 \rho_{\text{catalyst}} \quad (7)$$

where  $P_{\text{CO}}$ ,  $P_{\text{H}_2\text{O}}$ ,  $P_{\text{CO}_2}$ , and  $P_{\text{H}_2}$  denote the partial pressures in Pa of gas species, CO, H<sub>2</sub>O, CO<sub>2</sub>, and H<sub>2</sub>, respectively,  $\rho_{\text{catalyst}}$  (kg m<sup>-3</sup>) is the density of the WGS catalyst, and  $T$  (K) is the reactor temperature. The 16.6667 multiplier enforces unit conversion from mol g<sup>-1</sup> min<sup>-1</sup> to mol kg<sup>-1</sup> s<sup>-1</sup> units.  $K_{\text{eq}}$  is the equilibrium constant, given by [22]:

$$K_{\text{eq}} = 1.2 \times 10^{-2} e^{\left(\frac{4639}{T}\right)}. \quad (8)$$

### Momentum Balance:

Assuming constant pressure drop on both sides of the Pd-based membrane, the momentum balances in the WGS-MR are as follows:

$$\frac{dP_{\text{ret}}}{dz} = \text{constant}, \quad \frac{dP_{\text{perm}}}{dz} = \text{constant}. \quad (9)$$

A total pressure drop of 35 kPa was used on the retentate side and 0 kPa on the permeate side.

The boundary conditions for the WGS-MR module are:

$$z = 0: F_{\text{ret},i} = y_{\text{feed},i} F_0, \quad P_{\text{ret}} = P_{\text{feed}}, \quad (10)$$

$$z = l: F_{\text{perm},i} = y_{\text{sweep},i} F_{\text{sweep}}, \quad P_{\text{perm}} = P_{\text{sweep}}, \quad (11)$$

where  $F_0$  (mol s<sup>-1</sup>) is the total feed flowrate and  $F_{\text{sweep}}$  (mol s<sup>-1</sup>) is the total sweep flowrate.

### Model scaling

The material balances in **Eqs. (1) & (2)** result in poor solver convergence when the component flowrates,  $J_i$ , are near minimal values ( $\leq 10^{-4}$ ). To circumvent this, we apply dimensionless analysis to scale the flowrates and axial distance using the characteristic parameters  $F_0$  and  $l$ :

$$\bar{F} = \frac{F}{F_0} \leftrightarrow dF = F_0 d\bar{F} \quad (12)$$

$$\bar{z} = \frac{z}{l} \leftrightarrow dz = l d\bar{z}. \quad (13)$$

The material balances in **Eqs. (1) & (2)** become:

$$\frac{d\bar{F}_{\text{ret},i}}{d\bar{z}} = -J_i \frac{A_{\text{mem}}}{F_0} + v_i (-r_{\text{CO}}) \frac{V_{\text{rxn}}}{F_0} \quad (14)$$

$$\frac{d\bar{F}_{\text{perm},i}}{d\bar{z}} = -J_i \frac{A_{\text{mem}}}{F_0} \quad (15)$$

The boundary conditions in **Eqs. (10) and (11)** are updated accordingly by dividing the component flowrates ( $y_{\text{feed},i} F_0$  and  $y_{\text{sweep},i} F_{\text{sweep}}$ ) by  $F_0$ .

**Table 1** reports the model parameters used in this study.

**Table 1.** WGS-MR model parameters. These parameters are based on the work of Brunetti et al. [9]. \*Estimated from property data.

Parameter	Value	Description
$A_{\text{mem}}$ (m <sup>2</sup> )	1.57 $\times 10^{-2}$	Membrane area
$V_{\text{rxn}}$ (m <sup>3</sup> )	3.93 $\times 10^{-5}$	Reaction volume
$Q_0$ (mol m <sup>-2</sup> s <sup>-1</sup> Pa <sup>-0.5</sup> )	1.62 $\times 10^{-2}$	Permeance pre-exponential factor
$\frac{E_a}{R}$ (K)	3.10 $\times 10^3$	Activation energy of permeation per gas constant
$\rho_{\text{catalyst}}$ (kg m <sup>-3</sup> )	1.38 $\times 10^3$	Catalyst density
$n$ (unitless)	0.5	H <sub>2</sub> partial pressure exponent
$F_0$ (mol s <sup>-1</sup> )	1.26 $\times 10^{-3}$ *	Feed flowrate
$T$ (K)	553	Reactor temperature
$P_{\text{feed}}$ (MPa)	1	Feed pressure
$P_{\text{sweep}}$ (MPa)	0.1	Sweep-side pressure
$y_{\text{feed,CO}}$ (vol%)	20	Molar composition of gas species in the feed
$y_{\text{feed,H}_2\text{O}}$ (vol%)	20	
$y_{\text{feed,CO}_2}$ (vol%)	10	
$y_{\text{feed,H}_2}$ (vol%)	50	

### Numerical solution and computational environment

The WGS-MR model represented by the differential-algebraic system in **Eqs. (3) to (11) and (14) to (15)** was discretized using 20 finite volumes, resulting in 520 equations with 520 variables. The model was implemented in Pyomo v6.4.0 [23] and solved using Ipopt v3.13.2 [24] with linear solver ma27 [25], distributed as part of the Institute for the Design of Advanced Energy Systems Process Systems Engineering Framework (IDAES PSE) [26]. The model reliably converged in less than 0.1 CPU seconds on a laptop computer running Windows 11 with Intel® Core™ i5-8250U processor and 8GB of RAM.

### Performance criteria

CO conversion is the percentage of CO in the WGS reactor feed that is reacted to produce H<sub>2</sub> and CO<sub>2</sub> as expressed below [9]:

$$\text{CO Conversion} = \left(1 - \frac{\bar{F}_{\text{ret,CO}}^{\text{out}}}{\bar{F}_{\text{ret,CO}}^{\text{in}}}\right) \times 100\%. \quad (16)$$

H<sub>2</sub> recovery is the percentage of available H<sub>2</sub> in the reactor that is recovered in the permeate outlet [9, 14]:

$$\text{H}_2 \text{ Recovery} = \left( \frac{\bar{F}_{\text{perm,H}_2}^{\text{out}}}{\bar{F}_{\text{perm,H}_2}^{\text{out}} + \bar{F}_{\text{ret,H}_2}^{\text{out}}} \right) \times 100\%. \quad (17)$$

## Model initialization procedure

In the sensitivity analysis, we solved the WGS-MR model in three steps using a multi-start initialization strategy to improve solver convergence:

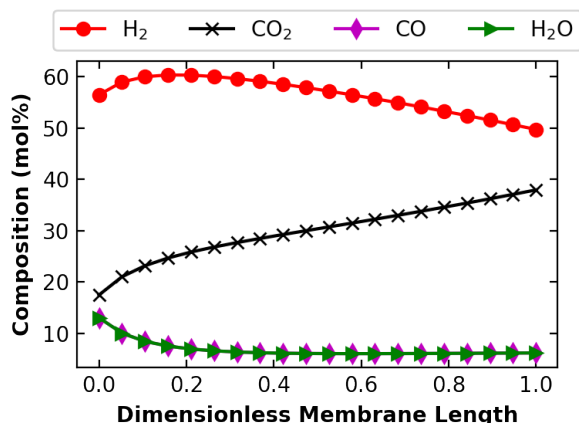
**Step 1:** Toggle off the reaction by fixing the reaction rate variable  $-r_{\text{CO}}$  to zero and deactivating **Eq. (7)**. Then, solve the resulting model. This reduces model complexity by eliminating the nonlinear reaction rate from the material balances in **Eq. (14)**.

**Step 2:** Activate the reaction rate expression in **Eq. (7)**, unfix and initialize the reaction rate,  $-r_{\text{CO}}$ , using partial pressure and temperature values from the solution from the previous step. Solve the entire model.

**Step 3:** Iteratively update the perturbed parameter (e.g., sweep ratio or steam/carbon ratio), defined as a mutable parameter in Pyomo, and re-solve the model for each point in the sensitivity analysis.

## RESULTS

### WGS-MR model reproduces the expected concentration profiles



**Figure 2.** The retentate side concentration profiles of gas species shows that the WGS-MR model captures the relevant physics (i.e., species production/consumption by WGS reaction and  $\text{H}_2$  depletion due to transmembrane permeation).

**Figure 2** shows the retentate side composition of gas species along the dimensionless length of the reactor. As expected, the concentration of the reactants, CO and  $\text{H}_2\text{O}$ , decreases along the reactor length due to their consumption in the WGS reaction. The concentration of  $\text{CO}_2$ , on the other hand, increases along the length of the reactor. The concentration of  $\text{H}_2$ , which is the only gas permeating through the Pd-based membrane, increases and then decreases with a peak near  $\bar{z} = 0.2$ . The

increasing  $\text{H}_2$  concentration corresponds to the reaction-dominated regime, whereas the decreasing  $\text{H}_2$  concentration corresponds to the transport-dominated regime. This concentration profile conforms to other WGS-MR concentration profiles reported in the literature [14, 27].

### Sweep gas increases $\text{H}_2$ production

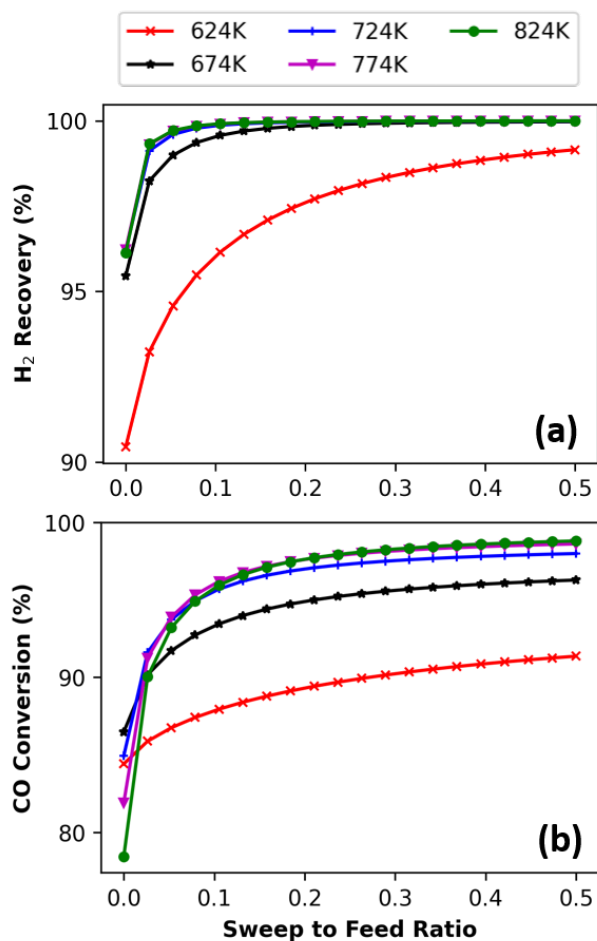
Using an inert sweep gas on the permeate side promotes  $\text{H}_2$  recovery and CO conversion in the WGS-MR by increasing the driving force for  $\text{H}_2$  permeation [9, 12].  $\text{N}_2$  was used as the sweep in this study. The sweep-to-feed ratio (i.e., the ratio of sweep gas flowrate to feed gas flowrate) was systematically varied to investigate the influence of sweep gas on CO conversion and  $\text{H}_2$  recovery in the WGS-MR (**Figure 3**).

**Figure 3(a)  $\text{H}_2$  recovery:** Using a  $\text{N}_2$  sweep-to-feed ratio of 0.1 resulted in a 4% increase in  $\text{H}_2$  recovery in the WGS-MR compared to baseline WGS-MR with no sweep. The sweep gas increases the  $\text{H}_2$  recovery by diluting the  $\text{H}_2$  in the permeate, which lowers its partial pressure and increases the transmembrane partial pressure difference in **Eq. (3)**, which is the driving force for  $\text{H}_2$  flux across the membrane. For reaction temperatures of 724 K and higher, the gain in  $\text{H}_2$  recovery plateaus at a sweep ratio of 0.1 as it approaches the theoretical maximum of 100%. Although sweep gas flow shows the potential for improving  $\text{H}_2$  recovery in the reactor, it also dilutes the recovered  $\text{H}_2$  which may necessitate additional purification based on the application.

**Figure 3(b) CO conversion:** At 724 K, a sweep-to-feed ratio of 0.5 raises CO conversion from 86% to 98%. This improvement in CO conversion is explained by the gain in  $\text{H}_2$  recovery, which translates to increased withdrawal of a reaction product, prompting the reversible WGS reaction to be favored in the forward direction.

Generally, improved performance is observed at higher temperatures. This is the expected behavior for  $\text{H}_2$  recovery because increasing the temperature boosts the  $\text{H}_2$  permeance of the membrane as shown in **Eq. (4)**. However, the increased CO conversion at higher temperatures is intriguing because WGS is an exothermic reaction. We expect lower conversions at higher temperatures. We hypothesize the trends in **Figure 3(b)** are an interplay between temperature effects on the equilibrium constant and  $\text{H}_2$  flux. Specifically, increasing the temperature enhances the flux of  $\text{H}_2$  across the membrane, according to **Eqs. (3)** and **(4)**, shifting the equilibrium to increase CO conversion. Concurrently, the elevated temperature diminishes the equilibrium constant according to **Eq. (8)**, thereby reducing the CO conversion. **Figure 4** shows additional sensitivity analysis for varying either **(a)** reactor temperature or **(b)** membrane temperature while the other is fixed at 624 K. **Figure 4** confirms the increase in CO conversion from the increased  $\text{H}_2$  flux outweighs the temperature shift in the reaction equilibrium.





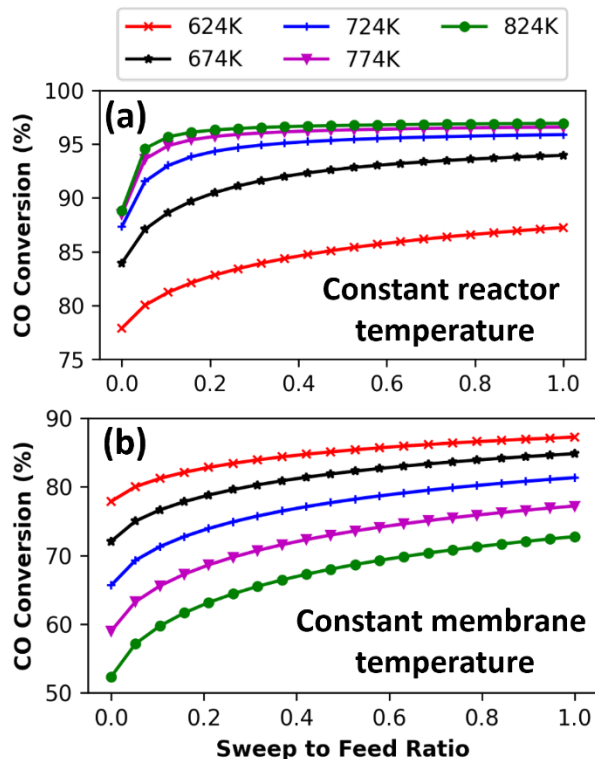
**Figure 3.** Increasing the sweep-to-feed ratio produces considerable gains in (a) H<sub>2</sub> recovery and (b) CO conversion in a WGS-MR.

### CO conversion sensitive to steam/carbon ratio in a WGS-MR

Next, we vary the steam/carbon ratio in the feed (**Figure 5**). The total feed flowrate (and, by extension, the gas hourly space velocity) is fixed as reported in **Table 1**; only the relative proportions of steam and carbon in the feed changed.

**Figure 5** shows that excess steam (greater than the stoichiometric ratio) promotes CO conversion in the WGS-MR. We observe that for a given feed pressure, CO conversion increases monotonically with the steam/carbon ratio up to 98 to 99% conversions at a steam/carbon ratio of 5.0. This observation is consistent with the experimental results of Bang et al. [28] for a Pd-Cu WGS-MR. Increasing the steam/carbon ratio increases reactant concentration, which triggers the forward reaction to nullify the disturbance in the equilibrium state as prescribed by La Chatelier's principle [28]. Increasing the steam/carbon ratio beyond 3.0 yields

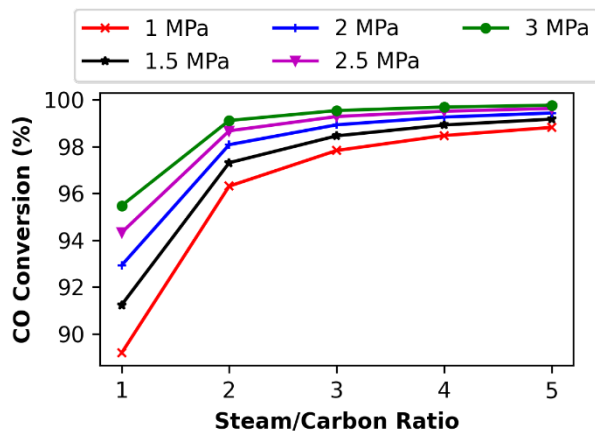
modest improvements in CO conversion as it approaches the theoretical maximum of 100%. The optimal choice of steam/carbon ratio would maximize CO conversion while constraining steam consumption in the reactor.



**Figure 4.** CO conversion as a function of sweep-to-feed ratio in a WGS-MR for two hypothetical conditions: (a) The reactor temperature is fixed at 624 K, and the membrane temperature varies. (b) The membrane temperature is fixed at 624 K and the reactor temperature varies.

### Model convergence with Ipopt solver

The WGS-MR model demonstrates good solver convergence over a wide range of input parameters (e.g., feed pressures of 1 to 3 MPa, reactor temperatures of 624 to 824 K, sweep-to-feed ratios of 0 to 0.5, and steam/carbon ratios of 1 to 5). For instance, the CPU time corresponding to the sensitivity analysis results in **Figure 3** for  $T = 624$  K is 0.26 CPU seconds for **Step 1**, 0.080 CPU seconds for **Step 2**, and 0.010 to 0.64 CPU seconds (average 0.074 CPU seconds) for **Step 3**.



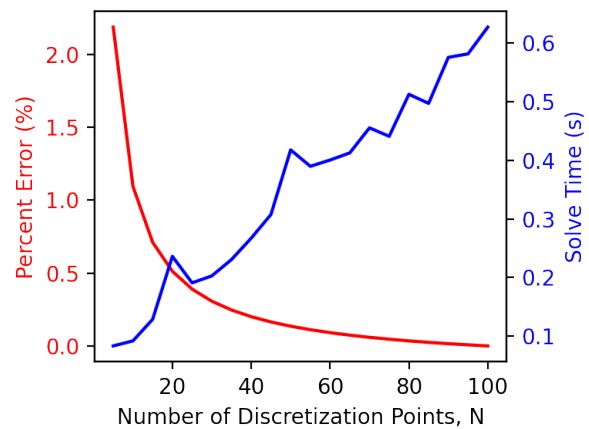
**Figure 5.** CO conversion in the WGS-MR increases with the steam/carbon ratio. Higher conversions are observed at higher feed pressures due to increased H<sub>2</sub> recovery at elevated pressures.

**Table 2** compares the solver convergence performance (i.e., number of iterations and CPU seconds) of the WGS-MR model before and after rescaling the model using the initialization procedure described above. These data correspond to the sensitivity analysis in **Figure 3** and show that the scaled model converges in about 50% of the CPU seconds for the unscaled model, emphasizing the role of proper model scaling in improving numerical performance. A similar trend is observed for other temperatures and sweep-to-feed ratios in **Figure 3**, which are omitted for brevity. We found the scaled model is more robust to a naive initialization, although we recommend the procedure described above.

**Table 2:** Solver convergence of WGS-MR model.

		624 K		824 K	
		scaled	un-scaled	scaled	un-scaled
Sweep ratio = 0.0	# of iterations	4	4	6	7
	CPU secs	0.019	0.036	0.043	0.111
Sweep ratio = 0.2	# of iterations	4	5	5	6
	CPU secs	0.028	0.058	0.034	0.046

**Figure 6** investigates the impact of the number of discretization points on the solve time and percent error. Here, percent error is defined relative to the numeric solution with  $N = 100$  discretization points. As expected, the solve time increases approximately linearly with the number of discretization points, whereas the percent error decays nearly exponentially as  $N$  increases. Based on these results, we conclude that  $N = 20$  is likely sufficient for process design and optimization.



**Figure 6.** Sensitivity analysis quantifies the impact of the number of discretization points on the solve time and percent error. The percent error was calculated for the flowrate of H<sub>2</sub> in the permeate relative to the numerical solution with  $N = 100$ .

## CONCLUSIONS

We present an EO modeling framework for WGS-MR and propose a scaling analysis and multi-start initialization procedure that promotes solver convergence. Through sensitivity studies with this model, we show that sweep gas on the permeate side could drive up conversion to over 95% and H<sub>2</sub> recovery to ~99% for temperatures 674 K and above. We also show that excess steam could promote CO conversions in the WGS-MR up to a steam/carbon ratio of 3.0 for the dataset considered in this study. This EO model converges rapidly (<1 CPU sec) and serves as a tool for the design and optimization of WGS-MRs.

Ongoing work utilizes this model to optimize the techno-economic benefits of WGS-MRs for blue H<sub>2</sub> production from gasified biomass. The EO modeling approach adopted in this work motivates further work into integrating the design of WGS-MR membrane modules into large-scale, EO-based flow sheets for process-wide optimization of H<sub>2</sub> production technologies.

## ACKNOWLEDGEMENTS

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Award Number DE-SC0022409, with prime recipient Precision Combustion, Inc. This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal

liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

D.A. also recognizes partial support from the U.S. National Science Foundation (EEC-164772).

## REFERENCES

- Lewis, E., McNaul, S., Jamieson, M., Henriksen, M. S., Matthews, H. S., Walsh, L., Grove, J., Shultz, T., Skone, T.J. & Stevens, R. *Comparison of commercial, state-of-the-art, fossil-based hydrogen production technologies*. No. DOE/NETL-2022/3241. National Energy Technology Laboratory (NETL), Pittsburgh, PA, Morgantown, WV, and Albany, OR (United States), 2022.
- Liguori, S., Kian, K., Buggy, N., Anzelmo, B. H., & Wilcox, J. Opportunities and challenges of low-carbon hydrogen via metallic membranes. *Progress in Energy and Combustion Science*, 80 (2020).
- National Grid. The hydrogen colour spectrum. URL: <https://www.nationalgrid.com/stories/energy-explained/hydrogen-colour-spectrum> (Accessed November 13, 2023).
- Baraj, E., Ciahotný, K., & Hlinčík, T. The water gas shift reaction: Catalysts and reaction mechanism. *Fuel*, 288 (2021), 119817.
- Ebrahimi, P., Kumar, A., & Khraisheh, M. A review of recent advances in water-gas shift catalysis for hydrogen production. *Emergent Materials*, 3 (2020).
- Iulianelli, A., Pirola, C., Comazzi, A., Galli, F., Manenti, F., & Basile, A. Water gas shift membrane reactors. In *Membrane reactors for energy applications and basic chemical production* (pp. 3-29) (2015). Woodhead Publishing.
- Schwartz, J., Porter, J., Patki, N., Kelley, M., Stanislawski, J., Tolbert, S., Way, J.D. and Makuch, D. *Advanced Hydrogen Transport Membrane for Coal Gasification*. Praxair, Inc., Tonawanda, NY (United States), 2015.
- Liu, C., Zhang, X., Zhai, J., Li, X., Guo, X., & He, G. Research progress and prospects on hydrogen separation membranes. *Clean Energy*, 7(1), (2023).
- Brunetti, A., Caravella, A., Barbieri, G., & Drioli, E. Simulation study of water gas shift reaction in a membrane reactor. *Journal of Membrane Science*, 306(1-2), (2007).
- Okazaki, J., Ikeda, T., Tanaka, D.A.P., Sato, K., Suzuki, T.M. and Mizukami, F. An investigation of thermal stability of thin palladium-silver alloy membranes for high temperature hydrogen separation. *Journal of Membrane Science*, 366(1-2), pp.212-219, (2011).
- Saw, S. Z., Nandong, J., & Ghosh, U. K. Optimization of steady-state and dynamic performances of water-gas shift reaction in membrane reactor. *Chemical Engineering Research and Design*, 134, 36-51, (2018).
- Karagöz, S., da Cruz, F. E., Tsotsis, T. T., & Manousiouthakis, V. I. Multi-scale membrane reactor (MR) modeling and simulation for the water gas shift reaction. *Chemical Engineering and Processing-Process Intensification*, 133, 245-262, (2018).
- Yonamine, W., Thangavel, S., Ohashi, H., & Fushimi, C. Performance analysis of a water-gas shift membrane reactor for integrated coal gasification combined cycle plant. *Energy Conversion and Management*, 174, 552-564, (2018).
- Gosiewski, K., & Tańczyk, M. Applicability of membrane reactor for WGS coal derived gas processing: Simulation-based analysis. *Catalysis today*, 176(1), 373-382, (2011).
- Giuliano, A., Poletto, M., & Barletta, D. Pure hydrogen co-production by membrane technology in an IGCC power plant with carbon capture. *International Journal of Hydrogen Energy*, 43(41), 19279-19292, (2018).
- Dowling, A. W., & Biegler, L. T. A framework for efficient large scale equation-oriented flowsheet optimization. *Computers & Chemical Engineering*, 72, 3-20, (2015).
- Ghosh, K., Vernuccio, S., & Dowling, A. W. Nonlinear reactor design optimization with embedded microkinetic model information. *Frontiers in Chemical Engineering*, 4, (2022).
- Biegler, Lorenz T. "New directions for nonlinear process optimization." *Current Opinion in Chemical Engineering* 21 (2018): 32-40.
- Agi, D.T., Jones, K.D., Watson, M.J., Lynch, H.G., Dougher, M., Chen, X., Carlozo, M.N., Dowling, A.W. Computational toolkits for model-based design and optimization. *Current Opinion in Chemical Engineering*, 43:100994, (2024).
- Caravella, A., Hara, S., Drioli, E., & Barbieri, G. Sieverts law pressure exponent for hydrogen permeation through Pd-based membranes: Coupled influence of non-ideal diffusion and multicomponent external mass

- transfer. *International journal of hydrogen energy*, 38(36), 16229-16244, (2013).
21. Amadeo, N. E., & Laborde, M. A. Hydrogen production from the low-temperature water-gas shift reaction: Kinetics and simulation of the industrial reactor. *International journal of hydrogen energy*, 20(12), 949-956, (1995).
  22. El Hawa, H. W. A., Paglieri, S. N., Morris, C. C., Harale, A., & Way, J. D. Application of a Pd-Ru composite membrane to hydrogen production in a high temperature membrane reactor. *Separation and Purification Technology*, 147, 388-397, (2015).
  23. Bynum, M.L., Hackebeil, G.A., Hart, W.E., Laird, C.D., Nicholson, B.L., Sirola, J.D., Watson, J.P. and Woodruff, D.L. *Pyomo-optimization modeling in python* (Vol. 67). Berlin/Heidelberg, Germany: Springer, 2021.
  24. Wächter, A., & Biegler, L. T. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical programming*, 106, 25-57, (2006).
  25. HSL A Collection of Fortran Codes for Large-Scale Scientific Computation. URL: <https://www.hsl.rl.ac.uk/>. (Accessed November 13, 2023).
  26. Lee, A., Ghouse, J.H., Eslick, J.C., Laird, C.D., Sirola, J.D., Zamarripa, M.A., Gunter, D., Shinn, J.H., Dowling, A.W., Bhattacharyya, D. and Biegler, L.T. The IDAES process modeling framework and model library—Flexibility for process simulation and optimization. *Journal of Advanced Manufacturing and Processing*, 3(3), p.e10095, 2021.
  27. Nekhamkina, O., & Sheintuch, M. Effective approximations for concentration-polarization in Pd-membrane separators. *Chemical Engineering Journal*, 260, 835-845, (2015).
  28. Bang, G., Moon, D. K., Kang, J. H., Han, Y. J., Kim, K. M., & Lee, C. H. (High-purity hydrogen production via a water-gas-shift reaction in a palladium-copper catalytic membrane reactor integrated with pressure swing adsorption. *Chemical Engineering Journal*, 411, 128473, 2021).

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Modeling the Maximization of Waste Heat Use in a Liquid Solvent Direct Air Capture Plant Through Hydrogen Production

Erick O. Arwa<sup>a\*</sup> and Kristen R. Schell<sup>a</sup>

<sup>a</sup> Carleton University, Department of Mechanical and Aerospace Engineering, Ottawa, Ontario, Canada

\* Corresponding Author: [erickarwa@mail.carleton.ca](mailto:erickarwa@mail.carleton.ca).

## ABSTRACT

Direct air capture (DAC) of carbon dioxide is a promising technology to enable climate change mitigation. The liquid solvent DAC (LSDAC) process is one of the leading technologies being piloted. However, LSDAC uses a high-temperature regeneration process which requires a lot of thermal energy. Although current LSDAC designs incorporate pre-heat cyclones and a heat recovery steam generator to enable heat recovery, these do not maximize the use of the heat in the products of calcination. In this paper, a linear optimization model is developed to minimize energy cost in a LSDAC that is powered by renewable energy and natural gas. First, the material flow network is modified to include a heat exchanger (HX) and water supply to a proton exchange membrane (PEM) electrolyser. Mass and energy balance constraints are then developed to include the water flow as well as the energy balance at the PEM and the HX. Results show that about 911 tonnes of hydrogen could be produced over 336 hours of operation using a 136MW PEM. Further analysis reveals that hydrogen production is only prioritized if the value is higher than the cost of natural gas.

**Keywords:** Climate change, Direct air capture, Hydrogen, Negative emission technologies, PEM

## INTRODUCTION

To forestall a possible climate crisis, negative emission technologies (NETs) such as direct air capture (DAC) of  $CO_{2(g)}$  from the atmosphere are considered a necessary addition to other mitigation measures [1]. This is due to its ability to capture the already emitted carbon from the atmosphere independent of the origin of the emission. Also, DAC is likely to reduce the cost of achieving net-zero power grids [2], [3]. DAC has special advantages over other NETs such as traceability, controllability, and modularity. Two crucial DAC technologies are currently being piloted in various parts of the globe: liquid solvent DAC (LSDAC) and solid sorbent DAC. Compared to the solid sorbent technologies, the LSDAC has a slightly lower energy requirement per tonne of  $CO_{2(g)}$ , better scalability and continuity of operation [1]. Further, LSDAC uses mature chemical process technologies such as calcium looping (CL) [4].

Despite the promising features of LSDAC, it faces

several challenges. First, its capture rate is heavily influenced by climate as the liquid solvent may freeze at temperatures below 0 °C [5]. Second, the high-temperature regeneration process is energy intensive requiring a stable power supply necessitating the use of high capacity factor power generation technologies, most of which depend on fossil fuels. DAC energetics is an important research agenda that would enable efficient integration of the technology into the energy system. Among the aspects of improving the energy profile of LSDAC is reducing the energy required for regeneration through electrochemical processes [6]. Another important aspect is making the regeneration process flexible so that it could be powered using renewable energy (RE) sources, which are intermittent by nature [7]. One technique of improving DAC energetics that is missing in the literature is the full utilization of heat recovered from the LSDAC. In [8],  $H_{2(g)}$  is produced using waste heat from an organic Rankine cycle generator. The waste heat is useful in providing the thermal needs of the PEM and raising the temperature of



the inflowing water to the PEM's operating temperature, thereby improving its performance [9], [10].

Keith et al [4] discussed several heat integration techniques that enhanced the use of heat recovered from the calciner and the slaker. In their model, heat is recovered from outgoing  $CO_{2(g)}$  from the calciner to pre-heat the incoming  $CaCO_{3(s)}$  pellets from the pellet reactor to 650 °C through two cyclones. Another cyclone is used to recover heat from  $CaO$  to pre-heat oxygen from the air separation plant (ASP) to 674 °C. Heat is also recovered from these two substances and the steam slaker through a heat recovery steam generator to produce steam which is used to drive a steam turbine to produce electricity for the plant. However, this elaborate heat recovery scheme still leaves a lot of thermal energy wasted. For instance, after all the heat recovery processes,  $CO_{2(g)}$  is still at 325 °C, which has energy that could still power the system if properly integrated. Furthermore, if the system is all electric, there is no need for an ASP, thus freeing the heat from  $CaO$  for reuse in the process.

This paper proposes the use of the recovered heat from the products of calcination to produce  $H_{2(g)}$ . It is an improvement of [7], where a linear programming (LP) model was proposed to lower energy costs in the LSDAC plant by flexibly scheduling regeneration process to maximize the utilization of RE in the plant. First, the process flow is modified to include a heat exchanger (HX) and water supply to a proton exchange membrane (PEM) electrolyser. Additional constraints are then developed to model the additional water flow as well as the energy balance at the PEM and the HX.

## METHODOLOGY

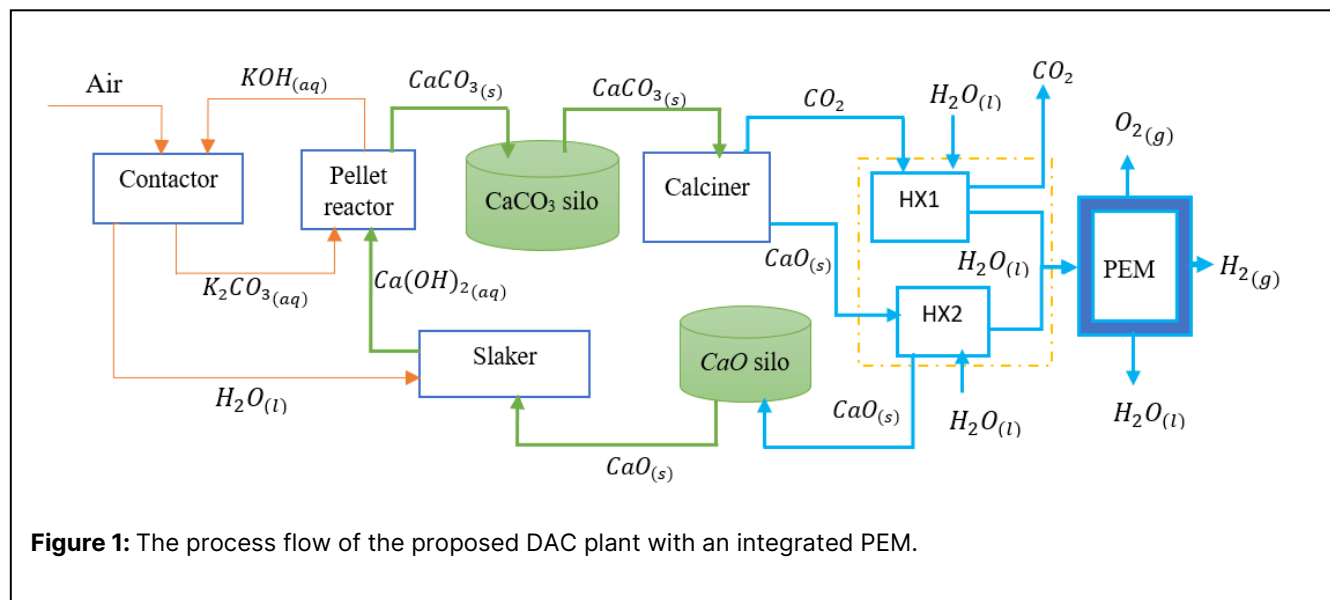
## Process Flow Design of the Proposed DAC Plant

Designed by Keith et al [4], a typical liquid solvent DAC plant has two major loops. The first loop – the potassium cycle – starts at the air contactor where  $KOH_{(aq)}$  reacts with  $CO_{2(g)}$  from the air, producing  $K_2CO_{3(aq)}$ , which is reacted with  $Ca(OH)_{2(aq)}$  in the pellet reactor, thereby regenerating the  $KOH_{(aq)}$  to complete the cycle. The second cycle starts with the thermal decomposition of  $CaCO_{3(s)}$  from the pellet reactor in a calciner to regenerate the captured  $CO_{2(g)}$ . This is followed by slaking of the resultant  $CaO_s$  using steam to produce  $Ca(OH)_{2(aq)}$  for use in the pellet reactor, thus completing the second cycle. In [7], the second loop is remodeled to include solids storage silos to enable flexible scheduling of the calcination process to maximize RE utilization in the energy-intensive process. In this paper, a further modification is proposed to improve the utilization of the energy contained in the  $CO_{2(g)}$  and  $CaO_s$  from the calciner. This is done by including  $H_{2(g)}$  production in the process flow to make use of the waste heat. Pre-heat cyclones included by Keith et al [4] are utilized here except for the one that is associated with the air separation plant, which is not necessary in our electrified calciner. The process flow of the proposed DAC plant with an integrated PEM is shown in Figure 1.

## Optimization Model of the Flexible DAC with Hydrogen Production

A new model is developed to include  $H_{2(g)}$  production constraints. The model is defined as follows:

$$\min F = \sum_{t \in T} (C_{ng} p_t^{ng} \Delta t + C_i C_c p_t^{ng} \Delta t - C_{H_2} m_t^{H_2}) \quad (1)$$



$$p_t^{pem} \geq \alpha m_t^{H_2O,pem}, \quad \forall t \in T \quad (2)$$

$$m_t^{H_2O,out} = m_t^{H_2O,in} - m_t^{H_2O,pem}, \quad \forall t \in T \quad (3)$$

$$m_t^{H_2} = \frac{2}{18} m_t^{H_2O,pem}, \quad \forall t \in T \quad (4)$$

$$m_t^{O_2} = \frac{16}{18} (m_t^{H_2O,pem}), \quad \forall t \in T \quad (5)$$

$$p_t^{pem} + m_t^{H_2O,in} c_{H_2O} (T_{hx}^{H_2O} - T_a) = \beta m_t^{H_2O,pem} + m_t^{H_2} c_{H_2} (T_{pem} - T_a) + m_t^{O_2} c_{O_2} (T_{pem} - T_a) + m_t^{H_2O,out} c_{H_2O} (T_{pem} - T_a), \quad \forall t \in T \quad (6)$$

$$m_t^{H_2O,in} c_{H_2O} (T_{hx}^{H_2O} - T_a) = \eta x_t^c (0.44 c^{O_2} (T_{dac}^{CO_2} - T_{hx}^{CO_2}) + 0.56 c^{CaO} (T_{dac}^{CaO} - T_{hx}^{CaO})), \quad \forall t \in T \quad (7)$$

$$P^{df} + p_t^{dv} + p_t^{pem} = p_t^{ng} + p_t^w + p_t^s, \quad \forall t \in T \quad (8)$$

$$p_t^{dv} = \phi m_t^{CO_2}, \quad \forall t \in T \quad (9)$$

$$0 \leq p_t^w \leq P_{max}^w, \quad \forall t \in T \quad (10)$$

$$0 \leq p_t^s \leq P_{max}^s, \quad \forall t \in T \quad (11)$$

$$0 \leq p_t^{ng} \leq P_{max}^{ng}, \quad \forall t \in T \quad (12)$$

$$m_t^{cc} = m_{t-1}^{cc} + r_t^{cc} \Delta t, \quad \forall t \in T \quad (13)$$

$$-R_{max}^{cc} \leq r_t^{cc} \leq R_{max}^{cc}, \quad \forall t \in T \quad (14)$$

$$0 \leq m_t^{cc} \leq M_{max}^{cc}, \quad \forall t \in T \quad (15)$$

$$m_t^{cc} = 0, \quad t \in \{0, T\} \quad (16)$$

$$m_t^c = m_{t-1}^c + r_t^c \Delta t, \quad \forall t \in T \quad (17)$$

$$-R_{max}^c \leq r_t^c \leq R_{max}^c, \quad \forall t \in T \quad (18)$$

$$0 \leq m_t^c \leq M_{max}^c, \quad \forall t \in T \quad (19)$$

$$m_0^c = m_T^c \quad (20)$$

$$x_t^c \leq X_{max}^c, \quad \forall t \in T \quad (21)$$

$$r_t^{cc} + x_t^c = X^p, \quad \forall t \in T \quad (22)$$

$$0.56 x_t^c = X^s + r_t^c, \quad \forall t \in T \quad (23)$$

$$m_t^{CO_2} = 0.44 x_t^c \Delta t, \quad \forall t \in T \quad (24)$$

The objective of the model is to minimize energy cost as given by equation (1), where the sales of hydrogen is subtracted to encourage heat recovery. Equation (2)-(7) are the PEM plant constraints defined as follows: the PEM electrical energy supply constraint (2); the water flow balance for the PEM (3); the  $H_{2(g)}$  output (4);  $O_{2(g)}$  output (5); power equilibrium at the PEM (6); and the HX thermal energy balance (7). In (2),  $\alpha$  is the Gibb's free energy,  $\Delta G$ , of water at the operating temperature (353K) in MWh/kg. This is obtained from the value of  $\Delta G$  in kJ/mol expressed as [11], [12], [13]:

$$\Delta G = nF (1.229 - 0.0009(\tau - 298)) \quad (25)$$

where  $n = 2$  is the number of electrons transferred in the electrolysis of a molecule of water,  $F = 96485C$  is Faraday's constant and  $\tau$  is the temperature in Kelvin. Equations (8)-(12) and (13)-(26) are the energy supply constraints and mass flow constraints for the flexible DAC plant; the reader is referred to [7] for a detailed description of these constraints. The energy supply constraints are briefly defined as follows: power equilibrium at the supply bus (8), variable calciner and compressor demand (9), wind power generator boundary (10), solar generation constraint (11), and NGG capacity constraint (12). The mass flow constraints for the DAC plant are defined as follows:  $CaCO_3$  storage silo mass flow dynamics (13), silo flowrate (14) and capacity limits (15); capture rate enforcement constraint for the  $CaCO_3$  silo (16), which ensures all the captured  $CO_{2(g)}$  is regenerated;  $CaO$  silo mass flow dynamics (17), flowrate (18) and capacity limits (19); capture rate enforcement constraint for the  $CaO$  silo (20); calciner throughput limit (21); coupling constraint between the calciner and the  $CaCO_3$  silo (22); coupling constraint between the calciner and the  $CaO$  silo (23) and a constraint to compute the mass of  $CO_{2(g)}$  regenerated. Table 2 shows the model decision variables that are optimized.

## Data

The parameters of the model are shown in Table 1. Details of the RE profiles and the DAC parameters have been provided in [7].

## RESULTS AND DISCUSSION

This paper models  $H_{2(g)}$  production from waste heat from a liquid solvent DAC plant using a PEM. The PEM is assumed to be able to make use of all the heat recovered to produce the  $H_{2(g)}$ , thus, the PEM size is not limiting. In this section, PEM output and the impact of this modification on the overall energy cost of the plant are discussed.

### PEM Performance and Energy Consumption

Over the 336 hours modeled, which represent variation in renewable energy supply, a total of about 910,890 kg of  $H_{2(g)}$  is produced, which translates to about 2,711 kg/h. This level of production would need a 136 MW PEM, if losses are ignored. Given that the higher heating value (HHV) of  $H_{2(g)}$  is about 0.0394 MWh/kg, the output energy is close to 35,889 MWh of energy produced at an average rate of 106 MW, which is equivalent to 78.4% efficiency. This efficiency value would be lower if all the losses are considered. However, it would still result in a better efficiency compared to using the recovered heat to generate power using an organic Rankine cycle, whose efficiency is less than 30%.

**Table 1:** Model Parameters

Parameter	Description	Value and Units
$C_c$	Carbon tax	37.1 \$/tonne- $CO_2$
$C_i$	NGG carbon intensity	0.3894 tonne- $CO_2$ /MWh
$C_{H_2}$	Price of $H_{2(g)}$	\$/kg
$C_{ng}$	Cost of natural gas	27.77 \$/MWh
$c^{CO_2}$	Specific heat capacity of $CO_2$	1.065 kJ/kg/K
$c^{O_2}$	Specific heat capacity of $O_2$	0.924 kJ/kg/K
$c^{H_2O}$	Specific heat capacity of $H_2O$	4.184 kJ/kg/K
$c^{H_2}$	Specific heat capacity of $H_2$	14.493 kJ/kg/K
$c^{CaO}$	Specific heat capacity of $CaO$	0.945 kJ/kg/K
$M_{max}^c$	$CaO$ silo capacity	370 tonnes
$M_{max}^{cc}$	$CaCO_3$ silo capacity	660 tonnes
$p^{df}$	Fixed DAC demand	20 MW
$p_{max}^{ng}$	NGG capacity	165 MW
$p_{max}^s$	Maximum PV power	110 MW
$p_{max}^w$	Maximum wind power	140 MW
$R_{max}^c$	$CaO$ silo maximum flowrate	134 tonnes/h
$R_{max}^{cc}$	$CaCO_3$ silo maximum flowrate	240 tonnes/h
$T$	Optimization horizon	336h
$T_a$	Ambient temperature	25 °C
$T_{dac}^{CO_2}$	Temperature of $CO_2$ leaving the pre-heat cyclones of DAC.	450 °C
$T_{dac}^{CaO}$	Temperature of $CaO$ leaving the calciner	900 °C
$T_{hx}^{CaO}$	Temperature of $CaO$ leaving the HX	300 °C
$T_{hx}^{H_2O}$	Temperature of $H_2O$ leaving the HX	80 °C
$T_{pem}$	PEM operating temperature	80 °C
$X_{min}^c$	Minimum calciner flowrate	20 tonnes/h
$X_{min}^{cc}$	Minimum calciner flowrate	500 tonnes/h
$X^p$	Pellet reactor flowrate	260 tonnes/h
$X^s$	Slaker flowrate	145.6 tonnes/h
$\alpha$	Gibb's free energy of $H_2O$ at $T_{pem}$	0.00351 MWh/kg
$\beta$	Enthalpy of formation of $H_2O$	0.00441 MWh/kg
$\eta$	HX efficiency	0.8
$\phi$	Calciner and compressor consumption	1.226 MWh/tonne- $CO_2$
$\Delta t$	Timestep size	1h

### Sensitivity Analysis: Factors Affecting the Mass of Hydrogen Produced

The production of  $H_{2(g)}$  requires both thermal and electrical energy. Although the amount of thermal energy recovered from the calciner is constant – provided the

Variable	Description	Units
$m_t^c$	Mass of stored $CaO$ at time, $t$	tonne
$m_t^{cc}$	Mass of stored $CaCO_3$ at time, $t$	tonne
$m_t^{CO_2}$	Mass of $CO_2$ regenerated at time, $t$	tonne
$m_t^{H_2}$	Mass of $H_2$ produced at time, $t$	kg
$m_t^{H_2O,in}$	Mass of $H_2O$ flowing into the PEM at time, $t$	kg
$m_t^{H_2O,out}$	Mass of $H_2O$ flowing out of the PEM at time, $t$	kg
$m_t^{H_2O,pem}$	Mass of $H_2O$ electrolyzed by the PEM at time, $t$	kg
$m_t^{O_2}$	Mass of $O_2$ produced by the PEM at time, $t$	kg
$p_t^{dv}$	Variable demand of the DAC plant at time, $t$	MW
$p_t^{ng}$	Power generated by the NGG at time, $t$	MW
$p_t^{pem}$	Electrical power supplied to the PEM at time, $t$	MW
$p_t^s$	Solar power consumed at time, $t$	MW
$p_t^w$	Wind power consumed at time, $t$	MW
$r_t^c$	Mass flow rate of the $CaO$ silo at time, $t$	tonne/h
$r_t^{cc}$	Mass flow rate of the $CaCO_3$ silo at time, $t$	tonne/h
$x_t^c$	Calciner flowrate at time, $t$	tonne/h

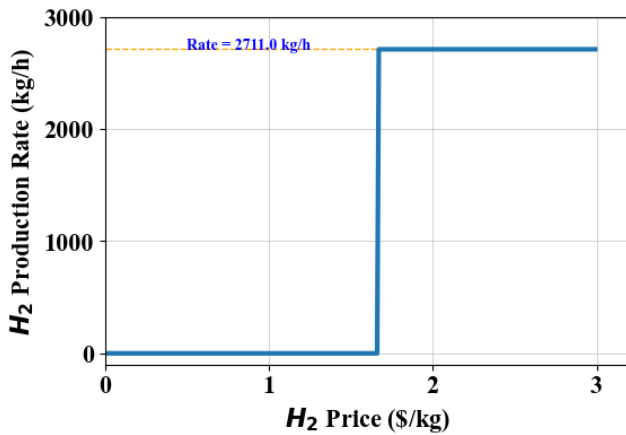
$CO_{2(g)}$  capture rate is enforced – the electrical input is derived from RE and the NGG. The objective function of the optimization model penalizes the use of NGG while encouraging the production of  $H_{2(g)}$ . Therefore, two main factors affect the amount of  $H_{2(g)}$  produced.

First, if the cost of NGG is constant, the price of  $H_{2(g)}$  determines whether  $H_{2(g)}$  is produced. Observing the amount of  $H_{2(g)}$  produced per hour for the value of  $C_{H_2}$  ranging from 0.0 to 3.0 \$/kg, it was established that the  $H_{2(g)}$  production becomes attractive at  $C_{H_2} = 1.67$  \$/kg. Considering that the higher heating value (HHV) of  $H_{2(g)}$  is about 0.0394 MWh/kg, this price coincides with an energy price of about 42.38 \$/MWh, which is very close to but slightly above the cost of NGG power from the NGG, which is 42.22 \$/MWh. Figure 2 shows the impact of the price of  $H_{2(g)}$  on its hourly production rate.

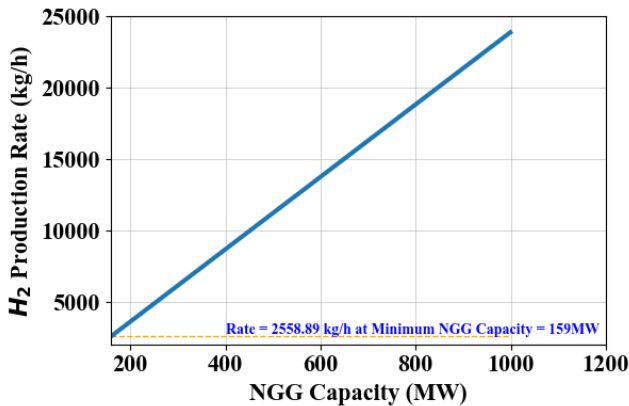
The second important factor affecting the  $H_{2(g)}$  production rate is the NGG capacity. If the price is set at 1.67 \$/kg, increasing the NGG capacity increases the rate of  $H_{2(g)}$  production provided the PEM capacity is not limited. This is because the increase in NGG capacity increases the amount of energy available for  $H_{2(g)}$  production. The production also increases because the price of

**Table 2:** Model Decision Variables

$H_{2(g)}$  is set slightly higher than the cost of energy from the NGG. Figure 3 shows the impact of the NGG capacity on  $H_{2(g)}$  production rate.



**Figure 2.** Impact of  $H_{2(g)}$  price on the production rate.



**Figure 3.** Impact of NGG capacity on  $H_{2(g)}$  production rate.

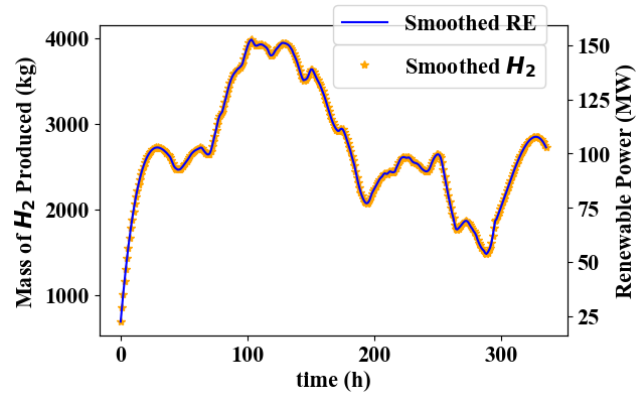
Though the rate of production of  $H_{2(g)}$  increases with NGG capacity, it also leads to more emission of  $CO_{2(g)}$ , therefore, only the capacity needed to sustain the system is allowed, which is set at 165 MW.

### Implications of Including the PEM on System Energy Use Dynamics

With flexibility in the  $H_{2(g)}$ , the role of the solids storage silos in minimizing curtailment vanishes, provided the PEM size is not a limiting factor. Therefore, the silo size no longer impacts the cost of energy, which departs from the findings in [7] where the silos played a major role in curtailment reduction. This is because the cheap RE facilitates the  $H_{2(g)}$  production so that none of the available renewable energy is curtailed. Furthermore, including the PEM increases the energy demand considerably such that no moments of oversupply exist.

Consequently, the calciner and the NGG operate at full capacity throughout the optimization horizon. This is because there is an incentive in the objective function for increasing hydrogen production. The calciner operates at full capacity to supply the thermal demands of the PEM while the NGG supplies the electrical energy demand.

However, the PEM's output profile coincides with that of the RE. This is because RE provides cheap electrical energy for hydrogen production, which could be used for both the PEM's thermal and electrical needs. Figure 4 shows the mass of  $H_{2(g)}$  produced by the PEM, which aligns with RE availability and use.



**Figure 4.**  $H_{2(g)}$  production profile alongside RE supply (Smoothing was done using Savitzky–Golay filter).

## CONCLUSION

This paper presents a linear programming optimization model which was developed to incorporate a PEM in the process flow design of an LSDAC plant to maximize the use of waste heat for hydrogen production. Novel mass and energy flow constraints are developed to maintain energy and water supply to the PEM electrolyser. Results show that 911 tonnes of hydrogen could be produced over 336 hours of operation by a 136 MW PEM. Sensitivity analyses reveal that hydrogen production is only prioritized if its sale value is higher than the cost of natural gas. Under such market conditions, the production of hydrogen would only be limited by the PEM size and availability of the required electrical energy. The inclusion of the PEM also eliminates the need for solids storage silos, which were previously required to enhance process flexibility and use of variable renewable energy. Future research could develop temperature-dependent models for the PEM hydrogen output profile to enable analysis of the impact of temperature on the amount of hydrogen produced from the recovered heat.

## ACKNOWLEDGEMENTS

This project was undertaken with the financial support of the Government of Canada.

## REFERENCES

1. International Energy Agency, "Direct Air Capture: A key technology for net zero," Paris, Apr. 2022. [Online]. Available: [www.iea.org/t&c/](http://www.iea.org/t&c/)
2. D. Y. Shu, S. Deutz, B. A. Winter, N. Baumgärtner, L. Leenders, and A. Bardow, "The role of carbon capture and storage to achieve net-zero energy systems: Trade-offs between economics and the environment," *Renewable and Sustainable Energy Reviews*, vol. 178, May 2023, doi: 10.1016/j.rser.2023.113246.
3. J. E. T. Bistline and G. J. Blanford, "Impact of carbon dioxide removal technologies on deep decarbonization of the electric power sector," *Nat Commun*, vol. 12, no. 1, Dec. 2021, doi: 10.1038/s41467-021-23554-6.
4. D. W. Keith, G. Holmes, D. St. Angelo, and K. Heidel, "A Process for Capturing CO<sub>2</sub> from the Atmosphere," *Joule*, vol. 2, no. 8, pp. 1573–1594, Aug. 2018, doi: 10.1016/j.joule.2018.05.006.
5. K. An, A. Farooqui, and S. T. McCoy, "The impact of climate on solvent-based direct air capture systems," *Appl Energy*, vol. 325, Nov. 2022, doi: 10.1016/j.apenergy.2022.119895.
6. Q. Shu, L. Legrand, P. Kuntke, M. Tedesco, and H. V. M. Hamelers, "Electrochemical Regeneration of Spent Alkaline Absorbent from Direct Air Capture," *Environ Sci Technol*, vol. 54, no. 14, pp. 8990–8998, Jul. 2020, doi: 10.1021/acs.est.0c01977.
7. E. O. Arwa and K. R. Schell, "Batteries or silos: Optimizing storage capacity in direct air capture plants to maximize renewable energy use," *Appl Energy*, vol. 355, p. 122345, Feb. 2024, doi: 10.1016/j.apenergy.2023.122345.
8. H. Nami, F. Mohammadkhani, and F. Ranjbar, "Utilization of waste heat from GTMHR for hydrogen generation via combination of organic Rankine cycles and PEM electrolysis," *Energy Convers Manag*, vol. 127, pp. 589–598, Nov. 2016, doi: 10.1016/j.enconman.2016.09.043.
9. D. J. Singh Aulakh, K. G. Boulama, and J. G. Pharoah, "On the reduction of electric energy consumption in electrolysis: A thermodynamic study," *Int J Hydrogen Energy*, vol. 46, no. 33, pp. 17084–17096, May 2021, doi: 10.1016/j.ijhydene.2021.02.161.
10. Z. Wang, X. Wang, Z. Chen, Z. Liao, C. Xu, and X. Du, "Energy and exergy analysis of a proton exchange membrane water electrolysis system without additional internal cooling," *Renew Energy*, vol. 180, pp. 1333–1343, Dec. 2021, doi: 10.1016/j.renene.2021.09.037.
11. C. Yilmaz and M. Kanoglu, "Thermodynamic evaluation of geothermal energy powered hydrogen production by PEM water electrolysis," *Energy*, vol. 69, pp. 592–602, May 2014, doi: 10.1016/j.energy.2014.03.054.
12. S. Sharifian, N. Asasian Kolar, and M. Harasek, "Transient simulation and modeling of photovoltaic-PEM water electrolysis," *Energy Sources, Part A: Recovery, Utilization and Environmental Effects*, vol. 42, no. 9, pp. 1097–1107, May 2020, doi: 10.1080/15567036.2019.1602220.
13. W. Li, H. Tian, L. Ma, Y. Wang, X. Liu, and X. Gao, "Low-temperature water electrolysis: fundamentals, progress, and new strategies," *Materials Advances*, vol. 3, no. 14. Royal Society of Chemistry, pp. 5598–5644, May 17, 2022. doi: 10.1039/d2ma00185c.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Technoeconomic Analysis of Chemical Looping Ammonia Synthesis Reactors to Enable Green Ammonia Production

Laron D. Burrows<sup>a</sup>, and George M. Bollas<sup>a\*</sup>

<sup>a</sup> University of Connecticut, Department of Chemical & Biomolecular Engineering, Storrs, CT, USA

\* Corresponding Author: [george.bollas@uconn.edu](mailto:george.bollas@uconn.edu).

## ABSTRACT

Chemical looping ammonia synthesis (CLAS) is a new ammonia synthesis method capable of efficiently synthesizing ammonia at atmospheric pressure. The low-pressure operation of CLAS systems could decrease the capital and operational costs of ammonia synthesis. Despite its early developmental stage, the use of standard process engineering equipment in CLAS makes it possible to reasonably assess its economic potential. In this study, we evaluated the technoeconomic potential of CLAS systems in comparison to a Haber-Bosch (HB) synthesis process in the context of green ammonia production. CLAS is more compatible with the separate nitrogen and hydrogen feedstocks used in green ammonia production, and cost savings from CLAS could improve the economic viability of green ammonia production. Ammonia synthesis loops were modeled in Aspen Plus and the levelized cost of ammonia (LCOA) of each system was calculated. Three CLAS systems; two high temperature and one low-temperature chemical loop, were compared to a conventional HB system of equivalent size. This study found that CLAS can reduce the synthesis cost by 90% and that the low temperature CLAS is more economically viable than the high temperature CLAS. The need for an external heater in the high temperature CLAS diminished any cost savings that would have been realized due to the low-pressure operation. This work highlights the potential of CLAS to reduce ammonia synthesis costs and emphasizes the need for further development of low-temperature CLAS systems.

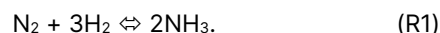
**Keywords:** Aspen Plus, Food & Agricultural Processes, Modelling and Simulations, Technoeconomic Analysis, Process Design

## INTRODUCTION

Ammonia is a vital chemical for our modern society. Most of the industrially produced ammonia is used as a fertilizer to feed the global population, and the remainder is used to manufacture pharmaceuticals, chemicals, or used as a refrigerant [1]. The high energy density of ammonia and hydrogen capacity per volume, makes ammonia an attractive solution as a carbon-free energy carrier [1]. Because of this, ammonia is expected to play a pivotal role in decarbonizing notoriously difficult to abate sectors like aviation, and maritime shipping [2-3]. However, ammonia production itself is a carbon-intensive and hard to abate sector. Global ammonia production emits 1.8% of annual CO<sub>2</sub> emissions while consuming 1% of the annual energy demand [1]. As global reliance on ammonia grows with the emergence of new applications, it is crucial to

ensure that the rising demand for ammonia production does not lead to a corresponding increase in CO<sub>2</sub> emissions.

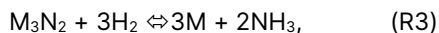
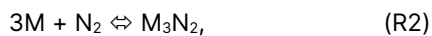
Industrial ammonia synthesis is performed by the Haber-Bosch (HB) process which combines nitrogen and hydrogen to synthesize ammonia via R1:



Although the ammonia synthesis reaction doesn't directly emit carbon dioxide, it results in the emission of 2 tons of CO<sub>2</sub> for every ton of ammonia produced [1]. This significant carbon footprint stems largely from generating hydrogen via steam methane reforming and combusting natural gas to heat the process [1]. Consequently, adopting carbon-free sources for hydrogen and heating is essential to curtail CO<sub>2</sub> emissions associated with ammonia production.

Green ammonia production offers a low-carbon solution for ammonia synthesis. This approach typically involves generating hydrogen through water electrolysis and extracting nitrogen from air using an air separation unit (ASU). These gases are then reacted in a Haber Bosch (HB) reactor to produce ammonia, with the entire process powered by renewable energy, making it environmentally friendly [2]. While technically viable, green ammonia production is not yet economically feasible in most cases [1]. Current efforts to lower production costs have primarily focused on reducing electrolyzer costs, with less emphasis on cost savings within the ammonia synthesis loop. Additionally, the HB reactor's sensitivity to minor fluctuations poses a challenge in integrating with intermittent renewable energy sources, underscoring the need for a more robust synthesis loop [4].

Chemical Looping Ammonia Synthesis (CLAS) is a new ammonia synthesis method that enables efficient ammonia production at atmospheric pressure and mild temperatures (<300°C) [5]. CLAS synthesizes ammonia in two or more steps mediated by a carrier material, typically a metal catalyst. A generalized CLAS process is shown in R2-R3:



where nitrogen is fixed to the metal catalyst in R2 and removed as ammonia in R3 using hydrogen. The primary benefit of CLAS is its ability to circumvent the equilibrium constraints of the traditional HB process (R1) [5]. This allows for ammonia production at reduced pressures and temperatures, which can significantly lower both capital and operational costs by enabling the use of more affordable compressors and reactors than those required for the HB process. Furthermore, CLAS is particularly well-suited for green ammonia production, which generates nitrogen and hydrogen in separate streams – a requirement for CLAS reactors. CLAS systems can vary in their number of steps, carrier materials, operating temperatures, and hydrogen sources making them difficult to assess [5]. Despite ongoing development of more CLAS systems, there's a gap in assessing their techno-economic viability, especially as a direct replacement for the HB synthesis loop in green ammonia production proposals. Currently, CLAS has a low technology readiness level (TRL) of 3. However, its use of established technologies such as compressors, heaters, heat exchangers, and fixed bed reactors allows for a reasonably confident early-stage economic assessment. In this work we aim to assess the techno-economic feasibility of CLAS reactors as a drop-in replacement for a HB synthesis loop in a green ammonia production system.

## METHOD

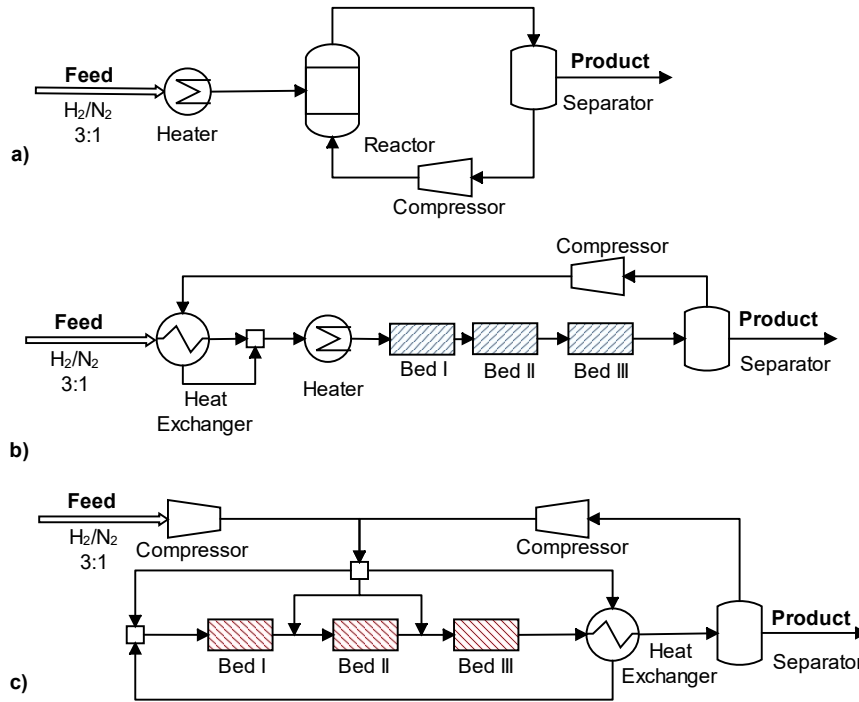
In this study, HB and CLAS systems were simulated in Aspen Plus and their levelized cost of ammonia (LCOA) were compared. For the HB process, we simulated a state-of-the-art loop featuring a three-bed reactor. For the CLAS systems, we examined two high-temperature processes operating at 400°C using data from two preliminary catalysts (Catalyst A & B) studied in-house reported in [6], and a low-temperature process operating at 270°C utilizing Ni-BaH<sub>2</sub> [7]. Each synthesis loop was fed with 2,600 kmol/h of nitrogen and hydrogen in a ratio of 3:1 at a temperature of 80°C and a pressure of 5 atm to replicate the pressure from a PEM electrolyzer [1]. The capital cost of each equipment was calculated using published correlations [8]. The levelized cost of ammonia was calculated using a 20-year period and 5% discount rate. The effect of conversion and CLAS reactor configuration was studied in relation to LCOA. The details of this analysis are described in this section.

### Chemical Looping Reactor Model

The chemical looping reactor models were simulated at cyclical steady state using adiabatic RSTOIC reactors with a fixed fractional conversion of N<sub>2</sub> [9]. In this context, cyclical steady state describes a condition in which the reactor performance is stabilized across multiple chemical looping cycles and reaches a steady state. In this work the single reactor model shown in Figure 1.a, and the three-reactor model shown in Figure 1.b were studied. In the single reactor model, the feed gas is heated to the set point temperature of the chemical loop before being reacted. Ammonia is separated from the reactor outlet using a separator, and the unreacted gases are recycled back to the reactor using a compressor. In the three-reactor setup three CLAS reactors are connected in series and the feed gas is preheated in a heat exchanger, before being sent to the heater that is set to the setpoint of the chemical looping reactor. Again, ammonia is separated from the reactor outlet, and unreacted N<sub>2</sub> and H<sub>2</sub> are recycled, forming the synthesis loop. Three CLAS configurations were examined, using preliminary production rates of 105 μmol/g<sub>cat</sub>·h for Catalyst A, 2 mmol/g<sub>cat</sub>·h for Catalyst B, and 28 mmol/g<sub>cat</sub>·h for Ni-BaH<sub>2</sub> [7].

### Haber-Bosch Reactor Model

The Haber-Bosch synthesis loop shown in Figure 1.c, was simulated as a three-bed reactor system with a preheater and quench splits [10]. The molar flow of the feed, shown in Figure 1.c, was split sending 23%, 13.9% and 12.7% to beds 1, 2, and 3, respectively and the balance sent to the preheater [10]. The catalyst beds were modeled as adiabatic plugged flow reactors with a catalyst void fraction of 0.33 and a density of 2200 kg/m<sup>3</sup>. The iron-based catalyst kinetics are described by Eq. 1 [10]:



**Figure 1:** Diagrams of the ammonia synthesis loops modeled in this work where; a) is the single CLAS reactor model, b) is the three reactor CLAS model, and c) is the three bed Haber Bosch model.

$$r_{Fe} = \frac{9.5}{\rho_{cat}} \left( \frac{k_{Fe} P_{N_2} P_{H_2}^{1.5}}{P_{NH_3}} - \frac{k_{-Fe} P_{NH_3}}{P_{H_2}^{1.5}} \right), \quad (1)$$

where  $\rho_{cat}$  is the catalyst bulk density,  $P_i$  is the partial pressure of component  $i$  in bar, and  $k_{Fe}$  and  $k_{-Fe}$  are the kinetic constants for the forward and reverse reactions of R1 respectively as:

$$k_{Fe} = 1.79 \times 10^4 e^{\left(-\frac{87,090}{RT}\right)}, \quad (2)$$

$$k_{-Fe} = 2.75 \times 10^{16} e^{\left(-\frac{198,464}{RT}\right)}, \quad (3)$$

where  $R$  is the universal gas constant, and  $T$  is the gas temperature in Kelvin.

### Auxiliary Equipment

Auxiliary equipment was simulated to replicate realistic operation of each synthesis loop. A multistage compressor was used to elevate the pressure from 5 atm to 200 atm for the HB synthesis loop. The compressor includes an interstage cooler to reduce the outlet gases temperature to 250°C. The cooler was modeled as a heat exchanger with pumped cooling water as the working fluid. The pumps were modeled to increase the cooling water pressure by 1 atm with a 95% pump and driver efficiency. The cost of cooling water was not considered in this work. Compressors were simulated as isentropic using the ASME method with an efficiency of 95%. The compressors were powered by explosion proof drivers

[8]. Heat exchangers were modeled with a 1 atm pressure drop across each side with a design specification to achieve a 50°C temperature change. The reported area of the heat exchangers was used to size the equipment. Heaters were modeled as electric heaters in Aspen Plus, and the duty of the heater was used to size the equipment. The separator used in each model assumed to be an absorbent based separator with a return stream of 250°C and a pressure drop of 1 atm [11]. The cost of separation was not included in the scope of this work.

### Economic Calculations

The capital cost of each synthesis loop was calculated using Eq 4-6 :

$$C^{cap} = \sum_j^n C_j, \quad (4)$$

$$C_j = C_e \times AF \times P_e \times \frac{CEPCI}{1000}, \quad (5)$$

$$\log_{10} C_e = K_1 + K_2 \log_{10}(A) + K_3 [\log_{10}(A)]^2, \quad (6)$$

where  $C^{cap}$  is the capital cost of the synthesis loop,  $C_j$  is the bare module cost of equipment  $j$ ,  $n$  is the number of equipment in the synthesis loop,  $AF$  is the alloy factor used as 2.7 for stainless steel,  $P_e$  is the pressure adjustment factor for each equipment, and the chemical engineering plant cost index,  $CEPCI$ , was taken as 800 for 2023 [8,12].  $C_e$  is the bare cost of the equipment defined by Eq. 6 where parameters  $K1 - K3$  and  $A$  are described

in Table 1 for each equipment.

**Table 1:** Equipment capital cost parameters for Eq. 6 [8].

Unit	Basis (A)	K1	K2	K3
Heater	Duty (kW)	3.068	0.6597	0.0194
Heat Exchanger	Area (m <sup>2</sup> )	2.7652	0.7282	0.0783
Compressor	Duty (kW)	2.2897	1.3604	-0.1027
Driver	Duty (kW)	2.4604	1.4191	-0.1798
Pump	Duty (kW)	3.3892	0.3161	0.1220

The reactor capital costs were calculated using Eq. 7:

$$C_{reactor}^{cap} = C_{fixed}^{cap} + f_p C_{ref}^{cap} \left( \frac{V}{20 \text{ m}^3} \right)^{0.52} + 15.50 w_{cat}, \quad (7)$$

where  $w_{cat}$  is the catalyst weight in kg,  $V$  is the reactor volume in m<sup>3</sup>,  $C_{ref}^{cap}$  is the reference reactor cost of \$268,000,  $C_{fixed}^{cap}$  is the fixed capital cost of the reactor at \$66,800,  $f_p$  is the pressure multiplier that accounts for the increase in vessel wall thickness at higher pressures, calculated as:

$$f_p = 0.125 \left( \frac{P}{10} \right) + 0.875, \quad (8)$$

where  $P$  is in units of bar [11]. To calculate the volume of the RSTOIC reactors, we first established the required weight of the catalyst by dividing the ammonia production rate of each reactor by the catalyst production rate per gram of catalyst. Next, we determined the catalyst's volume using its density. Finally, we increased this volume by 25% to estimate the reactor volume. The HB catalyst cost was assumed to be \$15.5/kg [11] and the CLAS catalyst costs were calculated using the bare metal costs [13-15] plus an 80% markup.

The operating cost of each synthesis loop,  $C^{op}$ , was calculated using Eq. 5:

$$C^{op} = \sum_j^n C_j^{op}, \quad (9)$$

$$C_j^{op} = C_{j,duty} \times e_{cost} \times O_p, \quad (10)$$

where  $C_j^{op}$  is the operating cost of equipment  $j$  in the synthesis loop defined by Eq 10.  $C_{j,duty}$  is the reported duty of equipment  $j$ ,  $n$  is the number of equipment in the synthesis loop,  $e_{cost}$  is the cost of electricity assumed as 7 cents per kWh, and  $O_p$  is the operating period in hours. The levelized cost of ammonia was then calculated using Eq. 11 [16]:

$$LCOA = \frac{\sum_{t=0}^T \frac{C_t^{cap} + C_t^{op}}{(1+r)^t}}{\sum_{t=0}^T \frac{m_{NH_3}}{(1+r)^t}}, \quad (11)$$

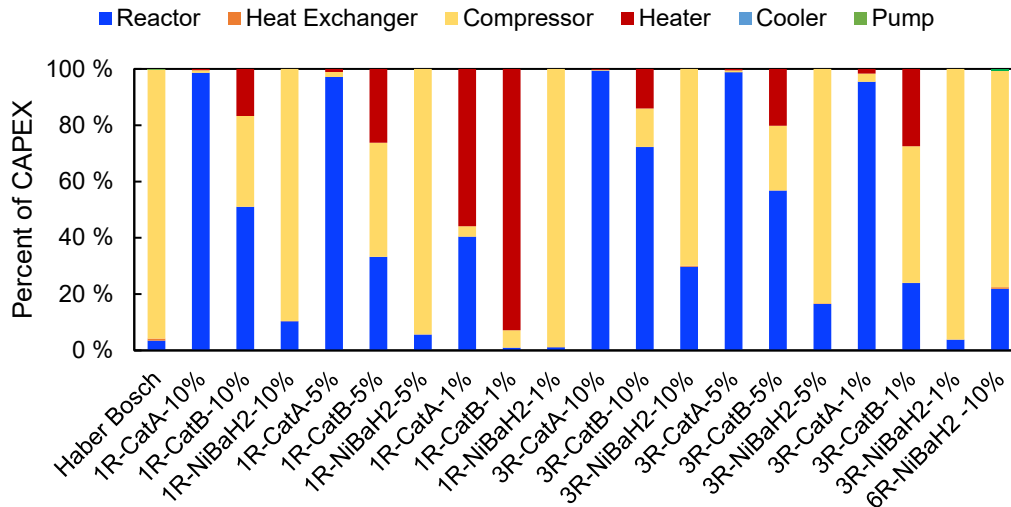
where  $T$  is the lifetime of the plant,  $t$  is the year,  $r$  is the discount rate, and  $m_{NH_3}$  is the annual ammonia production in tons. The LCOA was calculated for each synthesis loop using a 5% discount rate over the 20-year lifetime of the plant with an 90% uptime.

## Parameter Analysis

Four studies were conducted to understand how different model parameters affect the levelized cost of ammonia in CLAS systems. The first study varied the reactor conversion rates between 1%, 5%, and 10%. The second study changed the number of reactors in a series, while varying the reactor conversion. The third study involved a sensitivity analysis on the three reactor Ni-BaH<sub>2</sub> CLAS system, where N<sub>2</sub> conversion, electricity cost, catalyst cost, and plant uptime were varied by ±20% individually to determine their impact on the LCOA. Lastly, in the fourth study, a six-reactor CLAS system was studied using the Ni-BaH<sub>2</sub> configuration with a conversion of 10% to determine the marginal cost of adding more reactors.

## RESULTS

Figure 2 shows the levelized cost of ammonia for the single and three-reactor CLAS systems using Catalyst A, B, and Ni-BaH<sub>2</sub>. The figure identifies a "cost-effective" zone where CLAS systems are more economical than the HB synthesis loop, which costs \$63.9/ton. Catalyst A has the lowest production rate of the CLAS systems studied (105 μmol/g<sub>cat</sub>·h), and due to this it is the least economical option. The low production rate requires more catalyst resulting in larger more expensive reactors and a higher LCOA. Increasing the number of reactors with Catalyst A did not make the system more economical than the HB process as shown in Figure 2. Catalyst B has a higher production rate (2 mmol/g<sub>cat</sub>·h) than Catalyst A; however, it was only economically competitive in a three-reactor setup sized for a 10% conversion. In this case, the addition of more reactors reduced the LCOA of the configuration, to where it could be cost competitive with the HB process. The CLAS with Ni-BaH<sub>2</sub> is the only CLAS system that have economical configurations with the single reactor and the three-reactor setup. The Ni-BaH<sub>2</sub> configuration is less expensive overall due to its high production rate (28 mmol/g<sub>cat</sub>·h) and low temperature operation. The high production rate reduces the amount of catalyst needed which in-turn reduces the reactor size and cost. Additionally, the low temperature operation eliminates



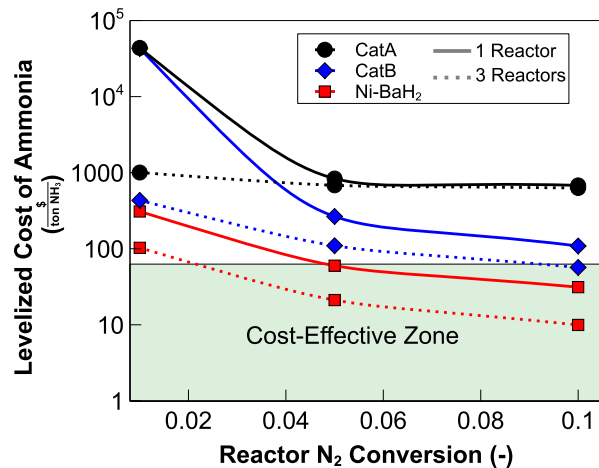
**Figure 3:** The capital contribution of each equipment in the synthesis loops studied where R represents the number of reactors and the percentage represents the N<sub>2</sub> conversion of the reactors.

the need for large external heaters. As a result, its single reactor configuration can achieve a LCOA lower than the HB process. A six-reactor configuration with Ni-BaH<sub>2</sub> was studied and shown to achieve a \$9.73/ton LCOA, which represents an 84.8% reduction over the HB loop. The capital cost of the six-reactor synthesis loop is compared to the HB process in Table 2, showing an 82% reduction in capital costs. Table 2 shows that the compressor costs, inclusive of the driver costs, are the main differences between the two systems.

compressor. In the CLAS systems, the major cost contributor varied significantly which each configuration. In the high temperature chemical loops, the need for an external heater drove up the capital cost making it a major cost contributor. In configurations with Catalyst A, the low production rate of the catalyst resulted in the reactor being the major cost driver in most configurations.

**Table 2:** Synthesis loop capital cost (\$Millions USD), comparing the six-reactor Ni-BaH<sub>2</sub> system to the HB.

Stream	Haber-Bosch	Chemical Loop
Compressor	29.1	4.2
Reactor + Catalyst	0.98	1.10
Heat Exchanger	0.21	0.06
Pump	0.04	0.04
Total	30.33	5.4

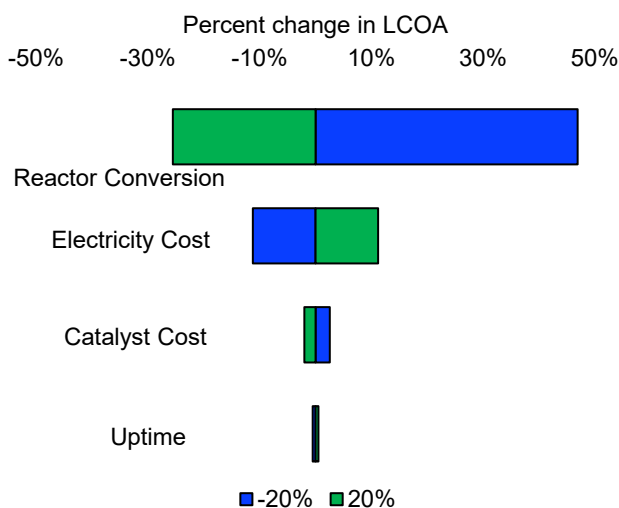


**Figure 2:** LCOA for each system, highlighting the “cost-effective zone” where systems outperform the HB loop.

The capital contribution of each equipment in the synthesis loop is shown in Figure 3. In the HB synthesis loop, over 95% of the capital expense is attributed to the

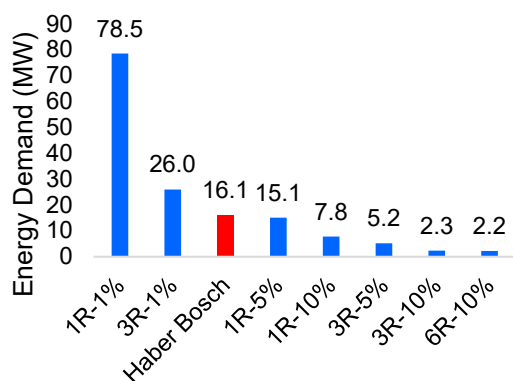
A sensitivity analysis was performed on the three-reactor Ni-BaH<sub>2</sub> CLAS system at 5% conversion. The reactor conversion, electricity cost, catalyst cost, and plant uptime were varied  $\pm 20\%$  and the effects on the LCOA were recorded. Figure 4 presents the change in LCOA as a result of the sensitivity analysis showing that the LCOA is most sensitive to changes in the reactor conversion. A 20% decrease in conversion increased the LCOA by 47%, while a 20% increase in conversion decreased LCOA by 25.5%. A  $\pm 20\%$  change in electricity costs resulted in a proportional  $\pm 11\%$  change in LCOA. Finally, a  $\pm 20\%$  change in catalyst cost and uptime had little effect on the LCOA.





**Figure 4:** Percent change in Levelized Cost of Ammonia (LCOA) in response to a  $\pm 20\%$  change in various model parameters: Tornado Diagram Analysis.

Figure 5 shows the energy demand of the Ni-BaH<sub>2</sub> CLAS configurations compared to the HB system. Figure 5 shows that the CLAS systems can reduce energy demand by 86% compared to the HB process, when a six-reactor configuration is used. However, CLAS systems with low conversion (< 1%) can consume more energy than the HB process. Low conversion systems have a higher recycling rate which requires more energy to operate their compressors. On the other hand, CLAS systems with higher conversion have reduced recycling needs and consequently lower energy demands. Nevertheless, CLAS is a more energy-efficient method for ammonia synthesis as the majority of configurations studied are more energy efficient than the HB system.



**Figure 5.** Energy required to produce 585 tons NH<sub>3</sub> per day from the Ni-BaH<sub>2</sub> CLAS systems compared to the HB system.

## DISCUSSION

This study demonstrates the significant potential of chemical looping ammonia synthesis reactors in reducing ammonia production costs. By operating at a lower pressure (5 atm) compared to the Haber Bosch (HB) reactor (200 atm), CLAS can decrease the synthesis loop cost from \$63.9/ton to \$9.95/ton at a scale of 585 tNH<sub>3</sub>/day. This study shows that the cost reduction in CLAS systems is largely due to lower compressor costs, a major cost factor in the HB synthesis loop. However, these savings are not just due to reduced pressure but also because CLAS operates at lower temperatures, eliminating the need for costly external heating. In contrast to the adiabatic HB reactor, CLAS requires active temperature control. Therefore, focusing on mild temperature ( $T < 300^{\circ}\text{C}$ ), low-pressure CLAS systems could significantly lower ammonia production costs compared to the HB process. This reduction in synthesis cost is achievable today using a Ni-BaH<sub>2</sub> CLAS with reported production values. Among the studied systems, the Ni-BaH<sub>2</sub> CLAS had the lowest LCOA due to its high production rate (28 mmol/g<sub>cat</sub>-h) and low temperature and pressure operation. A three-reactor system using Ni-BaH<sub>2</sub> had a LCOA of \$9.95/ton while the six-reactor system had a LCOA of \$9.73/ton. The additional reactors did not significantly improve the LCOA beyond the three-reactor system.

Despite its low TRL, the equipment similarities between CLAS and HB systems enable early economic evaluations of CLAS. This research highlights contrasting cost dynamics between the equipment in a HB system and a CLAS system. While the HB system has low catalyst costs and high reactor vessel costs, the CLAS system has low reactor vessel costs and high catalyst costs. These cost differences effectively balance each other out when the total reactor costs are considered. In the HB system, the compressor costs are high because of the elevated pressure, whereas the compressor costs in the CLAS system is driven by the high recycle rate. Notably, if reactor conversion in CLAS is too low ( $\leq 1\%$ ), its compressor costs can surpass those of the HB system. CLAS need an external heater, which is an important consideration to factor into future CLAS developments. In this work external heat was provided by an electrically driven heater, alternative renewable heating sources could be studied to reduce the heating costs.

This early economic evaluation sheds light on the economic potential of CLAS systems in development. However, there are three limitations in this work. Firstly, the catalyst production rate at 1 atm is assumed to be the same at 5 atm. CLAS production rates have not been extensively studied at elevated pressures therefore the reported production rate at 1 atm is used. Presumably, elevated pressures would increase the production rate of the CLAS improving the performance of the CLAS studied

and reducing their LCOA. Secondly, the analysis assumes the catalyst can last for the lifetime of the plant. There's a lack of long-term data on catalyst longevity and methods for their regeneration in CLAS systems. Shorter regeneration cycles and rising catalyst costs could significantly impact the levelized cost of ammonia by necessitating frequent expensive catalyst replacement. Lastly, the equipment sizing assumes that nitrogen and hydrogen are present in each equipment simultaneously. This is not accurate because CLAS alternate gas flows, which results in the CLAS equipment being oversized. While this is not optimal, it does provide a margin of contingency in relation to the CLAS cost estimates.

Future research into CLAS should focus on two key areas: the development of low-temperature CLAS systems and the exploration of strategies to maintain catalyst activity. The emphasis on low-temperature CLAS is crucial due to the economic challenges posed by high-temperature systems, which require substantial external heating. On the other hand, given the significant impact of catalyst costs on CLAS economics, finding ways to prolong catalyst life is essential. This involves conducting long-term tests to establish replacement timelines and understand degradation processes. Additionally, it is important to develop environmentally friendly methods for catalyst regeneration.

## CONCLUSION

This study evaluated the technoeconomic potential of chemical looping ammonia synthesis (CLAS) reactors compared to a state-of-the-art Haber Bosch loop using Aspen Plus simulations. Three CLAS configurations were studied, two high temperature CLAS using preliminary catalyst performance data and one low-temperature CLAS using Ni-BaH<sub>2</sub> as a catalyst. Studies were performed focusing on the number of reactors in series, their conversion, and a sensitivity analysis on operating variables. The chemical loop with the lowest LCOA was the low temperature Ni-BaH<sub>2</sub> chemical loop, due to its efficient operation at low temperature and pressure. The Ni-BaH<sub>2</sub> chemical loop could reduce the LCOA by 90% when compared to the HB process, when 6 reactors were simulated in series. The study emphasizes the significance of operating CLAS at moderate temperatures to maximize cost savings because the need for an external heater can diminish the cost benefits of low-pressure operation, as was seen in the high temperature chemical loops. Despite CLAS's low TRL, its integration of existing technologies like compressors, heat exchangers, heaters, and fixed bed reactors make it possible to evaluate their economic viability even at this early stage. Future research should focus on the long-term catalyst degradation and regeneration, and performance at varied pressures to optimize CLAS further. This study underscores the need for

ongoing research to make ammonia synthesis more cost-effective and sustainable.

## ACKNOWLEDGEMENTS

This work was partially sponsored by the Pratt & Whitney Institute for Advanced Systems Engineering of the University of Connecticut. Any opinions expressed herein are those of the authors and do not represent those of the sponsor.

## REFERENCES

1. Rouwenhorst, K., Castellanos, G., IRENA., & AEA. Innovation Outlook : Renewable Ammonia (2022)
2. Smith, C., Hill, A. K., & Torrente-Murciano, L. Current and future role of Haber–Bosch ammonia in a carbon-free energy landscape. *Energy Environ. Sci.*, 13(2), 331–344. (2020). <https://doi.org/10.1039/C9EE02873K>
3. Rouwenhorst, K. H. R., van der Ham, A. G. J., Mul, G., & Kersten, S. R. A. Islanded ammonia power systems: Technology review & conceptual process design. *Renewable Sustainable Energy Rev.*, 114 (2019). <https://doi.org/10.1016/j.rser.2019.109339>
4. Burrows, L., & Bollas, G. M. (2022). Stability Assessment of Small-Scale Distributed Ammonia Production Systems. *Ind. Eng. Chem. Res.*, 61(43), 16081–16092. <https://doi.org/10.1021/acs.iecr.2c00631>
5. Burrows, L., et. al. Thermodynamic feasibility analysis of distributed chemical looping ammonia synthesis. *J. Chem. Eng.*, 426, 131421 (2021). <https://doi.org/10.1016/j.cej.2021.131421>
6. Burrows, L., & Bollas, G.M. Methods and Apparatus for Ammonia Synthesis. U.S. Patent Application. 63/443,876. Washington, DC: U.S. Patent and Trademark Office. (2023).
7. Gao, W., et. al. Production of ammonia via a chemical looping process based on metal imides as nitrogen carriers. *Nat. Energy*, 3(12), 1067–1075 (2018).
8. Turton R., et.al. Analysis synthesis and design of chemical processes (Fourth edition International). Pearson Education International. (2012).
9. Chen, C., & M. Bollas, G. Design and Scheduling of Semibatch Chemical-Looping Reactors. *Ind. Eng. Chem. Res.*, 59(15), 6994–7006 (2020). <https://doi.org/10.1021/acs.iecr.9b05693>
10. Morud, J. C., & Skogestad, S. Analysis of Instability in an Industrial Ammonia Reactor. *AIChE Journal*, 44(4), 888–895 (1998). <https://doi.org/10.1002/aic.690440414>
11. Palys, M. J., et.al. Modeling and optimal design of absorbent enhanced ammonia synthesis.

Processes, 6(7) (2018).

<https://doi.org/10.3390/PR6070091>

12. Chemical Engineering Plant Cost Index Archive:  
<https://www.chemengonline.com/site/plant-cost-index/> (Accessed Nov. 2023)
13. London metals exchange:  
<https://www.lme.com/en/Metals/Non-ferrous/>  
(Accessed Nov. 2023)
14. Shanghai Metals Market Information & Technology Co., Ltd. (SMM) <https://metal.com/> (Accessed Nov. 2023)
15. Kresse, R., et.al. (2007). Barium and Barium Compounds. In Ullmann's Encyclopedia of Industrial Chemistry, (Ed.).  
[https://doi.org/10.1002/14356007.a03\\_325.pub2](https://doi.org/10.1002/14356007.a03_325.pub2)
16. Nayak-Luke, R., et.al. "Green" Ammonia: Impact of Renewable Energy Intermittency on Plant Sizing and Levelized Cost of Ammonia. *Industrial & Engineering Chemistry Research*, 57(43), 14607–14616 (2018).  
<https://doi.org/10.1021/acs.iecr.8b02447>

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Optimal Design and Control of Behind-the-Meter Resources for Retail Buildings with EV Fast Charging

Gustavo Campos<sup>a,\*</sup>, Roberto Vercellino<sup>a</sup>, Darice Guittet<sup>a</sup>, Margaret Mann<sup>a</sup>

<sup>a</sup> National Renewable Energy Laboratory (NREL), Golden, CO, USA

\* Corresponding Author: [gustavo.campos@nrel.gov](mailto:gustavo.campos@nrel.gov)

## ABSTRACT

The growing electrification of buildings and vehicles, while a natural step towards achieving global decarbonization, poses some challenges for the electric grid in terms of power consumption. One way of addressing them is by deploying onsite, behind-the-meter resources (BTMR), such as battery energy storage and solar PV generation. The optimal design of these systems, however, is a demanding task that depends on the integration of multiple complex subsystems. In this work, the optimal integrated design and dispatch of BTMR systems for retail buildings with electric vehicle fast charging stations is addressed. A framework is proposed, combining high-fidelity simulation (of buildings, electric vehicle fast charging stations, and BTMR), predictive control strategies with closed-loop implementation, and a derivative-free design method that explores parallelization and high-performance computing. Focus is given to the design layer, highlighting the effect of parallelization on the choice of the method, computational effort, and types of results. A case study of a big-box grocery store with an EV fast charging station is presented, and its optimal BTMR system is identified in terms of equipment sizes, costs (capital, utility, lifecycle, and levelized) and resiliency against outages, demonstrating great potential for real-world applications.

**Keywords:** Battery Energy Storage, Distributed Generation, Electric Vehicle Fast Charging, Model Predictive Control, Derivative-free Optimization.

## INTRODUCTION

Electrification is held as one of the fundamental pillars for achieving widespread decarbonization, especially for historically fuel-dominated sectors such as transportation (now the largest contributor to greenhouse gas emissions in the United States [1]). This trend, however, comes with its own challenges. Widespread electrification will generate loads not seen in the past – a single electric vehicle (EV) fast charger can have a comparable load to an entire commercial building or hundreds of single-home residences, while a fast charging station can overshadow its building counterpart [2]. These loads, if not addressed adequately, will incur hefty penalties in the form of utility charges and power distribution upgrades.

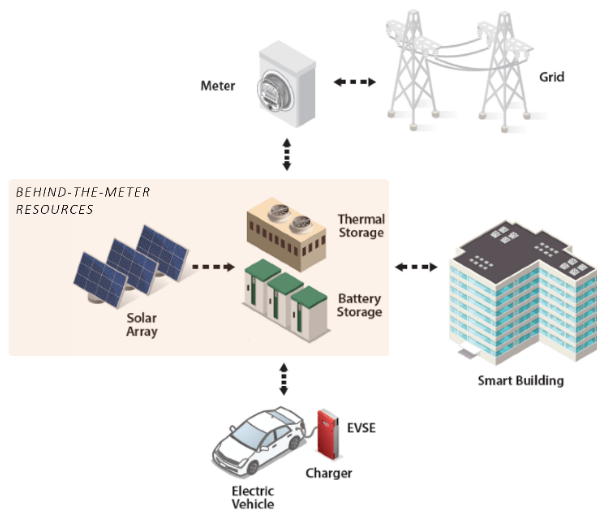
One way of reducing the peak demand and overall energy consumption of these systems is by employing onsite, distributed, behind-the-meter (BTM) resources (BTMR), such as energy storage and generation – battery storage and solar PV generation being the most notable

examples – as shown in Figure 1. This allows the system to generate renewable solar energy during the day and store it for later when either demand or prices are high (reducing both energy and demand charges), as well as to reduce the peak demand by charging the battery in advance and using it instead of the grid to meet the system's demand (reducing demand charges).

The design of BTMR systems, however, is a challenging task integrating multiple time scales, complex subsystems and application-specific parameters (e.g., expected building and EV charging loads, climate data, utility tariff, and power distribution capacity). The most cost-effective design of such systems also depends on their day-to-day operation in response to all these factors over long time periods. Moreover, the high variability of loads such as for EV fast charging requires a high-resolution time step to accurately predict system sizes.

Several tools and works have addressed this problem in the past. REopt [4], DER-CAM [5], and DER-VET [6] are examples of free software tools, while HOMER and

Energy Toolbase are examples of commercial software. The majority of these tools employ a structured, mixed-integer linear programming (MILP) optimization framework that solves an integrated design and dispatch problem. This allows the representation of a large variety of technologies [4], buildings and microgrid configurations [5, 6], but hinder accuracy. The building (and EV charging) loads are typically treated as exogeneous inputs, obtained from historical records or pre-computed hypothetical profiles. Detailed simulation models, essential for capturing equipment performance and financial mechanisms and validating the solutions generated by the simplified dispatching models, are not supported. The closed-loop implementation of the control solutions is also a critical missing component, overlooking forecast errors and their associated limitations (the projection of peak demand charges over the following month being one of the biggest challenges). Lastly, 1 minute time discretization, required to capture EV fast charging dynamics, are difficult to implement (using a year-long horizon, necessary for solar generation) without facing scalability issues.

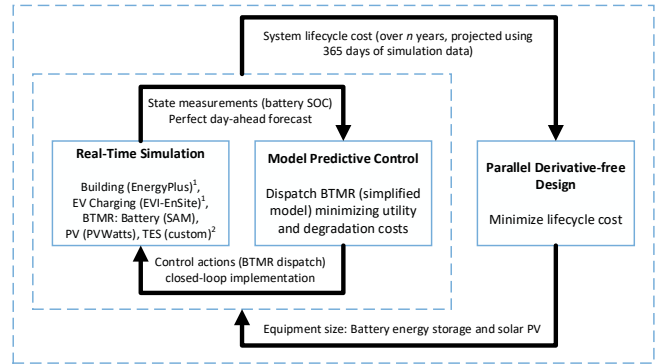


**Figure 1.** Diagram of behind-the-meter resources. (adapted from [3])

To address these limitations, a novel framework is proposed for simulation-based design of BTMR. This framework is used as the basis for the NREL tool EVI-EDGES (Electric Vehicle Infrastructure - Enabling Distributed Generation and Energy Storage) [7].

## METHODOLOGY

An overview of the methodology framework is presented in Figure 2. Each individual component is described in the following subsections.



<sup>1</sup>Electrical loads can be pre-computed depending on the level of detail required  
<sup>2</sup>Not included in this work

**Figure 2.** Methodology for sizing behind-the-meter resources.

## Simulation

The simulation environment is composed by three main components: building, electric vehicle charging station, and BTMR (battery energy storage, PV, thermal storage). In its simplest form, the first two models can be pre-solved before running the framework (for instance, when solving the base case without BTMR), while the third group of models needs to be solved in real-time with the control layer. When thermal storage (either active or passive) is present (in other words, when the operation of building cooling and heating equipment can be shifted in time), the building load depends on the BTMR dispatch and can no longer be detached and pre-simulated [8]. Thus, ideally, all the models would be integrated and solved simultaneously, allowing the consideration of interaction effects between them (e.g., battery thermal gain affecting the building indoor temperature, or EV charging schedule affected by the buildings loads).

## Building

The building thermal and electrical loads are generated using EnergyPlus (U.S. Department of Energy, 2023), a whole building energy simulation program with a rich database of construction material properties that allows the accurate representation of interactions between building zones and environment, and its resulting energy consumption (e.g., heating, cooling, ventilation, lighting, and plug loads). In its most general form, the framework allows passing operating setpoints from the control block to the building simulation, scheduling the operation of electrically-driven HVAC equipment such as chiller or heat pumps, in turn affecting the building's electrical load. The battery is placed inside of one of the building's zones with temperature control, modeling thermal interaction that determines zone cooling loads as well as battery thermal degradation and effects.

## Electric Vehicle Charging Station

The EV fast charging station is simulated using an



agent-based model (NREL's EVI-EnSite [9]). Input data includes probability distributions of vehicle arrival time and arrival state-of-charge (SOC), the number of vehicle arrivals per day or week, as well as general vehicle type-related characteristics, such as battery size and charge acceptance curves (power x SOC). The vehicle charging process is simulated by assigning arriving vehicles to unoccupied charging ports following a certain metric (usually either first-in-first-out, or a managed-charging schedule based on cost minimization). In case all charging ports are occupied, the arriving vehicles are queued and must wait in order to charge. Typically, the charging station design (i.e., number of charging ports and charging port power level) is performed to minimize the queuing time distribution (mean or a certain percentile), the latter constructed empirically by repeatedly sampling the input distributions in a Monte-Carlo fashion.

### BTMR: Battery Energy Storage and PV

A detailed technoeconomic BTM battery storage model from NREL's System Advisor Model (SAM) [10] is employed to simulate the real system's performance (as opposed to the simplified battery model inside of the MPC formulation). This model is able to represent different battery types (lithium-ion, lead-acid, redox flow) and multiple chemistries of lithium-ion batteries (LMO, LTO, LCO, LFP, NMC, NCA) with associated typical voltage curves, thermal behavior, and lifetime degradation profiles.

The solar PV panels are modeled using NREL's PVWatts model [11, 12], which considers a grid-connected photovoltaic system with modules/panels composed by crystalline silicon or thin film photovoltaic cells. Inputs related to the system's physical characteristics include power capacity, module and array types, losses, array orientation (angles) and mounting type, as well as a few built-in module and inverter characteristics.

### Optimal Dispatch/Control

The battery dispatch/control problem is formulated as a Linear Programming (LP) model (Nonlinear Programming depending on the type of battery degradation considered [13], or Mixed-Integer Nonlinear Programming if thermal storage is included [8]). This problem optimizes the adjusted cost of operation of the system (utility costs and battery degradation). A condensed formulation of the LP problem is presented next:

$$\min \phi = C^E + C^D + C^R \quad (1)$$

subject to

$$C^E = \Delta t \sum_t \phi_t^E (P_t^{purch} - \gamma^{nm} P_t^{sold}) \quad (2)$$

$$C^D = \sum_d \phi_d^D \bar{P}_d^{max} \quad (3)$$

$$C^R = (E^{max} - E_{t=t_f}^h) (RCP * EPR + RCE) \quad (4)$$

$$-P_t^{load} + P_t^{pv} + (P_t^{purch} - P_t^{sold}) + (P_t^{out} - P_t^{in}) = 0 \quad (5)$$

$$E_t - E_{t-1} = \Delta t (\eta^{in} P_t^{in} - \eta^{out} P_t^{out}) \quad (6)$$

$$E_{t=t_0}^h S^{min} \leq E_t \leq E_{t=t_0}^h S^{max} \quad (7)$$

$$E_{t=t_f}^h = E^{max} - D E_{t=t_f}^{tp} \quad (8)$$

$$E_{t=t_f}^{tp} = E_{t=t_i}^{tp} + \frac{\Delta t}{2} \sum_t (P_t^{in} + P_t^{out}) \quad (9)$$

$$\bar{P}_{\bar{t}}^{purch} = \frac{\Delta t}{\Delta \bar{t}} \sum_{t \in T_{\bar{t}}} P_t^{purch} \quad (10)$$

$$\bar{P}_d^{max} \geq \bar{P}_{\bar{t} \in \bar{T}_d}^{purch}, \bar{P}_d^{max} \geq \bar{P}_d^{max,i} \quad (11)$$

where **indices** include the time step  $t \in T$  (initial and final time steps  $t_i$  and  $t_f$ ); 15-min average time step  $\bar{t} \in \bar{T}$ ; and demand charge periods  $d \in D$ . **Special sets** include  $T_{\bar{t}}$  (all time steps  $t$  inside of each 15-min average time step  $\bar{t}$ ); and  $\bar{T}_d$  (all average time steps  $\bar{t}$  inside of each demand period  $d$ ).

**Parameters** include the difference between time steps  $\Delta t$  and  $\Delta \bar{t}$  (hours); battery installed cost per power  $ICP$  (\$/kW) and per energy  $ICE$  (\$/kWh); battery replacement cost per power  $RCP$  (\$/kW) and per energy  $RCE$  (\$/kWh); time-of-use energy prices  $\phi_t^E$  and demand prices  $\phi_d^D$ ; net-metering factor (representing the lesser value of selling power to the grid as opposed to purchasing it)  $\gamma^{nm}$  (decimal); battery charging and discharging efficiencies  $\eta^{in}$  and  $\eta^{out}$  (decimal), battery degradation per energy throughput  $D$  (decimal/kWh); energy to power ratio  $EPR$  (kWh/kW) (from design layer); minimum and maximum state-of-charge  $S^{min}$  and  $S^{max}$  (decimal); and the 15-min peak power demand from the previous day for each demand charge period  $\bar{P}_d^{max,i}$  (kW). **Forecasted parameters** include the system electrical load (sum of building and EV charging)  $P_t^{load}$  (kW), and PV generation  $P_t^{pv}$  (kW).

**Variables** include the battery power charged and discharged  $P_t^{in}$  and  $P_t^{out}$  (kW); the battery energy  $E_t$  (kWh), throughput  $E_t^{tp}$  (kWh), and health  $E_t^h$  (decimal); power purchased from and sold to the grid,  $P_t^{purch}$  and  $P_t^{sold}$  (kW); 15-min average purchased power  $\bar{P}_{\bar{t}}^{purch}$  and maximum purchased power in each demand period  $\bar{P}_d^{max}$  (for demand charge calculations). All variables are non-negative reals.

The **objective function** (1) minimizes energy cost  $C^E$  (\$), demand cost  $C^D$  (\$), and battery replacement cost  $C^R$  (\$). **Constraints** include cost calculations (2-4); power balance (5); battery energy storage difference equation (6); battery energy bounds based on its health (7); battery health at the end of the control window as a function of its energy throughput (8); battery energy throughput definition (9); 15-min averaging of power purchased from the grid (10); and 15-min peak demand power calculation (11).

A control horizon of 1 day is used (typical for short-duration energy storage problems with diurnal patterns), with a recalculation frequency of 1 day and a time step discretization of 1 minute to accurately capture the high variability from EV charging loads (battery and power distribution must be sized with the instantaneous load instead of averages). After each horizon is solved, a time series of control setpoints is provided to the simulation for evaluating the response of the subsystems and updating the stationary battery health. Then the final system state from the simulation is set as initial point of the successive control problem.

## Optimal Design

Under the formulation described above, the overall design problem can be posed as a nested optimization problem, in which each evaluation of the design layer depends on the solution of a lower-level layer consisting of several coupled control optimization subproblems, further complicated by their closed-loop implementation in the simulation environment. The objective function (system lifecycle cost) is also time-consuming to evaluate (in the order of hours), which makes approximating the derivatives via finite-difference methods an ineffective approach. For these reasons, only derivative-free (black-box) optimization methods were considered [14, 15].

These methods are generally classified according to the properties of the black-box oracle function [15], which in the present case is assumed to be deterministic, generally nonconvex and non-smooth, multimodal (although empirically observed to be unimodal), unconstrained, with bounded feasible region (decision variables are typically nonnegative real variables, with upper bounds determined from space and weight constraints or heuristics), and time-consuming evaluation (hours). The latter is a key characteristic that prohibits or hinders the application of direct-search methods such as Nelder-Mead (with a low number of simplex vertices) or trust-region methods, which require several iterations to converge.

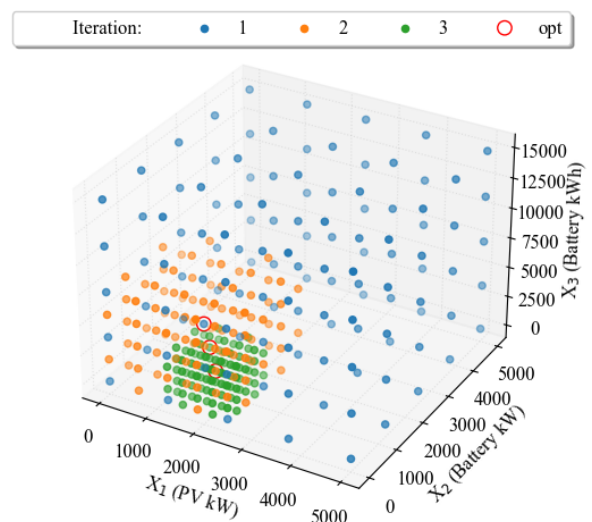
The ability to parallelize (concurrent) function evaluations is considered as a critical feature to determine the choice of method. Most optimization methods tend to minimize the total number of function evaluations over the entire algorithm, as opposed to minimizing the time required to solve the problem (typically proportional to each other). When each function evaluation is time consuming, however, it is more efficient to parallelize as many concurrent evaluations as possible, then use the ensemble information to perform the search.

## Sequential Grid Search Method

A sequential grid search method is employed, which samples the decision space using symmetric hypercubes that begin by covering the whole feasible space and shrink around the observed optimum after each iteration.

In other words, after each iteration, the optimum point is identified between two (if the solution lies at the bounds) or three (if it falls in the middle) grid points in each direction, which are set as the new bounds for the reduced grid at the next iteration (previously evaluated points are reused for the new grid to reduce effort). The number of function evaluations for each iteration depends on the number of discrete points in each of the grid's independent coordinates (variables). For a problem with three decision variables, the number of evaluations will then be equal to the number of points to the power of three. This points out the major deficiency of this approach, which is its scalability with the number of decision variables. While it is efficient for this problem setup, it becomes more challenging when including other design variables (e.g., TES or charging station). The method is illustrated in Figure 3, showing the evolution of the grid and optimal point throughout subsequent iterations.

After each iteration, a surrogate model can be fitted and the next-iteration grid can be generated around the surrogate optimum instead of the sampled one. However, the benefit of this approach is not guaranteed, as it depends on how well the surrogate represents the real underlying function. If the surrogate model does not capture the function appropriately, it can mislead the search and affect it negatively, adding iterations and increasing the runtime. It was empirically observed that low order polynomial models do not add a significant value in predicting the true optimum location and were thus not employed.



**Figure 3.** Illustration of a 3-dimensional sequential grid search with 5 points in each direction and 3 iterations. used to interpolate the optimum location.

## Accuracy and Computational Effort

The upper bound on the final accuracy level  $\delta x_{n_i}$  (difference between two evaluated points) after  $n_i$

iterations, using  $n_p$  points in each direction (assuming the same number of points for all directions), and starting with an initial range  $\Delta x_0$ , is given in (12). Solutions obtained at the grid boundaries at any iteration will have a better accuracy due to its one-sided shrinkage (as opposed to two-sided shrinkage necessary for non-boundary solutions).

$$\delta x_{n_i} = \frac{(2^{n_i-1} \Delta x_0)}{(n_p-1)^{n_i}} \quad (12)$$

Two ways of refining the solution accuracy consist in increasing the number of evaluation points ( $n_p$ ) or number of iterations ( $n_i$ ). To guide the parameter tuning decision, the final solution accuracy (relative to the initial range  $\Delta x_0$ ) and computational effort (in terms of core-hours, a typical measure for HPC systems) are presented in Table 1. It is assumed that each design evaluation point (one year simulation with closed-loop optimal control) has a runtime of approximately 2 hours, and that the number of cores equals the number of evaluated design points. Cells with final solution accuracy under 1% of the initial range are highlighted in bold in both tables.

**Table 1:** Solution accuracy and computational effort over varying number of grid points (cols) and iterations (rows).

$n_p$	4	5	6	7	8	9	10	$h$
$n_i$	<b>Final Solution Accuracy (Upper Bound on Error Margin)</b>							
1	33.3%	25.0%	20.0%	16.7%	14.3%	12.5%	11.1%	2
2	22.2%	12.5%	8.0%	5.6%	4.1%	3.1%	2.5%	4
3	14.8%	6.3%	3.2%	1.9%	1.2%	<b>0.8%</b>	<b>0.5%</b>	6
4	9.9%	3.1%	1.3%	<b>0.6%</b>	<b>0.3%</b>	0.2%	0.1%	8
5	6.6%	1.6%	<b>0.5%</b>	0.2%	0.1%	0.0%	0.0%	10
6	4.4%	<b>0.8%</b>	0.2%	0.1%	0.0%	0.0%	0.0%	12
7	2.9%	0.4%	0.1%	0.0%	0.0%	0.0%	0.0%	14
8	2.0%	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%	16
9	1.3%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	18
10	<b>0.9%</b>	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	20
	<b>Computational Effort (Core-Hours)</b>							
1	128	250	432	686	1024	1458	2000	2
2	256	500	864	1372	2048	2916	4000	4
3	384	750	1296	2058	3072	<b>4374</b>	<b>6000</b>	6
4	512	1000	1728	<b>2744</b>	<b>4096</b>	5832	8000	8
5	640	1250	<b>2160</b>	3430	5120	7290	10000	10
6	768	<b>1500</b>	2592	4116	6144	8748	12000	12
7	896	1750	3024	4802	7168	10206	14000	14
8	1024	2000	3456	5488	8192	11664	16000	16
9	1152	2250	3888	6174	9216	13122	18000	18
10	<b>1280</b>	2500	4320	6860	10240	14580	20000	20

For the presented setup, the number of grid points should be equal or greater to 5 (using 4 points requires at least 10 iterations, which would entail 20 hours of runtime), while the appropriate number of iterations varies between 3 and 6 (runtime of 6 and 12 hours). The bottom table shows that computational effort increases more rapidly with the number of grid points (total number of evaluations/cores per iteration is given by  $n_p^3$ ) than iterations, although the latter increases runtime linearly by 2 hours. The final setup (grid coarseness and number of iterations) should thus be a function of several factors, including required solution time, number of parallel cores

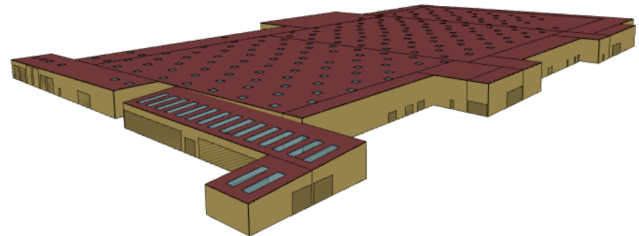
available, and acceptable solution accuracy.

In summary, a few conditions that encourage the application of this method include time-consuming objective function evaluation (in the order of hours); small number of design variables (up to 3 or 4); high parallelization (100s of concurrent function evaluations); bounded feasible region; and an acceptable final solution accuracy within 1% of initial range.

## CASE STUDY

### Building Model

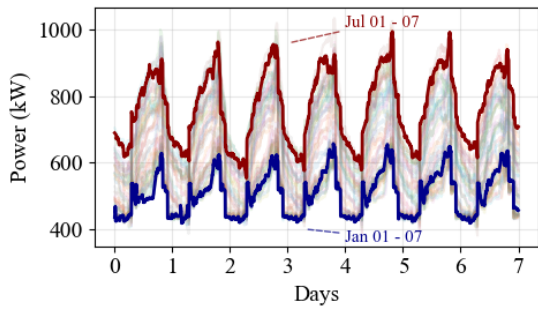
A big-box grocery store building energy model, developed and validated in [2], was employed in this work. This represents a common type of commercial building in the United States, at which EV fast charging stations are expected to be increasingly deployed in the near future – fast charging stations in retail buildings will attract customers and bring additional revenue, increasing business competitiveness, while also providing charging access to disadvantaged areas with lacking access to home charging. The building model has 18 thermal zones, 136 surfaces, and 263 sub-surfaces (e.g., windows, doors), and an area of approximately 20,000 m<sup>2</sup>. Figure 4 presents the 3D model of the building, with different zones roughly separated by solid roof lines, while the electrical load generated by the building (from HVAC and plug loads) is presented in Figure 5.



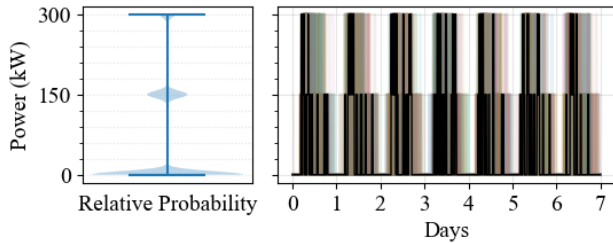
**Figure 4.** 3D model of the generic big-box grocery store from OpenStudio.

### Electric Vehicle Charging Loads

A fast charging station with two 150 kW charging ports and 6 events per port per day was simulated. The load profile is presented in Figure 6. Default inputs for commonly used EVs were employed (vehicle battery size, charge acceptance curves), while the probability distributions (vehicle arrival time and SOC) were correlated to the building occupation patterns. The load has a high variability (spikiness), which may not significantly increase the total energy consumption, but will likely result in a higher peak demand, affecting the sizes of both the power distribution system and energy storage, as well as utility demand charges.



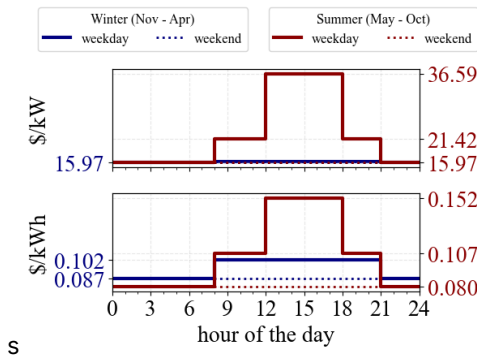
**Figure 5.** Building electrical load over a week for all weeks in a year.



**Figure 6.** Violin plot (left) and weekly profile (right) of EV fast charging electrical load (one week shown as the dark line, remaining weeks in the background).

### Utility Rate Structure

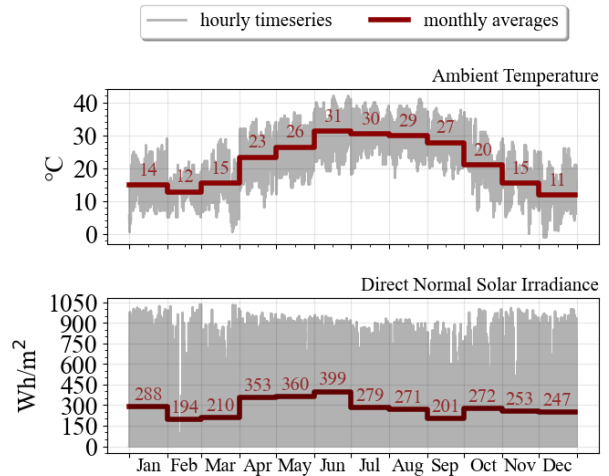
A utility rate tariff structure based on a high-price region (PG&E in Northern California) is used. The demand and energy prices are shown in Figure 7.



**Figure 7.** Utility rate tariff: demand (top) and energy (bottom) prices.

### Location/Weather Inputs

The most relevant location-dependent weather parameters are presented in Figure 8, namely, ambient temperature (which drives the building energy consumption and equipment efficiency) and solar irradiance (which drives building heat gains and solar PV power generation). The city of Tucson, AZ, was chosen as the location for the study.



**Figure 8.** Weather inputs.

### Remaining Inputs

A summary of the remaining inputs and settings is presented in Table 2.

**Table 2:** Remaining inputs.

Input	Value
Life-cycle Analysis Horizon	20 years
Discount Rate	8.3%
Battery Installed Cost	\$540/kW + \$120/kWh
PV Installed Cost	\$1600/kW
PV initial range	0 – 2,000 kW
Battery initial range	0 – 5,000 kW 0 – 10,000 kWh
Forecast Method	Perfect
PV annual degradation	0.5%
Battery usable SOC range	10 – 90%
Battery Chemistry	LMO-LTO
Battery cost de-escalation rate for replacement	6%/year
Battery replacement capacity threshold	60%
MPC time horizon	24 h
MPC time step	1 min
MPC demand charge time averaging	15 min

## RESULTS

The problem was solved using NREL's high-performance computing system Kestrel with dual socket Intel Xeon Sapphire Rapids (52-core) processors, 104 cores per CPU node, and 256 GB DDR5 memory. The sequential grid method was applied with 5 points and 6 iterations, achieving an accuracy upper bound of 0.8% of the initial range (or lower).



## Design Solution: System Sizes and Costs

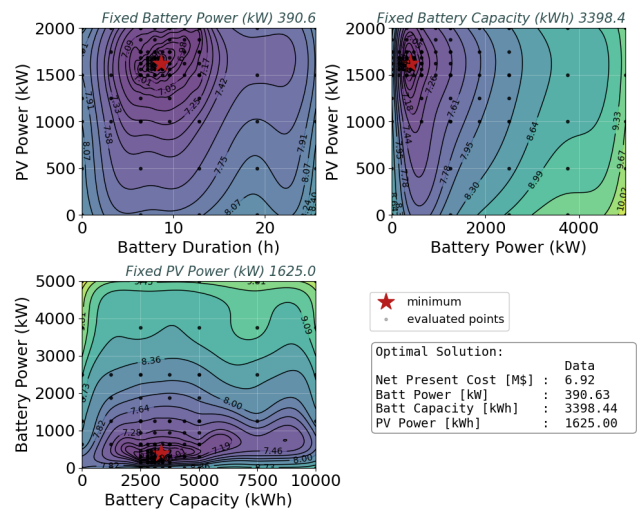
The main results of the analysis are presented in Table 3 for the base case (without BTMR), and the BTMR case (both cases with the EV charging station). The identified BTMR system consists of a battery with 390.6 kW capacity and 8.7 hours of duration, and a PV capacity of 1625 kW, highlighting the potential for PV deployment in locations with high solar incidence. Their associated capital costs follow a similar trend, being dominated by the large PV deployment and its higher unit cost. In this setting, the EV supply equipment (EVSE) capital cost is only a fraction of the BTMR system (which is not always the case, specially for commercial fleets). The benefits of BTMR deployment are shown in the metrics associated with electricity consumption and utility costs. Annual energy and peak demand are significantly reduced (by over half), in turn reducing their associated utility-related charges. The Levelized Cost of Charging (LCOC) (as defined in [13]) demonstrates the positive impact that BTMR systems can have on fast charging station deployment, decreasing from \$0.33/kWh (a typical value in the US) to -\$0.96/kWh (negative value representing the fact that the BTMR system reduces the total operating cost more than the EVSE increases it). Finally, the lifecycle net-present cost (adjusted capital and operating cost) indicates that the BTMR system increases profit gains by \$1.41 million over the system's lifetime.

## Sensitivity Analysis

One of the benefits of evaluating a grid of points (as opposed to a narrow search employed by most derivative-based methods) is that the sensitivity of the objective with respect to the decision variables can be evaluated. This allows the decision maker to identify design solutions with low sensitivity (in which adjusting a design variable does not significantly affect the objective function), adding more flexibility and confidence to the final design solution. Figure 9 shows contour plots of net-present cost with respect to two decision variables at a time (fixing the remaining one at its optimal value). Black dots represent points evaluated during the design procedure, while contour lines are generated performing interpolation using cubic radial-basis functions (RBF). The top-left plot shows a balanced effect between the two variables, with nonconvexities arising for large battery duration values. The top-right plot indicates that the battery power affects the design solution more significantly than PV size (contour lines are closer to each other). A similar trend is observed in the bottom-left plot, where battery power seems to affect the objective more significantly than battery capacity/duration. The ripples and non-convexities shown in the contour lines (more pronounced in the top-left plot) indicate a possibly low quality of the interpolation fit. One way of improving it would be to use a larger number of grid points in the design method.

**Table 3:** Overview of results: system sizes and costs.

Variable	Base Case	BTMR
<i>System Sizes</i>		
Battery Power	--	390.6 kW
Battery Storage (Duration)	--	3398.44 kWh (8.7 h)
PV Power	--	1625.0 kW
<i>Capital Costs</i>		
Battery Capital Cost	--	\$618,750
PV Capital Cost	--	\$2,600,000
EVSE Capital Cost	\$308,400	\$308,400
<i>Electricity Consumption</i>		
Energy from Grid	5,392 MWh/year	2,280 MWh/year
Monthly Peak Demand	1031.81 kW max 891.65 kW mean	695.39 kW max 579.62 kW mean
<i>Utility Costs</i>		
Energy charges	\$527,449/year	\$193,558/year
Demand charges	\$321,538/year	\$147,930/year
<i>Levelized Costs</i>		
EV Charging	\$0.33/kWh	-\$0.96/kWh
Energy (Electricity)	\$0.16/kWh	\$0.13/kWh
<i>Lifecycle Costs</i>		
Net Present Cost	\$8.33 M	\$6.92 M
Net Present Value	--	\$1.41 M



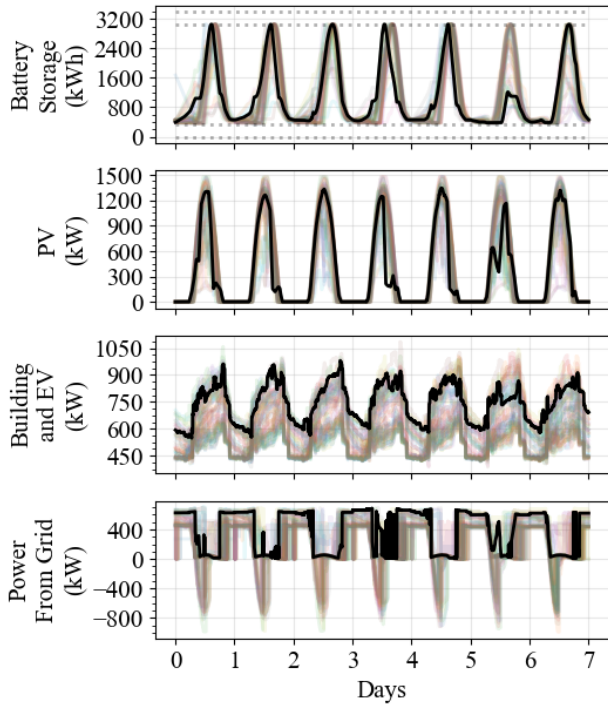
**Figure 9.** Contour plots of net present cost with respect to system design variables.

## Dispatch and Load Profiles

The storage dispatch and load profiles (PV generation, building and EV loads, and power purchased from or sold to the grid) are presented in Figure 10. The first two subplots highlight the synergistic effect between battery



storage and PV generation, the former being discharged soon after the latter starts to decrease. The timing between PV generation and the building and EV charging loads also works in favor of BTMR, allowing most of the solar electricity generated onsite to be directly used to address the load. Finally, the power purchased from the grid shows a significant reduction in peak power demand (as well as energy consumption), inverting the time in which it happens from day to night, and the utilization of the net-metering mechanism to sell electricity back to the grid (despite its price being substantially lower than the purchase) during periods with solar generation surplus.

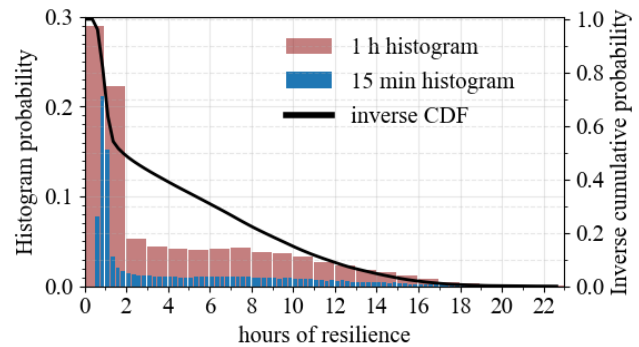


**Figure 10.** Dispatch and load profiles over all weeks in a year (week 26 highlighted as the dark solid line).

### Resiliency against power outages

The cost-optimal BTMR system sized in the previous section also has the added benefit of providing resiliency against grid interruptions. That is, whenever there is an interruption in the power grid service (average of 4 hours, 1.4 times per year in the United States [16]), the battery storage (and PV generation) can be used to meet the system's electrical load. However, since the battery operates according to a cost-optimal dispatch, there is no guarantee that it will be able to meet that load. To assess this, the battery's SOC at all time steps in the simulated year was used to estimate its resiliency potential (how many time steps into the future it would be able to meet the system's load). The result is a probability distribution of resiliency hours, presented in Figure 11. The bar plots show the resiliency histogram with two levels of

time step granularity (hourly and 15-min intervals). This graph represents the probability that the provided resiliency will be inside of each time interval (e.g., 29% between 0 and 1 hours, or 21% between 45 and 60 minutes). The solid line represents the inverse cumulative distribution function (with 15-min intervals), or the probability that the BTMR will provide over  $n$  hours of resiliency (e.g., 75% of the time the system provides over 1 hour of resiliency). The BTMR system provides over 30 minutes of resiliency 100% of the time, meaning that it can naturally address small outages without having to change its dispatching priorities, and can potentially meet up to 23 hours of future load.



**Figure 11.** Resiliency provided by the cost-optimally-sized BTMR (battery + PV).

## CONCLUSIONS

In this work, the optimal integrated design and dispatch of Behind-the-Meter Resources (BTMR) (i.e., battery energy store and PV generation) for retail commercial buildings with EV fast charging stations was addressed. A novel modeling framework was proposed, integrating detailed simulation models (of building, electric vehicle fast charging station, and BTMR equipment), model predictive control (MPC) strategies for equipment dispatch with closed-loop implementation, and a derivative-free design layer that leverages parallelization and high-performance computing. A sequential grid search method was employed to solve the design problem, balancing accuracy and computational effort. Results from a case study involving a big-box grocery store retail building model highlight the benefits of BTMR in reducing energy and peak demand charges, by generating clean electricity onsite, and shifting power purchases from the electric grid. Contour plots of lifecycle cost, generated by leveraging the parallel evaluations from the design method, indicate the design variables that have the biggest impact on the objective. Finally, the system's resiliency under grid outages was also evaluated, demonstrating a good potential for replacing fuel-based alternatives.

Ideas for future work include expanding the case

study to assess the potential of BTMR for retail buildings across multiple locations and utility rates; including a resiliency-based cost in the dispatch objective; considering the effect of forecast errors of future electrical load (from both building and EV charging station) in the control problem; and including carbon emissions in the design and dispatch optimization objectives.

## ACKNOWLEDGEMENTS

Funding provided by the U.S. Department of Energy Vehicle Technologies Office, Buildings Technologies Office, the Assistant Secretary for Energy Efficiency and Renewable Energy.

## REFERENCES

1. U.S. Environmental Protection Agency (EPA). Fast Facts on Transportation Greenhouse Gas Emissions. (2023) <https://www.epa.gov/greenvehicles/fast-facts-transportation-greenhouse-gas-emissions>
2. Gilleran M, Bonnema E, Woods J, Mishra P, Doebber I, Hunter C, Mitchell M, Mann M. Impact of electric vehicle charging on the power demand of retail buildings. *Adv. Appl. Energy* 4:100062 (2021).
3. NREL. Behind-the-Meter Storage Consortium. (N.D.) <https://www.nrel.gov/research/behind-the-meter-storage-consortium.html>
4. Ogunmodede O, Anderson K, Cutler D, Newman A. Optimizing design and dispatch of a renewable energy system. *Appl. Energy* 287:116527 (2021)
5. Mashayekh, S, Michael S, Gonçalo C, Heleno M. A mixed integer linear programming approach for optimal DER portfolio, sizing, and placement in multi-energy microgrids. *Appl. Energy*, 187:154-168 (2017)
6. EPRI. Distributed Energy Resource Value Estimation Tool (DER-VET™ v1.2), Program 94. <https://www.der-vet.com/>
7. NREL. EVI-EDGES: Electric Vehicle Infrastructure – Enabling Distributed Generation Energy Storage Model. (N.D.) <https://www.nrel.gov/transportation/evi-edges.html>
8. Guittet D, Bonnema E, Mitchell M, Mahvi A, Woods J. To freeze or not to freeze: model-predictive vs. schedule-based control of an ice-based thermal energy storage system. *Submitted*.
9. NREL. EVI-EnSite: Electric Vehicle Infrastructure – Energy Estimation and Site Optimization Tool. (N.D.) <https://www.nrel.gov/transportation/evi-ensite.html>
10. NREL. System Advisor Model Version 2022.11.29 (SAM 2022.11.21). <https://sam.nrel.gov>
11. NREL. System Advisor Model - Photovoltaic Models. (N.D.) <https://sam.nrel.gov/photovoltaic.html>
12. NREL. PVWatts Calculator. (N.D.) <https://pvwatts.nrel.gov/>
13. Guittet DL, Gasper P, Shirk M, Mitchell M, Gilleran M, Bonnema E, Smith K, Mishra P, Mann M. Levelized cost of charging of extreme fast charging with stationary LMO/LTO batteries. *J. Energy Storage* 82:110568 (2024)
14. Conn AR, Scheinberg K, Vicente LN. Introduction to Derivative-Free Optimization. Philadelphia: SIAM. (2009)
15. Larson J, Menickelly M, Wild SM. Derivative-Free Optimization Methods. Cambridge University Press. (2019)
16. U.S. Energy Information Administration (EIA). Reliability Metrics of U.S. Distribution System. (2023) [https://www.eia.gov/electricity/annual/html/epa\\_11\\_01.html](https://www.eia.gov/electricity/annual/html/epa_11_01.html)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Optimization of Retrofit Decarbonization in Oil Refineries

Sampriti Chattopadhyay<sup>a</sup>, Rahul Gandhi<sup>b</sup>, Ignacio E. Grossmann<sup>a</sup> and Ana I. Torres.<sup>a\*</sup>

<sup>a</sup> Carnegie Mellon University, Department of Chemical Engineering, Pittsburgh, PA, USA

<sup>b</sup> Shell, USA

\* Corresponding Author: [aitorres@cmu.edu](mailto:aitorres@cmu.edu).

## ABSTRACT

The chemical industry is actively pursuing energy transition and decarbonization through renewables and other decarbonization initiatives. However, navigating this transition is challenging due to uncertainties in capital investments, electricity costs, and carbon taxes. Adapting to decarbonization standards while preserving existing valuable infrastructure presents a dilemma. Early transitions may lead to inefficiencies, while delays increase the carbon footprint. This research proposes a framework to find an optimal retrofit decarbonization strategy for existing oil refineries. We start with a generic process flowsheet representing the refinery's current configuration and operations, and consider various decarbonization alternatives. Through superstructure optimization, we identify the most cost-effective retrofit strategy over the next three decades to achieve decarbonization goals. We develop a Mixed-Integer Linear Programming (MILP) model, integrating simplified process equations and logical constraints to identify the most economical retrofit decarbonization strategy. The paper presents numerical results from the MILP model. Furthermore, the trends exhibited by the outcomes across various scenarios considering distinct electricity costs and carbon tax levels are presented. These results provide valuable insights into the economic feasibility of retrofit electrification strategies for decision-makers in the chemical industry.

**Keywords:** Optimization, Process Design, Electricity & Electrical Devices, Process Operations, Renewable and Sustainable Energy

## INTRODUCTION

The 2021 UN Climate Change Conference (COP26) stressed the urgent need to reduce global greenhouse gas emissions to limit global warming to 1.5°C [25]. This goal coupled with a global increase in energy consumption has urged nations to accelerate the adoption of low-emission energy systems. There is a pressure for the chemical industry to embrace emerging low-carbon technologies [16] [8].

Decarbonization of oil refineries is challenging due to their diverse configurations and high operational efficiencies. One potential pathway for decarbonization involves incorporating low-carbon feed into refinery operations. Retrofitting a fossil-based refinery to use a biomass-based feed over a time horizon of 10 years has been studied in [30] using a Mixed-Integer Linear Programming (MILP) model. Currently, significant effort is directed towards reducing scope 1 [32] emissions by electrification of process heat & hydrogen and the use of

carbon capture and storage techniques. Other possible pathways involve the use of carbon capture technologies and/or electrification of hydrogen production and steam generation integrated within existing refinery infrastructure (see Fig. 1 (a)) which are described below.

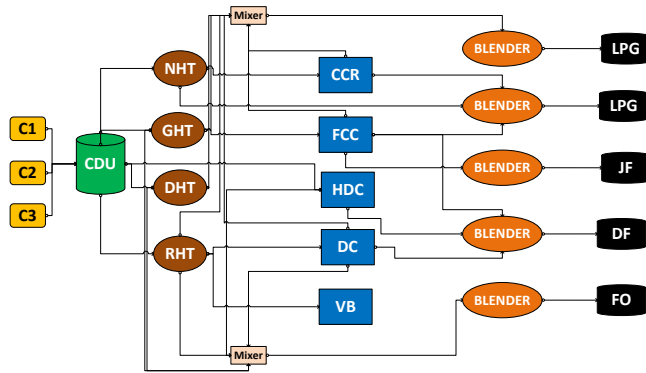
### Carbon capture (CC) Techniques

Post-combustion, pre-combustion, and oxy-combustion are the primary techniques for carbon capture (CC) [18]. Post-combustion capture suits low CO<sub>2</sub> concentration flue gas, while pre-combustion applies to gasification plants. Oxy-combustion involves burning fuel in an oxygen-rich environment. This study exclusively focuses on pre- and post-combustion capture due to their ease of retrofitting existing operations. [22, 24, 28, 12]

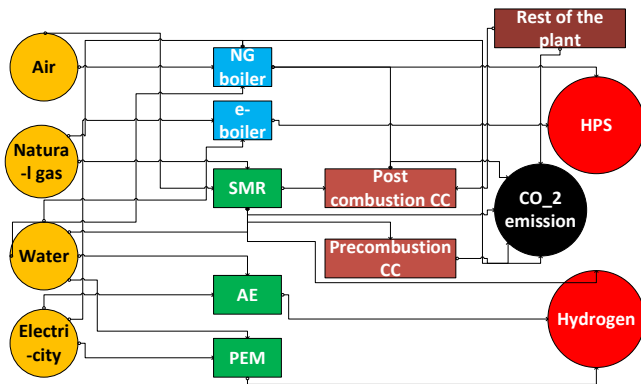
### H<sub>2</sub> production

One approach for decarbonizing H<sub>2</sub> production is through blue H<sub>2</sub> production, where SMR (Steam Methane Reforming) and WGS (Water Gas Shift) reactors are augmented with pre-combustion CC [9]. Alternatively, the flue gas

from these reactors, after extracting hydrogen through adsorption-based separation, can be mixed with flue gas from other plant sources. This mixture allows for absorption-based post-combustion CC. Another method for decarbonizing H<sub>2</sub> production involves using renewable electricity for the electrolysis of water, producing green hydrogen [21, 20, 15, 27]. We focus on the use of low-temperature electrolyzers, such as Alkaline Electrolyzers (AE) and Proton Exchange Membrane Electrolyzer (PEME), due to their level of maturity (AE), efficiency, and adaptability to handling variable operational conditions (PEME) [7].



**Fig 1(a)** : Generic Flowsheet for an oil refinery; CT denotes different types of crude oil fed to Hydrotreating units—NHT (Naphtha Hydrotreater), DHT (Distillate Hydrotreater), GHT (Gas Oil Hydrotreater), and RHT (Residue Hydrotreater)—further treated in processing Units: CCR (Continuous Catalytic Reformer), HDC (Hydrocracking), FCC (Fluid Catalytic Cracking), DC (Delayed Coking), and VB (Visbreaking); Output streams are blended in desired proportions to produce LPG (liquified petroleum gas), JF (Jet Fuel), DF (Diesel Fuel), FO (Fuel Oil); adopted from [29]



**Fig 1(b)**: Superstructure with possible pathways for hydrogen and HPS (high-pressure steam) production: considered routes for H<sub>2</sub> production include Gray H<sub>2</sub>, Blue H<sub>2</sub>, and Green H<sub>2</sub> (using PEME or AE); HPS generation alternatives encompass utilization of Natural Gas (NG) Boilers or e-boilers; Blue hydrogen can be synthesized employing either pre-combustion or post-combustion

techniques; the flue gas released from the rest of the plant is captured by post-combustion CC.

## Process heat

In the refinery, most of the heating demands are met using steam. Boilers in Fig 1 (b) generate only HPS. MPS and LPS are obtained from HPS by reducing the pressure with let-down valves. Other heating needs are fulfilled by furnaces tailored to individual unit operations in the refinery. To simplify the analysis, we assume that flue gas from these sources can be collected and treated by post-combustion CC units. Currently, steam production relies on gas boilers that burn natural gas. Electric boilers (e-boilers) are a possible greener replacement for gas boilers due to their easy installation, control, and maintenance, as well as higher thermal efficiency. e-boilers can also be easily retrofitted with renewable power sources, offering a more sustainable option for steam generation [26].

## The gap: Electrification from a systems perspective

While market-ready technologies exist and techno-economic analysis are available for individual decarbonization initiatives, there is a lack of literature on how these initiatives can be optimally combined to retrofit existing units cost-effectively. As emphasized in [33] since the heat and mass flow between units are interconnected in a chemical industry it is extremely critical to perform a systems level analysis to perform for decarbonization or electrification planning. This work develops a tool for planning a transition towards such a retrofit solution, which can cater to each specific oil refinery at a given location.

## PROBLEM STATEMENT

Existing literature discusses some market-ready technologies for alternative decarbonized technologies [17] but lacks a clear framework for understanding which ones should be chosen, when these alternative technologies should be implemented, and how they could be optimally integrated. This paper aims to fill this gap by developing a comprehensive decarbonization plan that is applied to a given oil refinery operating over 29 years (2022-2050). The plan involves strategic selection and timely implementation of the best set of initiatives for retrofitting the supply of HPS and hydrogen to the given refinery (by addition and removal of units and interconnections), considering operating data for a typical crude-based refinery with a distillation capacity of 100 kbbbl crude oil per day. The objective is to minimize the Present Value of the cost for decarbonization retrofit design, while meeting predefined environmental commitments.

## METHODOLOGY

Fig. 1 (b) shows a superstructure representing potential pathways for High-Pressure Steam (HPS) generation, hydrogen production, and CC.

The variables used in the model are,  $F_{i,j,k,yr}$ , a positive continuous variable representing the amount of component  $k$  present in a stream flowing from unit  $i$  to unit  $j$  in year  $yr$ .  $y_{i,yr}$  ( $m_{i,yr}$ ), a binary variable indicating whether the installation of a specific unit of technology  $i$  (e.g. PEME, AE) is initiated (completed) in year  $yr$ .  $z_{i,yr}$  ( $p_{i,yr}$ ), a non-negative integer variable indicating the number of initiated (completed) installations in year less than or equal to  $yr$  for units of technology  $i$ .  $Q_{i,yr}$ ; ( $QE_{i,yr}$ ), a continuous variable indicating the total (expansion in) installed capacity of a technology  $i$  in a given year  $yr$ .

### Logic Timing constraints

Eqs. (1) and (2) capture the relationships between the described binary variables.

$$\sum_{yr=1}^l y_{i,yr} = z_{i,l} \quad \forall i, l \quad (1)$$

$$\sum_{yr=1}^l m_{i,yr} = p_{i,l} \quad \forall i, l \quad (2)$$

These Eqs. are valid for all years in the range 1-29 unless specified otherwise. Furthermore, additional logical constraints are added to specify that  $z$  and  $p$  are time-lagged copies of  $y$  and  $m$  respectively depending on the duration of installation of each equipment. Flow to units of technology  $i$  is non-zero if the installation of any unit of  $i$  has been completed by a given year. This relationship can be expressed as an upper bound constraint, shown in Eq. (3):

$$F_{i,j,k,yr} \leq UB * p_{i,yr} \quad (3)$$

### Mass Balance constraints

The proposed formulation requires specifying that flow can take place only when the units are connected in the superstructure.

$$F_{i,j,k,yr} = 0 \quad \forall j \notin \text{Out}(i), \forall k, \forall yr$$

Additionally, equations specify that air, natural gas, water, and electricity are obtained from their respective supply nodes.

Mass Balance constraints given by Eqs. (4) and (5) relate the amount of components present in inlet and outlet streams to reactors and separators through simplified yield-based models. In Eq. (6),  $\mu_{in,j,k}$  represents the amount of reactant  $k$  reacting with a certain amount of reactant  $k'$  in unit  $j$ , denoted as  $\mu_{in,j,k'}$ . Similarly, in Eq. (5),  $\mu_{out,j,k}$  denotes the amount of product  $k$  in the outlet stream of unit  $j$ , given a certain amount of reactant  $k'$  in the inlet stream.

$$\sum_{i \in \text{In}(j)} F_{i,j,k,yr} = \frac{\mu_{in,j,k}}{\mu_{in,j,k'}} \sum_{i \in \text{In}(j)} F_{i,j,k',yr} \quad (4)$$

$$\sum_{i \in \text{Out}(j)} F_{j,i,k,yr} = \frac{\mu_{out,j,k}}{\mu_{in,j,k'}} \sum_{i \in \text{In}(j)} F_{i,j,k',yr} \quad (5)$$

$$k \in \{\text{Relevant Components, Electricity}\} \\ \forall i \in \{\text{Reactors, electrolyzers, boilers}\}$$

Here, the notations  $\text{In}(j)$  and  $\text{Out}(j)$  represent the set of nodes connected at the inlet and outlet of node  $j$  respectively.

For the CC units, a component-wise mass balance constraint is applied, as no reactions occur. The constraint is represented by Eq. (6)

$$\sum_{i \in \text{Out}(j)} F_{j,i,k,yr} = \sum_{i \in \text{In}(j)} F_{i,j,k,yr} \quad \forall j \in \text{CC units} \quad (6)$$

Additional constraints are imposed to guarantee the specified purity of the final streams.

### Energy Balance Constraints

Energy balances for Natural Gas (NG) boiler, e-boiler, and electrolyzers are given by Eqs. (7), (8) and (9) respectively.

$$F_{\text{NG Supply, NG Boiler, NG, yr}} LHV_{\text{NG Boiler}} \eta_{\text{NG boiler}} = F_{\text{Water Supply, NG Boiler, NG, yr}} \Delta H_{\text{vap}} \quad (7)$$

$$F_{\text{Grid, e-boiler, electricity, yr}} \eta_{\text{e-boiler}} = \Delta H_{\text{vap}} F_{\text{H}_2\text{O supply, e-boiler, H}_2\text{O, yr}} \quad (8)$$

$$F_{\text{electrolyzer, H}_2\text{ requirement, Hydrogen, yr}} \Delta H_{\text{electrolysis}} = F_{\text{grid, electrolyzer, Electricity, yr}} \eta_{\text{electrolyzer}} \quad (9)$$

$\eta$  denotes the energy efficiency of equipment and LHV is the Lower Heating Value of natural gas.

### Design Constraints

The values of the total HPS and hydrogen produced by all the different pathways must match the total hydrogen and HPS requirements of the refinery. The amount of CO<sub>2</sub> captured and emitted after passing through CC units are linearly related by a constant factor given by the efficiency of the respective CC equipment. Similarly, the steam produced and the losses through the purge stream (blowdown) from the boilers are linearly related by a constant factor.

The amount of CO<sub>2</sub> and other gases present in the flue-gas released from the rest of the operations in the refinery are specified by constraints. Decarbonization goals of the refinery impose an upper limit on the emissions as shown in Eq. (10)

$$\sum_{i \in \text{In}(\text{CO}_2 \text{ emission})} F_{i, \text{CO}_2 \text{ emission, CO}_2, yr} \leq \text{CO}_2 \text{ cap}_{yr} \quad (10)$$

Eq. (11) imposes a logical constraint on the capacity expansion in any year  $\forall i \in \text{PEM, AE, e-boiler, CC units}$ .

$$LB_1 m_{i,yr} \leq QE_{yr} \leq UB_1 m_{i,yr} \quad (11)$$



Eq. (12) evaluates the total capacity at the end of any year  $yr$ . [23, 19]

$$QE_{i,yr} + Q_{i,yr-1} = Q_{i,yr} \quad (12)$$

Eqs. (13) and (14) state that the operating flowrate or power is less than the installed capacity.

$$F_{grid,i,electricity,yr} \leq Q_{i,yr} \quad (13)$$

$$\forall i \in e - boiler, electrolyzer$$

$$F_{CC unit,CO_2 captured,CO_2,yr} \leq Q_{CC unit,yr} \quad (14)$$

We consider that the capacity expansion for e-boiler and CC facilities can take any value within a continuous range, whereas the expansion of electrolyzer capacity can only take discrete values based on available standard sizes. To handle this, we introduce binary variables  $w_{i,h,s,yr}$  which denote whether, in year  $yr$ , the  $h^{th}$  electrolyzer ( $h \in \{1, \dots, H\}$ ) of type  $i \in \{PEM, AE\}$  is of the  $s^{th}$  size  $s \in \{Available\ discrete\ sizes\}$ , with  $d_{i,s}$  denoting the available sizes for each type of electrolyzer.

Eq. (15) adds the size of all installed electrolyzers to calculate the total capacity expansion in any given year [10]. Eq. (16) ensures that at most one size is chosen for every electrolyzer. Eq. (17) is added in order to avoid degenerate solutions.[6]

$$\sum_{h=1}^H \sum_{s=d_{1,s}}^{d_{1,s}} w_{i,h,s,yr} d_{i,s} = QE_{i,yr} \forall i \in PEM, AE \quad (15)$$

$$\sum_s w_{i,h,s,yr} \leq 1 \forall i \in PEM, AE \quad (16)$$

$$\sum_s w_{i,h,s,yr} d_{i,s} \geq \sum_s w_{i,h+1,s,yr} d_{i,s} \quad (17)$$

## Cost Constraints and Objective function

Eq. (18) & (19) incorporate an upper limit on the annual capital expenditure (CAPEX) that can be allocated.

Eq. (18) shows that the total CAPEX for an expansion is denoted by  $\alpha + \beta QE$ , where  $\alpha$  represents the fixed cost and  $\beta QE$  represents the variable cost.

$$CAPEX_{yr} = \sum_i \frac{\alpha_{i,yr}(z_{i,yr} - p_{i,yr})}{dur(i)} + \sum_{p=1}^{dur(i)} \frac{\beta_{i,yr} QE_{i,yr+p}}{dur(i)} \quad (18)$$

$$CAPEX_{yr} \leq CAPEX_{CAP} \quad (19)$$

Here  $dur(i)$  is the time taken to install the unit  $i$ . The OPEX incurred for pursuing the decarbonization initiatives each year is given by Eq. (20). The OPEX consists of two primary components: the first pertains to the consumption of natural gas and electricity from their respective sources, while the second, linked to the CC units, is directly proportional to emission reductions.

$$OPEX_{yr} = \left( \sum_{j \in Out(NG\ Supply)} F_{NG\ Supply, j, NG, yr} \frac{\$}{unit\ mass} + \sum_{j \in Out(NG\ Supply)} F_{Grid, j, Electricity, yr} \frac{\$}{unit\ energy} + \sum_{i \in CCS\ Units} F_{i, Captured\ CO_2, CO_2, yr} OPEX_i \right) time\ units$$

(20)

Finally, we formulate the objective function as the Present Value of the Cost of the decarbonization project in Eq. (21) which is subject to Eqs. (1) to (20), the constraints of the optimization problem.

$$Minimize \sum_{yr=1}^{29} (CAPEX_{yr} + OPEX_{yr}) / (1+i)^{yr} \quad (21)$$

## RESULTS

The formulated retrofit MILP model was implemented in Pyomo Pyomo 6.6.1 with the Gurobi 10.0.1 solver [11], for a 29-year horizon containing 204,700 constraints and 102,918 variables (97,233 continuous, 5,684 integer). Discounting for the inequality constraints and dependent equations, there are ~3,900 degrees of freedom

### Base Case

As shown in Fig. 2 (a) for the base case when pre-defined targets for reduction in CO2 emissions. (i.e. 50% reduction by year 10 and reduction to the minimum attainable value using the given superstructure by year 28) are implemented, carbon capture technology is favored over electrified options. Electrified technologies are not chosen due to the high operating costs associated with them [2, 1, 14, 3, 13, 31, 4]. Due to space constraints details regarding the data used for natural gas and electricity prices and capital costs for e-boilers, electrolyzers and CC technologies shall be disseminated in a full-length journal publication which is under preparation. Among the CC technologies, post-combustion capture is preferred for both SMR and other combustion-related flue gases, as the additional CAPEX for pre-combustion technology outweighs operational cost savings.

Transition to electrified technologies for both steam and H<sub>2</sub> production occurs in 2049 when stringent restrictions are enforced (as shown in Fig. 2 (a) and (c)) that limits the emissions to the minimum attainable value. For electrification of H<sub>2</sub>, PEM electrolyzers are preferred at given electricity price forecasts due to the associated savings in OPEX.

As shown in Fig. 2 (b) the capital expenditure for CC alternatives was very high compared to electrified technologies and reached the upper limit. Fig. 2 (c) shows that carbon neutrality was not achieved in the final years as the current superstructure only allows post-combustion CC of the flue gas from the rest of the plant, which is not 100% efficient. Fig 2(c) also shows that the emission cap constraint is always active at the optima.

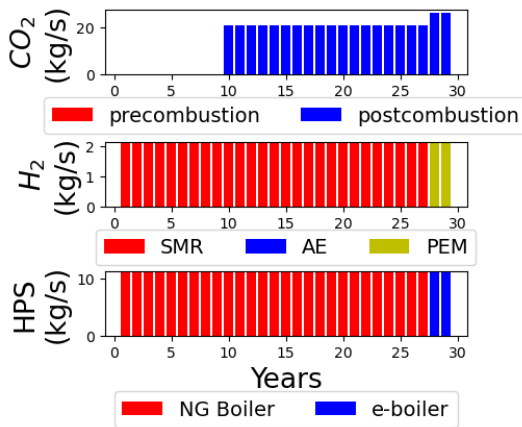
We can see in Fig. 2 (d) that lowering the maximum allowed CAPEX spending per year leads to a preference for installation and use of e-boilers until the CAPEX intensive capture facility is built. Hence, enterprise specific economic policies can alter the optimal solution.

However, adding a penalty for unused capital could alter the optima away from the presented solution.

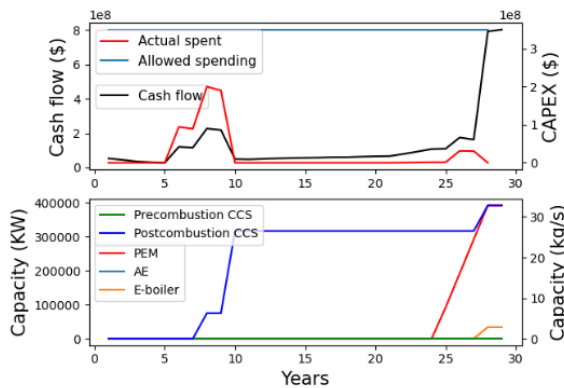
To gain insight on scenarios that could lead to adoption of electrified options, we studied how the optimal solution varied when (i) Carbon taxes are applied, and (ii) Electricity prices are reduced by different percentages.

### Case 2: Introduction of Carbon taxes

Given carbon tax implementation at specified rates as shown in Figure 3 (b), (mimicking Canadian rates) it was observed that transitioning to CC based decarbonization technologies relatively earlier is economically beneficial as shown in Fig 3(a), decreasing the overall CO<sub>2</sub> emissions. Fig 3(b) also shows that economically optimal emission levels are below the maximum allowed levels. However, C-taxes do not expedite the adoption of carbon-neutral electrified technologies at the optimal design as can be seen in Fig. 3 (a). Additional case studies have also shown us that the results remain qualitatively similar if carbon credits are introduced instead of carbon taxes.

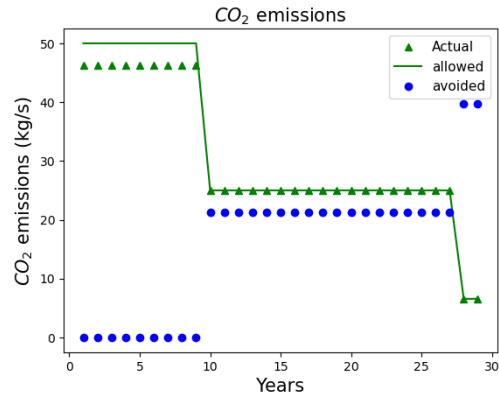


**Fig 2 (a)** The panels from top to down show the contribution of each technology in carbon capture; Hydrogen production and HPS production. Postcombustion CC is preferred over years 10-27; e-boilers and PEMS operated in years 28-29

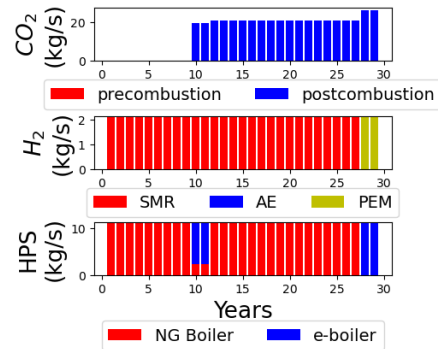


**Fig 2 (b)** Upper panel shows the actual CAPEX spending and the maximum CAPEX spending limit; along with the total cash outflow associated with hydrogen and HPS

production. The panel below shows the rate at which the electrified alternatives should be installed in KW capacity. It also shows the optimal rate of CC capacity installation in kg/s. CAPEX spending reaches maximum limit when CC infrastructure is built; CAPEX for electrified alternatives is relatively cheaper; large cash outflow after adoption of electrified alternatives



**Fig 2(c)** The total carbon-di-oxide emissions as compared to the maximum allowed emissions over the years. The emissions avoided as a result of the decarbonization initiatives are also shown. Emissions are restricted to 50% of current values from years 10-27; Emissions are capped at the minimum attainable value in years 28-29. Emission cap constraint is always active.



**Fig 2(d)** With \$100,000,000 cap in annual capital expenditure e-boiler is operated until CC technology is built

### Case 3: With reduction in electricity prices

Examining the impact of reductions in electricity costs, it was observed that with 25-70% reduction in electricity costs, the overall solution remained qualitatively similar to the base case. However, a mix of PEME and AE was chosen in the final solution due to the trade-off in CAPEX and OPEX as shown in fig 4(a). Fig 4 (b) shows that upon further reducing electricity costs to 80% below the forecasted levels, e-boiler technology is selected earlier. Nevertheless, the adoption of electrolyzer technology is delayed due to high energy requirements for electrolysis. Fig 4(c) shows that when electricity costs are reduced by 90%, a partial shift from SMR-based

hydrogen production to electrolysis-based hydrogen production is observed, suggesting the use of a mix of both types of electrolyzers. A higher capacity of AEs (Alkaline Electrolyzers) is chosen in the final mix due to the cheap electricity prices in this scenario.

### Case 4: Electricity price reduction and carbon taxes

With reduction in electricity prices and carbon taxes implemented, we further analyze three sub-cases where the electricity prices are reduced by 85%, 90%, and 95% compared to the currently forecasted prices. In all the scenarios, the introduction of carbon capture units is accelerated due to carbon tax introduction as shown in figures 4(d)-4(f). Simultaneously, electrified technologies were also chosen earlier as compared to the base case. In other words, the trends observed are a superposition of the effects obtained in the 2 previous cases. In summary, while carbon taxes Encourage early transition to CC-based technologies, substantial electricity cost reductions or stringent environmental norms are needed for favoring electrified technologies.

### CONCLUSION

This paper has proposed a MILP-based superstructure optimization model to minimize the Present Value of the Cost for retrofit decarbonization of an oil refinery. Higher electricity prices relative to natural gas favor CC over electrification alternatives. Lowering electricity costs by up to 70% has little impact, but further reductions significantly affect the optimal solution. Carbon taxes accelerate adoption of carbon capture technologies. Substantial electricity cost reductions make e-boilers financially attractive.

It should be noted that for the two different carbon capture technologies we have assumed MEA based absorption operating at different conditions and with different costs for dilute (post-combustion) and concentrated (pre-combustion) CO<sub>2</sub> concentrations of the flue gas. For each technology, the cost per unit of captured CO<sub>2</sub> was assumed to remain constant. Future work will enhance the tool by taking into account the differential pricing for carbon capture technologies as a function of the CO<sub>2</sub> concentration of the flue gas stream using piecewise linear approximations. In addition, future work may include the carbon intensity of the fuel and electricity as they are expected to change over time and affect the optimal retrofit plan.

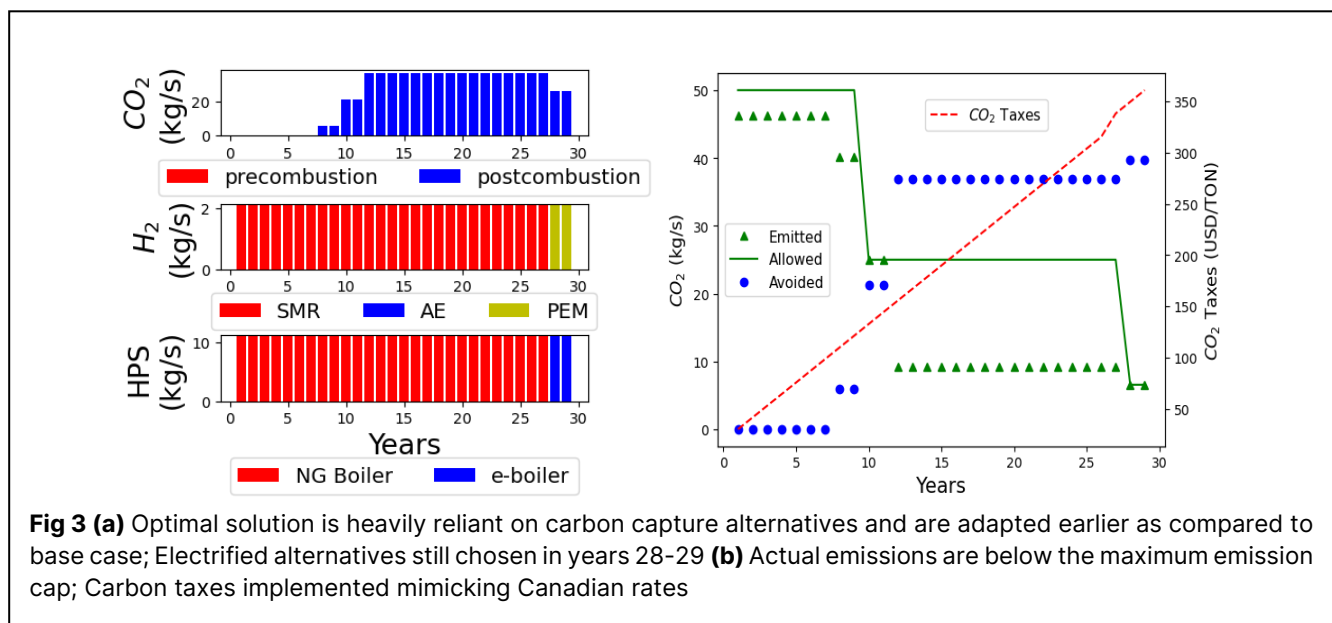
Finally, obtaining accurate cost forecasts are challenging due to energy and carbon market uncertainties. Hence, an enhancement of the solution's robustness is needed. For this, we plan to explore stochastic programming techniques to account for uncertainties in the MILP model [5].

### ACKNOWLEDGEMENTS

This work was financed [in part] by a grant from the Commonwealth of Pennsylvania, Department of Community and Economic Development.

### REFERENCES

1. "Is Carbon Capture Too Expensive – Analysis - IEA." n.d. Adam Baylin-Stern, [Niels Berghout](https://www.iea.org/commentaries/is-carbon-capture-too-expensive) <https://www.iea.org/commentaries/is-carbon-capture-too-expensive>.
2. "U.S. Energy Information Administration - EIA - Independent Statistics and Analysis." n.d.



- <https://www.eia.gov/outlooks/aeo/data/browser/#/?id=1-AEO2023&region=0-0&cases=ref2023&start=2021&end=2050&f=A&lin echart=~ref2023-d020623a.44-1-AEO2023&map=&ctype=linechart&sid=ref2023-d020623a.44-1-AEO2023&sourcekey=0>
3. IEA greenhouse gas R&D programme. "Understanding the Cost of Retrofitting CO<sub>2</sub> Capture in an Integrated Oil Refinery." SINTEF. [www.sintef.no/recap](http://www.sintef.no/recap). (2017)
  4. The International Renewable Energy Agency, The International Renewable Energy. 2020. Green hydrogen cost reduction: Scaling up electrolyzers to meet the 1.5°C climate goal. ISBN 9789292602956. [www.irena.org/publications](http://www.irena.org/publications). (2020)
  5. Birge, John R., and François Louveaux. 1997. Introduction to stochastic programming. New York, NY, USA: Springer-Verlag, New York, NY, USA, (1997)
  6. Thomas J. Watson IBM Research Center, and N. Megiddo. A note on degeneracy in linear programming. IBM Thomas J. Watson Research Division. (1985)
  7. Corengia, Mariana, and Ana I. Torres. "Coupling Time Varying Power Sources to Production of Green-Hydrogen: A Superstructure Based Approach for Technology Selection and Optimal Design." *Chemical Engineering Research and Design* 183 (July): 235–49. (2022)
  8. Cresko, Joe, Edward Rightor, Alberta Carpenter, Kathryn Peretti, Neal Elliott, Sachin Nimbalkar, William R. Morrow III, et al. "DOE Industrial Decarbonization Roadmap," September. (2022)
  9. George, Jan Frederick, Viktor Paul Müller, Jenny Winkler, and Mario Ragwitz. "Is Blue Hydrogen a Bridging Technology? - The Limits of a CO<sub>2</sub> Price and the Role of State-Induced Price Components for Green Hydrogen Production in Germany." *Energy Policy* 167: 113072. (2022)
  10. Grossmann, I. E., V. T. Voudouris, and O. Ghattas. "Mixed-Integer Linear Programming Reformulations for Some Nonlinear Discrete Design Optimization Problems." *Recent Advances in Global Optimization*, December, 478–512. (1991)
  11. Gurobi Optimization, LLC. 2023. "Gurobi Optimizer Reference Manual." <https://www.gurobi.com>.
  12. Ho, Minh T., Guy W. Allinson, and Dianne E. Wiley. "Reducing the Cost of CO<sub>2</sub> Capture from Flue Gases Using Pressure Swing Adsorption." *Industrial and Engineering Chemistry Research* 47 (July): 4883–90. (2008)
  13. Jeremiah, Shelor. n.d. "EIA Predicts 'Record-High' U.S. Oil, Natural Gas Output Through 2050 - Natural Gas Intelligence." <https://www.naturalgasintel.com/eia-predicts-record-high-u-s-oil-natural-gas-output-through-2050/>
  14. LAZARD. n.d. "Lazard's levelized cost of hydrogen analysis." <https://www.lazard.com/media/12qcx1j/lazards-levelized-cost-of-hydrogen-analysis-vf.pdf>.
  15. Liu, K., C. Song, and V. Subramani. Hydrogen and syngas production and purification technologies. (2009)
  16. Lively, Ryan P. 2021. "The Refinery of Today, Tomorrow, and the Future: A Separations Perspective." *AIChE Journal* 67. (2021)
  17. Mallapragada, Dharik S., Yury Dvorkin, Miguel A. Modestino, Daniel V. Esposito, Wilson A. Smith, Bri-Mathias Hodge, Michael P. Harold, et al. "Decarbonization of the Chemical Industry Through Electrification: Barriers and Opportunities." *Joule* 7 (1): 23–41. (2023)
  18. National Energy Technology Laboratory (NETL). n.d. "Carbon Dioxide Capture Approaches." <https://netl.doe.gov/research/carbon-management/energy-systems/gasification/gasifipedia/capture-approaches>.
  19. Norton, Leon C, and Ignacio E Grossmann. "Strategic Planning Model for Complete Process Flexibility." *Ind. Eng. Chem. Res* 33: 69–76. (1994)
  20. Ogden, Joan M. "Review of small stationary reformers for hydrogen production." (2001).
  21. Oni, A. O., K. Anaya, T. Giwa, G. Di Lullo, and A. Kumar. "Comparative Assessment of Blue Hydrogen from Steam Methane Reforming, Autothermal Reforming, and Natural Gas Decomposition Technologies for Natural Gas-Producing Regions." *Energy Conversion and Management* 254 (February): 115245. (2022)
  22. Psarras, Peter, Jiajun He, Hélène Pilorgé, Noah McQueen, Alexander Jensen-Fellows, Kourosh Kian, and Jennifer Wilcox. "Cost Analysis of Carbon Capture and Sequestration from u.s. Natural Gas-Fired Power Plants." *Environmental Science and Technology* 54 (May): 6272–80. (2020)
  23. Sahinidis, N. V., I. E. Grossmann, R. E. Fornari, and M. Chathrathi. "Optimization Model for Long Range Planning in the Chemical Industry." *Computers and Chemical Engineering* 13 (September): 1049–63. (1989)
  24. Siegelman, Rebecca L., Phillip J. Milner, Eugene J. Kim, Simon C. Weston, and Jeffrey R. Long. "Challenges and Opportunities for Adsorption-Based CO<sub>2</sub> Capture from Natural Gas Combined Cycle Emissions." *Energy and Environmental Science* 12: 2161–73. (2019)
  25. UNFCCC. 2021. "Report of the Conference of the Parties on its twenty-sixth session, held in

Glasgow from 31 October to 13 November 2021.”  
[https://unfccc.int/sites/default/files/resource/cp20\\_21\\_12\\_add1E.pdf](https://unfccc.int/sites/default/files/resource/cp20_21_12_add1E.pdf).

26. Wallace, William D., and Lawrence G. Spielvogel. “Field Performance of Steam and Hot Water Electric Boilers.” *IEEE Transactions on Industry Applications* IA-10 (November): 761–69. (1974)

27. Wang, Qinhui. “Hydrogen Production.” *Handbook of Climate Change Mitigation*, 1091–1130. (2012)

28. Zhai, Haibo. “Advanced Membranes and Learning Scale Required for Cost-Effective Post-Combustion Carbon Capture.” *iScience* 13 (March): 440–51. (2019)

29. Zhang, Lifeng, Zhihong Yuan, and Bingzhen Chen. “Refinery-Wide Planning Operations Under Uncertainty via Robust Optimization Approach Coupled with Global Optimization.” *Computers and Chemical Engineering* 146 (March): 107205. (2021)

30. Zhang, Lifeng, A. I Torres, Chen Bingzhen, Yuan Zhihong, and I. E. Grossmann. “Optimal Retrofitting of Conventional Oil Refinery into Sustainable Bio-Refinery Under Uncertainty” In *Submitted for Publication*.(2023)

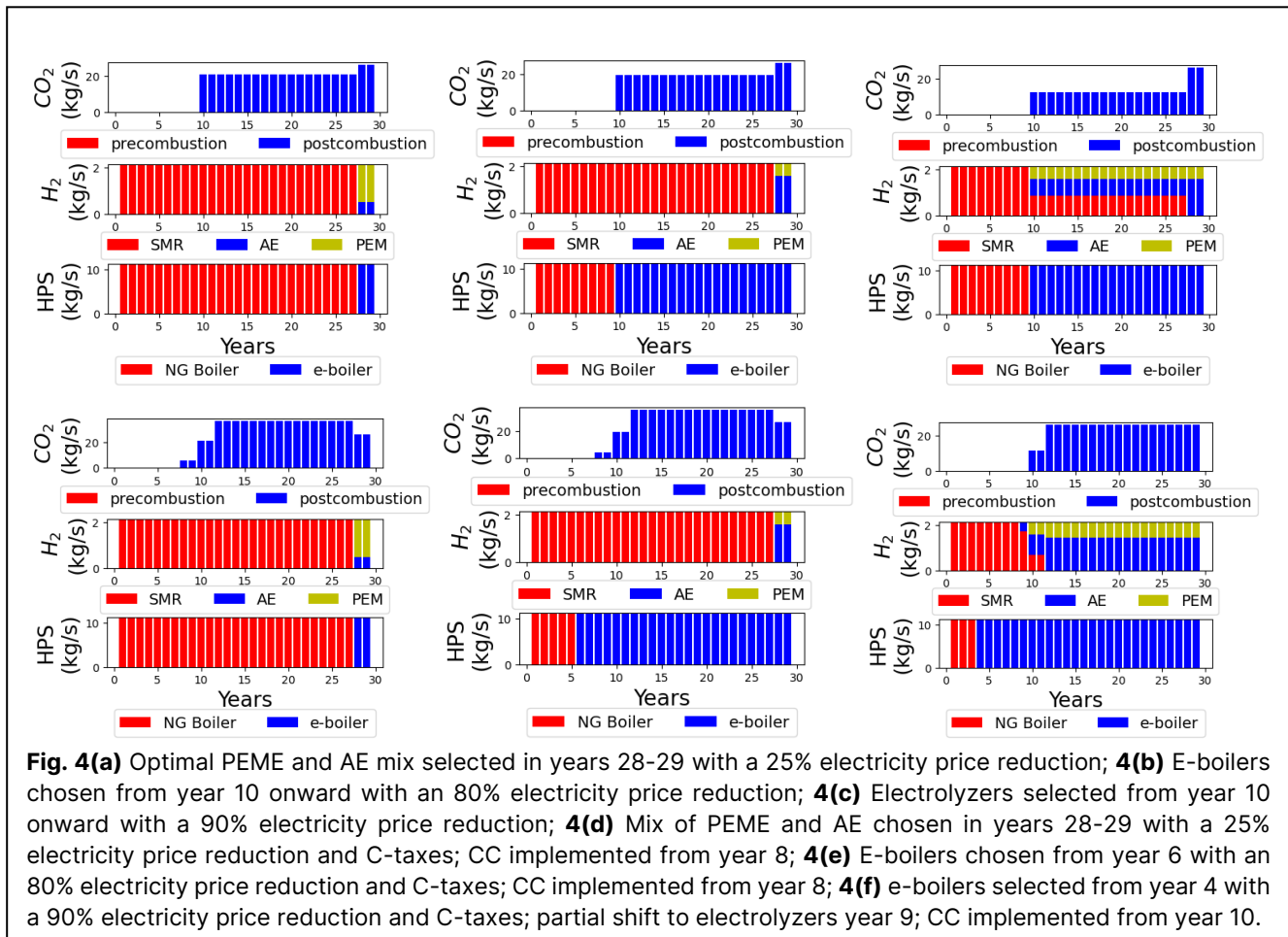
31. Zuberi, M. Jibrán S., Ali Hasanbeigi, and William Morrow. 2022“Electrification of Industrial Boilers in

the USA: Potentials, Challenges, and Policy Implications.” *Energy Efficiency*. 70. (2022).

32. EPA. Scope 1 and scope 2 inventory guidance, (2022)  
<https://www.epa.gov/climateleadership/scope-1-and-scope-2-inventory-guidance>

33. Agrawal Rakesh, Siirola. Jeff, Decarbonization of Chemical Process Industries via Electrification. 2023, 32–40. Volume 53 No. 2, Summer 2023, “The Bridge.” NAE Website,  
<https://nae.edu/21020/Bridge>. Accessed 4 Apr. 2024.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Conceptual Design of Integrated Energy Systems with Market Interaction Surrogate Models

Xinhe Chen<sup>a</sup>, Radhakrishna Tumbalam-Gooty<sup>b,c</sup>, Darice Guittet<sup>d</sup>, Bernard Knueven<sup>d</sup>, John D. Siirola<sup>e</sup>, and Alexander W. Dowling<sup>a\*</sup>

<sup>a</sup> University of Notre Dame, Department of Chemical and Biomolecular Engineering, South Bend, IN 46556, United States

<sup>b</sup> National Energy Technology Laboratory (NETL), Pittsburgh, PA 15236, United States

<sup>c</sup> NETL Support Contractor, Pittsburgh, PA 15236, United States

<sup>d</sup> National Renewable Energy Laboratory, Golden, CO 80401, United States

<sup>e</sup> Sandia National Laboratories, Albuquerque, NM 87185, United States

\* Corresponding Author: adowling@nd.edu

## ABSTRACT

Most integrated energy system (IES) optimization frameworks employ the price-taker approximation, which ignores important interactions with the market and can result in overestimated economic values. In this work, we propose a machine learning surrogate-assisted optimization framework to quantify IES/market interactions and thus go beyond price-taker. We use time series clustering to generate representative IES operation profiles for the optimization problem and use machine learning surrogate models to predict the IES/market interaction. We quantify the accuracy of the time series clustering and surrogate models in a case study to optimally retrofit a nuclear power plant with a polymer electrolyte membrane electrolyzer to co-produce electricity and hydrogen.

**Keywords:** Integrated Energy System, Surrogate Models, Machine Learning, Optimization, Time Series Clustering

## INTRODUCTION

Integrated energy systems (IES) exploit synergies between different technologies and energy carriers [1] such as fossil, nuclear, renewable (e.g., solar, wind), and storage to produce multiple products such as electricity, heat, and chemicals while increasing energy efficiency and supporting greater renewable integration into the grid. Moreover, IES can provide more flexibility to the grid, which is critical for increasing the integration of non-dispatchable renewable energy sources and meeting decarbonization goals.

The price-taker approximation is a widely used approach to incorporate dynamic market signals into the optimization of IES. The price-taker approximation treats the electricity grid as an “infinite bus” that can receive any amount of electricity produced by each generator or IES at any time without affecting the location marginal price (LMP) of electricity. Price-taker allows for co-optimizing IES design and operating decisions using historical or forecasted time-series LMPs. Lakey et al. [2]

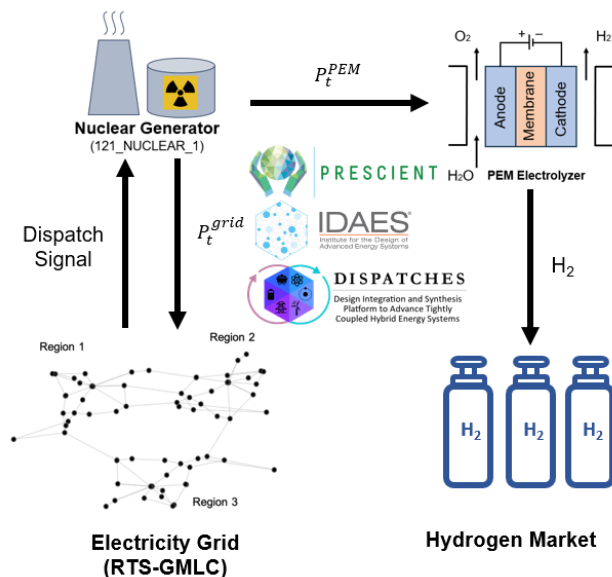
proposed a framework that allows users to generate a multi-period, price-taker model instance and automatically generate common operational constraints for their model in the IDAES-PSE platform [3].

Recent papers [2-5] highlight the inaccuracies of the price-taker approximation compared to more rigorous production cost models (PCM) for analyzing IES. Specifically, Martinek et al. [4] compare the revenue, dispatch, and scheduling of concentrating solar power plants using the PCM and price-taker models. They show the price-taker approximation can over-aggressively respond to short-duration peaks in LMP. Likewise, Frew et al. [5] show the price-taker approximation can overestimate the value of IES that co-produces electricity and hydrogen. Sousa et al. [6] evaluate the integration of wind energy and pump storage systems and find that in the PCM, storage operation will decrease the electricity price and dispatch less wind power than the price-taker. Finally, Xian et al. [7] propose a multiscale simulation framework that combines rigorous process (IES) and grid (PCM) models to quantify IES-grid interactions across

hourly to annual timescales. They find that optimizing the design, operation, and control of a single IES impacts the prices and dispatch of generators across the electric grid.

We recently proposed a machine learning surrogate-assisted optimization framework [8] to integrate the process-centric model (price-taker) and grid-centric model (PCM) into the IES conceptual design optimization. Specifically, we contemplate how to size a replacement generator while considering market impacts. We compare algebraic and neural network surrogate models to predict market revenue, annual capacity factor distribution, and the number of shutdowns as a function of the characteristics of the replacement generator. We embed these surrogates into a nonlinear optimization problem and show more accurate results versus the price-taker approximation compared to PCM simulations (ground truth). While promising, this approach only considers steady-state models and surrogates.

In this short conference paper, we extend our prior work [8] by using time series clustering to generate time series representative dispatch scenarios. Then, we train a frequency surrogate model that predicts the frequency of each scenario according to the different IES designs. Together, these surrogate models provide dynamic operation profiles that may be embedded in multiscale IES optimization problems. We develop the surrogate models and clustering methods using a baseload nuclear power plant (NPP) with a polymer electrolyte membrane (PEM) electrolyzer that produces hydrogen ( $H_2$ ) as a case study.



**Figure 1.** The nuclear + PEM IES can divert the power to co-produce hydrogen when the LMP is low to increase its profitability and flexibility.

## Nuclear and $H_2$ IES

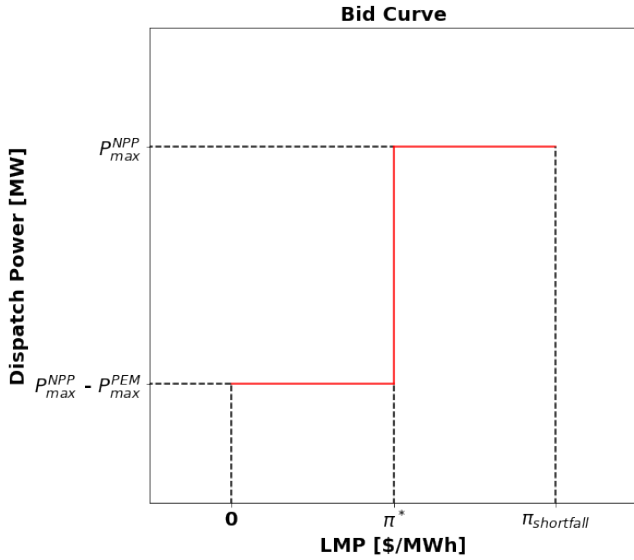
We consider retrofitting an existing nuclear power plant with a polymer electrolyte membrane (PEM) electrolyzer to co-produce electricity and hydrogen, as shown in Figure 1. The NPP has a nameplate capacity of 400MW and operates strictly as a baseload generator. After retrofitting, the IES (i.e., NPP + PEM) operates as a hybrid baseload and peaking generator when power from the NPP can be diverted to the PEM to produce hydrogen. Thus, from the perspective of the electric grid, the PEM provides flexibility. Hydrogen is an important resource for the future energy industry and is essential to the decarbonization of many industries [9]. The retrofit thus increases and diversifies the NPP's revenue.

## Production Cost Models

A PCM simulates the operation of an electricity system by scheduling generators and clearing the market (i.e., setting time-varying prices). In summary, each generator communicates its (time-varying) production costs and operational constraints (e.g., minimum up and down times, startup costs) to the market. This includes a bid curve (see Figure 2 for an example), which communicates the cost of producing electricity as a function of the market price. Using this information, the market operator solves unit commitment and economic dispatch optimization problems, which seek to minimize the total system-wide generation costs subject to the costs and constraints of each generator and forecasted demand and renewables production. Most regions schedule most electricity generation in the day-ahead market (DAM). The real-time market (RTM) operates (sub)hourly to correct for forecasting errors and unplanned events. The PCM simulates the rolling horizon operation of the DAM and RTM. See [10] for details.

In this work, all PCM simulations were performed using the open-source production cost model Prescient [11] and the RTS-GMLC [12] dataset. RTS-GMLC is an open-source test network that approximates the characteristics of the southwest United States [7]. We consider retrofitting the 121\_NUCLAER-1 nuclear generator in RTS-GMLC region 1 with a PEM.

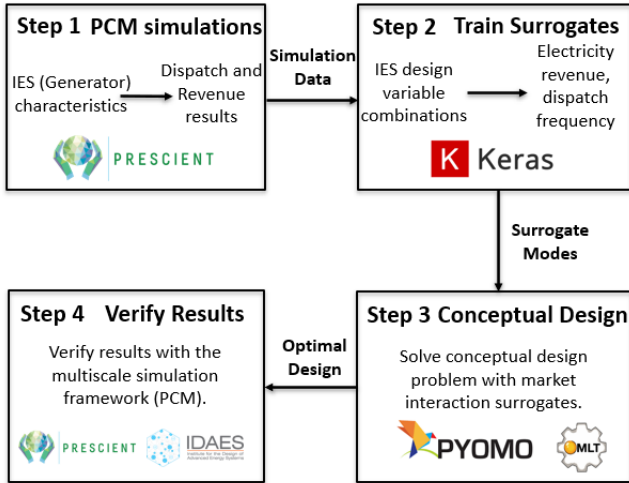
## METHODS



**Figure 2.** Bid curve of the NPP-PEM IES. The IES will dispatch all power to the grid when the electricity price exceeds  $\pi^*$ . Otherwise, the IES will divert  $P_{max}^{NPP} - P_{max}^{PEM}$  power to co-produce  $H_2$ .

### Surrogate-Assisted Optimization Workflow

Figure 3 illustrates the four-step surrogate-assisted optimization workflow.



**Figure 3.** Surrogate-assisted IES conceptual design workflow, adapted from [6].

**Step 1: Perform PCM simulations to generate training data for surrogate models.** We consider two IES design variables: (1) the ratio of maximum PEM power ( $P_{max}^{PEM}$ ) to the maximum power plant power ( $P_{max}^{NPP}$ ), and (2) the threshold LMP price ( $\pi^*$ ) that the IES diverts from the electricity production mode to the hydrogen co-production mode. In PCM simulations, we use different bidding curves [13] (production-cost pairs, see Figure 2) to communicate the IES's flexibility and economics to the

market. We discretize the continuous IES design space and simulate all combinations in the PCM. We performed 192 annual IES PCM simulations in the NPP-PEM IES case study. PCM simulation results give detailed hourly LMP and dispatch profiles within the simulation horizon (366 days).

**Step 2: Train IES-market surrogate models.** Using the LMP and dispatch profile, we calculate the total electricity revenue of each IES design combination. Then, we train a neural network revenue surrogate model (2 hidden layers, 25 nodes each layer and activated by hyperbolic tangent) to predict the annual electricity revenue according to the IES design. We subdivide the annual hourly time series IES dispatch profile into 366 daily dispatch profiles. Next, we use time series clustering [14] to identify 24-hour representative dispatch profiles for the IES conceptual design optimization problem. A neural network dispatch frequency surrogate model (3 hidden layers, 75 nodes each layer and activated by sigmoid) is trained to predict the occurrence probability of each representative day. Inputs of surrogate models are two design variables mentioned in Step 1, plus PCM simulation reserve factor and load shed price. Both surrogate models are trained by Keras [15], and we use 80% of the data for training and 20% for validation.

**Step 3: Solve conceptual design optimization of IES with surrogate models.** We then formulate and solve a stochastic optimization problem, Equations (1) to (5), with IES-market surrogate models to obtain the optimal IES design and operation decisions:

$$\max \varphi[R(x) + \sum_s \omega_s(x) \sum_t R_{H2}(x, u_{s,t}, \delta_{s,t}) - C(x, u_{s,t}, \delta_{s,t})] - CAPEX(x) \quad (1)$$

$$\text{s. t.} \quad h(x, u_{s,t}, \delta_{s,t}) \leq 0, s \in S, t \in T \quad (2)$$

$$g(x, u_{s,t}, \delta_{s,t}) = 0, s \in S, t \in T \quad (3)$$

$$R(x) = f_{rev}(x) \quad (4)$$

$$\omega_s(x) = f_{dis}(x), s \in S \quad (5)$$

The objective function is the 30-year NPV value of the IES retrofit. The two-stage stochastic optimization problem aims to optimize the IES design and operation together under the operation uncertainty. Table 1 summarizes the nomenclature.

For brevity, we do not report optimization results for (1) – (5) in this short conference paper. We recommend using OMLT [16] to embed neural network surrogate models in Pyomo [17]. Because the neural network uses nonlinear activation functions, we recommend using IPOPT [18] with HSL linear algebra [19] distributed as part of the IDAES platform [3].

**Step 4: Verify the optimal IES design with PCM.** Perform PCM simulation of the optimal design IES and compare the result obtained from Step 3. If the verification result varies dramatically from the optimization

results, the training data can be expanded and Steps 2 to 4 repeated.

This conference paper focuses on the time-series clustering in Steps 1 and 2.

**Table 1.** Conceptual design model nomenclature.

Symbol	Meaning
$s, S$	Scenario (Set)
$t, T$	Time step (Set)
$R()$	Revenue surrogate model
$\omega_s()$	Frequency surrogate model
$C()$	Operation cost
$CAPEX()$	Capital cost
$h()$	Inequality constraints
$g()$	Equality constraints
$x$	Surrogate input variables
$u$	Operation variables
$\delta$	Representative dispatch profiles
$\varphi$	NPV multiplier

### Time Series Clustering of PCM Data

Time series clustering is employed to generate representative IES operation profiles. In Step 1, we perform PCM simulation of different IES designs and obtain an annual IES operation profile. We slice the dispatch profile into 24-hr daily time series and use the K-means clustering algorithm to minimize the within-cluster sum of squares Euclidean distance to get the representative days.

In the NPP-PEM model, due to the different PEM

max capacity, the minimum power of the IES will be different. To increase the clustering accuracy and avoid the potential infeasibility in the conceptual design model, we scale the capacity factor of NE/PEM IES between 0 and 1, given in Equation (6). The 0 corresponds to the minimum power output to the grid, i.e., when the PEM is operating at the maximum power.

$$cf_t^{NPP} = \frac{p_t^{NPP} - p_{min}^{NPP}}{p_{max}^{NPP} - p_{min}^{NPP}} \quad (6)$$

Before applying K-means clustering, we filter the time series with capacity factors equal to 0 (minimum dispatch at  $p_{max}^{NPP} - p_{max}^{PEM}$ ) or 1 (max dispatch at  $p_{max}^{NPP}$ ) at every time step. These time series are placed into two clusters labeled “min” and “max” in Figure 4. We find this filtering improves the accuracy of K-means clustering and reduces the size of the clustering dataset by 43.5%.

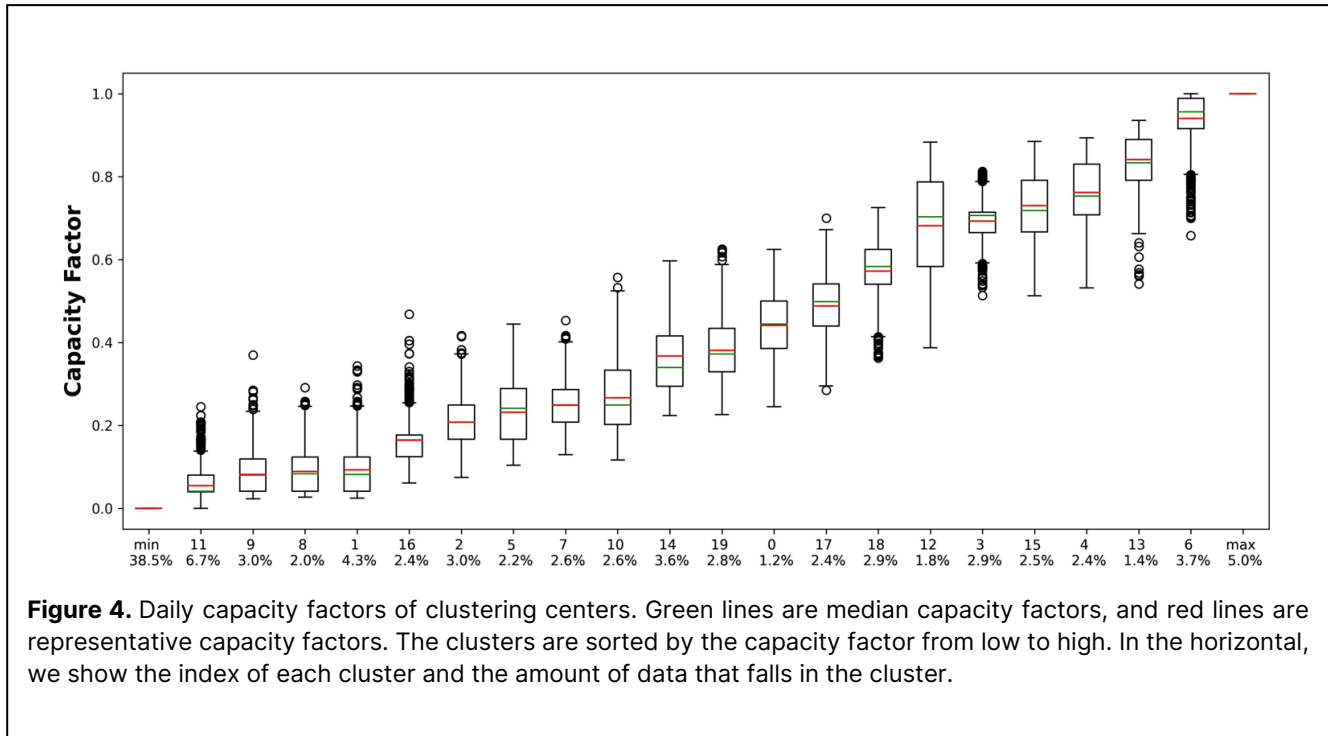
## RESULTS AND DISCUSSION

### Time Series Clustering Centers Reliably Represent the Entire Dispatch Dataset

In Figure 4, we use a box plot to sort the daily capacity factor of all clusters from “min” to “max” calculated via:

$$cf_{daily}^{NPP} = \frac{\sum_{t=1}^T cf_t^{NPP}}{T} \quad (7)$$

Figure 4 shows that cluster centers accurately represent the data in the cluster, and each cluster center has a substantial amount of data. As we can see there, except for the “min” and “max,” all other clusters hold the time



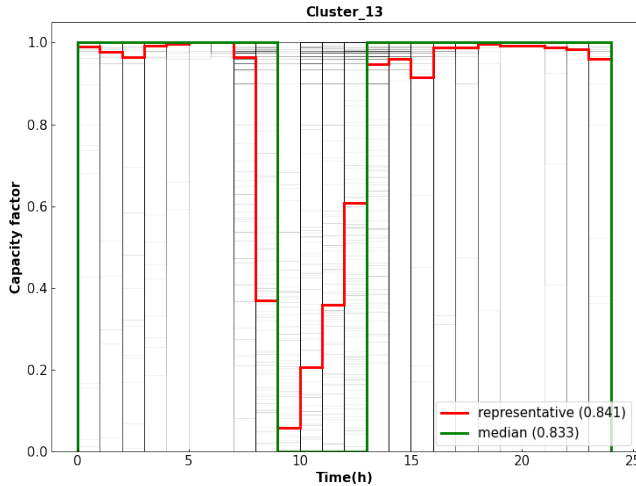
**Figure 4.** Daily capacity factors of clustering centers. Green lines are median capacity factors, and red lines are representative capacity factors. The clusters are sorted by the capacity factor from low to high. In the horizontal, we show the index of each cluster and the amount of data that falls in the cluster.

series from 1.2% to 6.7%, and there is no single cluster with a significant large or small amount of data. None of the clusters have a significant number of outliers.

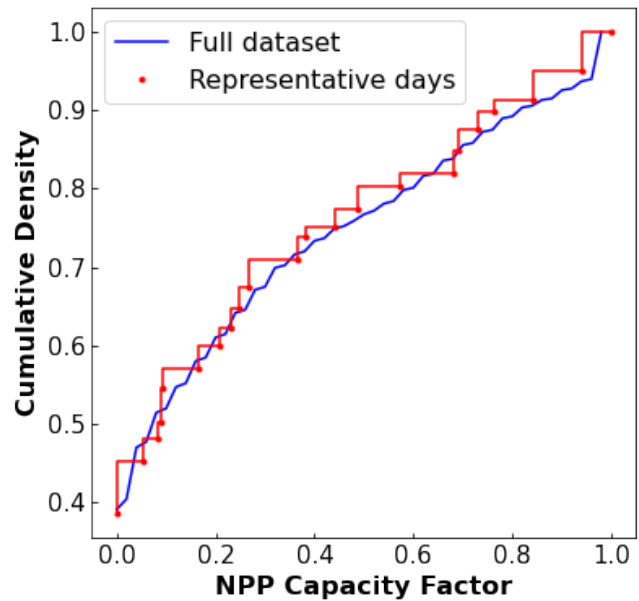
As shown in Figure 2, the NPP-PEM IES offers power between  $P_{max}^{NPP} - P_{max}^{PEM}$  (min) and  $P_{max}^{NPP}$  (max). Moreover, because the PEM model is linear (constant efficiency assumption), the bid curve is piecewise constant with two segments. This means the NPP-PEM is almost always dispatched at min or max (unless it is the marginal generator). However, the K-means clustering algorithm gives the average of the time series data within the cluster, whose values may deviate from the true NPP-PEM IES operation. To avoid this, we can choose the time series data with the median capacity factor in each cluster to replace the representative time series generated by the K-means algorithm. Figure 4 shows that the median and mean capacity factors are very similar.

Figure 5 shows the training data and results for cluster 13. The median and representative time series have similar patterns and capacity factors. However, all values in the median time series are either 0 or 1, which mimics the NPP-PEM IES dispatch from the PCM simulations.

Figure 6 compares the cumulative density of the NPP capacity for the entire dataset and representative days. Using a large number of representative days ( $N = 22$ ) helps ensure the capacity factors of the representative days match the entire dataset, as all red dots are distributed closely around the blue line.



**Figure 5.** Clustering results of cluster 13. Black lines are the time series data in this cluster. The red and green lines are the representative time series calculated by K-means and the time series data from cluster 13 with the median capacity factor, respectively.



**Figure 6.** Comparison of the cumulative density of the capacity factor. The blue line is the full dataset used as input to the time series clustering algorithm, and the red line and dots are the representative days.

### Surrogate Models are Accurate

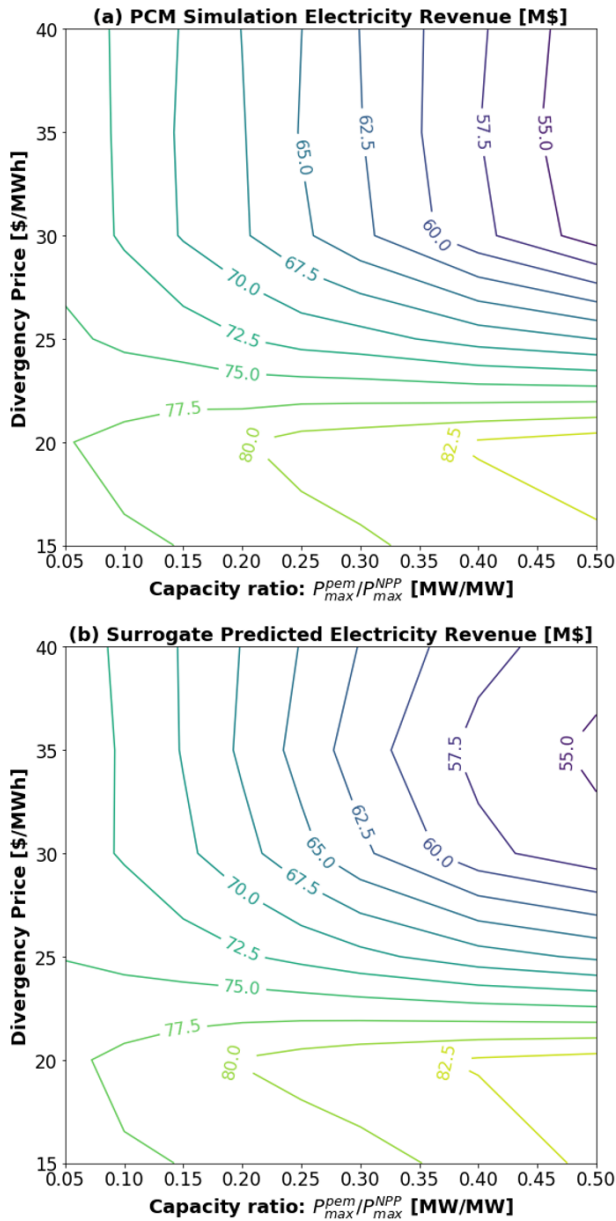
Figure 7 compares contour plots of the electricity revenue from PCM simulations and surrogate model predictions. These plots are strikingly similar, with the root-mean-squared error (RMSE) of 0.71 M\$ and  $R^2$  of 0.995, emphasizing the surrogate’s accuracy. Moreover, in both plots, the electricity revenue does not monotonically decrease with the increasing power of the PEM, which highlights the non-trivial IES-market interactions. Diverting power to co-produce hydrogen reduces the supply of electricity in the market, which increases the LMP; the average day-ahead LMP of all PCM simulations (22.43 \$/MWh) is 1.5% higher than the price-taker signal (22.09 \$/MWh).

The output of the frequency surrogate model predicts the weight of each representative day regarding the given IES design. For the different input  $x$ , the frequency surrogate model predicts the frequency of each representative day within a year. Table 2 summarizes the  $R^2$  value of all 22 representative days.

**Table 2:**  $R^2$  of dispatch frequency surrogate models

	Value
Mean $R^2$	0.9362
Standard Deviation	0.0747
Max $R^2$	0.9988
Min $R^2$	0.7631





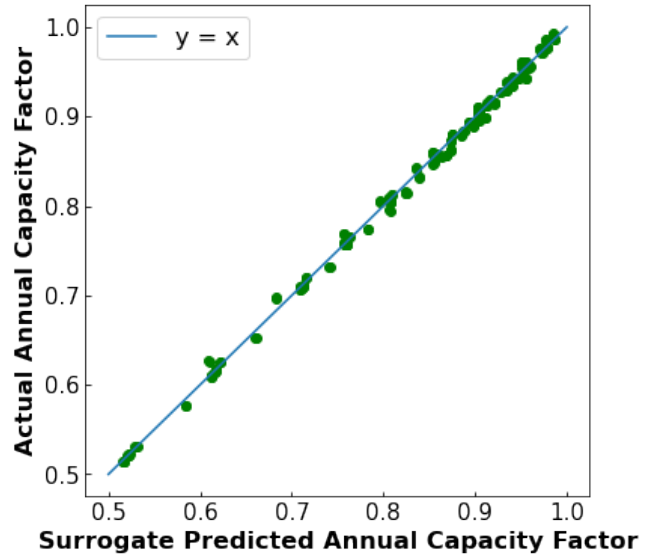
**Figure 7.** Contour plots of electricity revenue from (a) PCM Simulation and (b) revenue market surrogate model. In (a), electricity revenue is calculated from 48 PCM simulations with different divergency price and capacity ratio combinations. In (b), we use the surrogate model to predict the electricity revenue using divergency price and capacity ratio combinations as inputs.

To further test the accuracy of the frequency model, we calculated the annual NPP-PEM IES capacity factor predicted by the frequency surrogate model. We compared then compare these predictions with the PCM simulation data. The annual capacity factor is calculated by Equation (8):

$$c_{annual}^{NPP} = \sum_s N_{day} \cdot \omega_s \cdot \sum_t \frac{c_t^{NPP}}{T} \quad (8)$$

$\omega_s$  is the weight predicted by the frequency surrogate model.  $c_t^{NPP}$  is the capacity factor of the IES at time  $t$  in each representative day,  $N_{day}$  is the number of days in a year (equal to 366).

Figure 9 shows the parity plot of the predicted versus actual annual IES capacity factors. The predicted annual capacity factors closely agree with the PCM simulation results, with the root-mean-squared error (RMSE) of 0.006.



**Figure 9.** Annual capacity factor parity plot. The green dots correspond to the 192 annual simulations considered in the training dataset. The parity line is shown in blue.

## CONCLUSIONS AND FUTURE WORK

This work develops time-series clustering methods for surrogate-assisted IES design and operation optimization considering IES/market interactions. The result shows that time series clustering identifies accurate representative days. The representative days and frequency surrogate models accurately predict the annual capacity factor. Finally, the neural network surrogate models for revenue accurately capture complex IES/market interactions as revealed by the PCM simulations (ground truth).

In future work, these surrogate models should be compared within the context of the IES optimization problem (Step 3) and compared to PCM verification simulations (Step 4). Moreover, the modeling strategy should be extended to consider energy storage systems [19], uncertainty from renewables, and additional revenues from ancillary services [20]. We also plan to extend the approach to consider price fluctuations in co-products, which is especially important as hydrogen and other renewable-energy-derived chemical markets continue to mature.

## ACKNOWLEDGEMENTS

This work was conducted as part of the Design Integration and Synthesis Platform to Advance Tightly Coupled Hybrid Energy Systems (DISPATCHES) project through the Grid Modernization Lab Consortium with funding from the U.S. Department of Energy's Office of Fossil Energy and Carbon Management, Office of Nuclear Energy, and Hydrogen and Fuel Cell Technologies Office. Additional work was conducted as part of the Institute for the Design of Advanced Energy Systems (IDAES) with support through the Simulation-Based Engineering within the U.S. Department of Energy's Office of Fossil Energy and Carbon Management. Additional work was conducted in part by appointments for A.W.D. and X.C. to the U.S. Department of Energy (DOE) Postgraduate and Faculty Research Programs at the National Energy Technology Laboratory administered by the Oak Ridge Institute for Science and Education (ORISE). This work was authored in part by authors at the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

## DISCLAIMER

This project was funded by the Department of Energy, National Energy Technology Laboratory an agency of the United States Government, through a support contract. Neither the United States Government nor any agency thereof, nor any of its employees, nor the support contractor, nor any of their employees, makes any warranty, expressor implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

The U.S. Government retains and the publisher, by

accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

## REFERENCE

1. Arent, D. J., Bragg-Sitton, S. M., Miller, D. C., Tarka, T. J., Engel-Cox, J. A., Boardman, R. D., ... and Garfield, D. J. Multi-input, multi-output hybrid energy systems. *Joule*, 5(1), 47-58 (2021).
2. Laky, D.J., Tumbalam-Gooty, R., Jaffe, T., Holly, M., Atia, A., Chen, X., & Dowling, A.W. IDAES-PSE Software Tools for Optimizing Energy Systems and Market Interactions. FOCAPD, 2024.
3. Lee, A., Ghose, J.H., Eslick, J.C., Laird, C.D., Sirola, J.D., Zamarripa, M.A., Gunter, D., Shinn, J.H., Dowling, A.W., Bhattacharyya, D. and Biegler, L.T. The IDAES process modeling framework and model library—Flexibility for process simulation and optimization. *Journal of Advanced Manufacturing and Processing*, 3(3), p.e10095 (2021).
4. Sorourifar, F., Zavala, V.M. and Dowling, A.W. Martinek, J., Jorgenson, J., Mehos, M., and Denholm, P. A comparison of price-taker and production cost models for determining system value, revenue, and scheduling of concentrating solar power plants. *Applied energy*, 231, 854-865 (2018).
5. Frew, B., Levie, D., Richards, J., Desai, J. and Ruth, M. Analysis of multi-output hybrid energy systems interacting with the grid: Application of improved price-taker and price-maker approaches to nuclear-hydrogen systems. *Applied Energy*, 329, p.120184 (2023).
6. Sousa, J.A., Teixeira, F. and Faias, S. Impact of a price-maker pumped storage hydro unit on the integration of wind energy in power systems. *Energy*, 69, pp.3-11 (2014).
7. Gao, X., Knueven, B., Sirola, J.D., Miller, D.C. and Dowling, A.W. Multiscale simulation of integrated energy system and electricity market interactions. *Applied Energy*, 316, p.119017 (2022).
8. Jalving, J., Ghose, J., Cortes, N., Gao, X., Knueven, B., Agi, D., Martin, S., Chen, X., Guittet, D., Tumbalam-Gooty, R., Bianchi, L., Beattie, K., Gunter, D., Sirola, J.D., Miller, D.C. and Dowling, A.W. Beyond price taker: Conceptual design and optimization of integrated energy systems using machine learning market surrogates. *Applied Energy*, 351, p.121767 (2023).
9. Mekhilef, S., Saidur, R. and Safari, A. Comparative study of different fuel cell technologies. *Renewable*

and Sustainable Energy Reviews, 16(1), pp.981-989 (2012).



10. FER Commission. Energy Primer A Handbook for Energy market Basics (2020).
11. RTS-GMLC, Reliability test system-grid modernization lab consortium. <https://github.com/GridMod/RTS-GMLC>. [Accessed 16 November 2023].
12. Prescient production cost modeling platform. <https://www.osti.gov/doecode/biblio/47057>. [Accessed 16 November 2023].
13. Gao, X. and Dowling, A.W. Making money in energy markets: Probabilistic forecasting and stochastic programming paradigms. In 2020 American Control Conference (ACC) (pp. 168-173). IEEE (2020).
14. Tavenard, R., Faouzi, J., Vandewiele, G., Divo, F., Androz, G., Holtz, C., Payne, M., Yurchak, R., Rußwurm, M., Kolar, K. and Woods, E. Tslearn, a machine learning toolkit for time series data. The Journal of Machine Learning Research, 21(1), pp.4686-4691 (2020).
15. Keras, a Python interface for artificial neural networks. <https://keras.io>. [Accessed 16 November 2023].
16. Ceccon, F., Jalving, J., Haddad, J., Thebelt, A., Tsay, C., Laird, C.D. and Misener, R. OMLT: Optimization & machine learning toolkit. The Journal of Machine Learning Research, 23(1), pp.15829-15836 (2022).
17. Hart, W.E., Watson, J.P. and Woodruff, D.L. Pyomo: modeling and solving mathematical programs in Python. Mathematical Programming Computation, 3, pp.219-260 (2011).
18. Wächter, A. and Biegler, L.T. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. Mathematical programming, 106, pp.25-57 (2006).
19. HSL, A collection of Fortran codes for large scale scientific computation. <http://hsl.rl.ac.uk/>. [Accessed 16 November 2023].
20. Integrated multiscale design, market participation, and replacement strategies for battery energy storage systems. IEEE Transactions on Sustainable Energy, 11(1), pp.84-92 (2018).
21. Dowling, A.W., Kumar, R. and Zavala, V.M. A multi-scale optimization framework for electricity market participation. Applied Energy, 190, pp.147-164 (2017).

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>

# Preliminary Examination of the Biogas-to-Hydrogen Conversion Process

Hegwon Chung<sup>a</sup>, Minseong Park<sup>a</sup>, and Jiyong Kim<sup>a\*</sup>

<sup>a</sup> Sungkyunkwan University, School of Chemical Engineering, Suwon, Gyeonggi-do, South Korea

\* Corresponding Author: [jiyongkim@skku.edu](mailto:jiyongkim@skku.edu).

## ABSTRACT

Biogas is a promising energy source for sustainable hydrogen production due to its high concentration of CH<sub>4</sub>. However, determining the optimal process configuration is challenging due to the uncertainty of the fed biogas composition and the sensitivity of the operating conditions. This necessitates early-stage evaluation of the biomass-to-hydrogen process's performance, considering economics, energy efficiency, and environmental impacts. A data-driven model was introduced for early-stage assessment of hydrogen production from biogas without whole process simulation and optimization. The model was developed based on various biogas compositions and generated parameters for mass and energy balance. A database of unit processes was created using simulation models. Sensitivity analysis was performed under four techno-economic and environmental evaluation criteria: Unit Production Cost (UPC), Energy Efficiency (EEF), Net CO<sub>2</sub> equivalent Emission (NCE), and Maximum H<sub>2</sub> Production (MHP). The early-stage evaluation of the biogas-to-hydrogen process can guide the establishment of biogas utilization strategies and propose effective biogas enhancement process development solutions to respond to market disturbances.

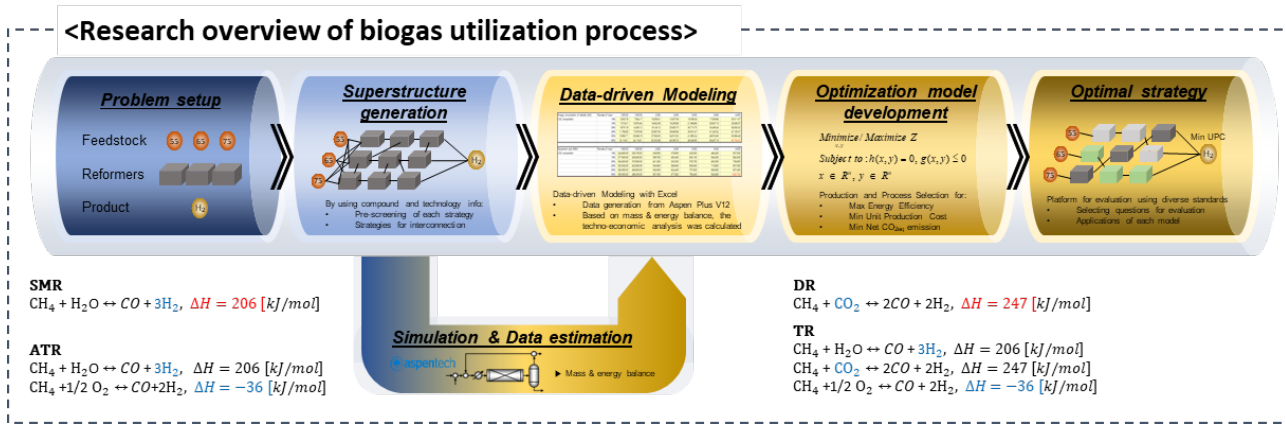
**Keywords:** Hydrogen, Biosystems, Optimization, Environment, Technoeconomic Analysis, Data-driven model

## 1. INTRODUCTION

Continuous concerns over global warming and the depletion of fossil fuels have spurred research into efficient technologies and sustainable energy generation. Currently, hydrogen is gaining increasing attention as a substantial energy carrier, primarily due to its environmental advantages and high calorific value. Although hydrogen is a potentially valuable material especially for the environment as a fuel, the conventional steam methane reforming plant, responsible for 50% of conventional hydrogen production emits 13.7 kg eq. CO<sub>2</sub> per kg of net H<sub>2</sub> produced [1]. Biogas is a viable alternative raw material for hydrogen production, as the methane within biogas can be utilized to prevent greenhouse gas emissions into the atmosphere [2]. There are various sources from which biogas can be obtained, such as sewage sludge digesters, organic waste digesters, or landfills. The biogas typically comprises 55 – 75% of methane, 25-45% of CO<sub>2</sub>, and trace elements like nitrogen (N<sub>2</sub>), oxygen (O<sub>2</sub>), hydrogen sulfide (H<sub>2</sub>S), siloxanes, and some dust

particles [3]. Before constructing a biogas utilization plant, an early-stage assessment of the biogas-to-hydrogen process is necessary. Zhao X, et al investigated various biogas reforming technologies utilizing different catalysts to convert biogas to syngas. They conducted techno-economic analysis of biogas conversion technologies [4]. However, the previous research has not explored and evaluated the most efficient pathways for biogas utilization.

In this study, we have presented a superstructure-based data-driven model that can provide optimal strategies for biogas-to-hydrogen conversion. We carried out a comprehensive analysis of economic, environmental, and technological aspects by generating mass and energy balance data for every possible pathway and operating conditions from the superstructure. The main operating variables include the types of reformers and each operating conditions for various criteria. The techno-



**Figure 1:** Research overview of biogas utilization process.

economic analysis and sensitivity analyses were conducted prior to optimization to assess the overall process parameters. As a result of the optimization, we can recommend the appropriate technologies with defined operating conditions for the process.

## 2. RESEARCH OVERVIEW AND METHODOLOGY

### Research Overview

The section demonstrates the conceptual framework for hydrogen production through the reforming of biogas. Figure 1 provides an overall view of the research on the utilization of biogas, which comprises five steps: Problem setting, superstructure generation, data-driven modelling, development of optimization models, and optimal strategy formulation. To begin the process, we identify the feedstock, types of reformers, and product. The superstructure for the biogas-to-hydrogen process can be generated via various pathways, each involving different types of reformers and operating conditions for the applied technologies. Afterwards, data generated from the process simulation can be utilised to estimate additional mass and energy balance. This is done using the Excel-based data-driven model and is key to calculating techno-economic and environmental parameters. Prior to recommending optimal strategies, an optimization model was created based on diverse evaluation criteria evaluating the impact of economic and environmental factors. The optimization outcomes enable the recommendation of appropriate strategies for the distinct elements of biogas.

We considered three biogas compositions, comprising 55%, 65%, and 75%  $\text{CH}_4$ . To attain the best pathways for each composition, we evaluated four reformers: Steam Methane Reforming (SMR), Dry Reforming (DR), Auto-Thermal Reforming (ATR), and Tri-Reforming (TRI).

In the case of conventional hydrogen production from natural gas, SMR is typically employed in conjunction with water gas shift reactions to maximize hydrogen production. However, the significant quantity of  $\text{CO}_2$  emissions produced by the conventional process necessitates the assessment of the biogas utilization process on various parameters. Steam methane reforming, which has a 3:1  $\text{H}_2/\text{CO}$  ratio, is highly endothermic, whereas dry reforming reduces  $\text{CO}_2$  by consuming it as a reactant. Nonetheless, DR is also an extremely high endothermic process. To reduce the consumption of the reforming process, the ATR that employs partial oxidation with steam methane reforming has emerged. The TRI, which simultaneously employs partial oxidation and dry reforming with steam methane reforming, is a viable technology for biogas utilization. Both ATR and TRI can minimise energy consumption for the reforming process, although  $\text{H}_2/\text{CO}$  ratio may decrease to less than 3. Prior to constructing a biogas utilization process, it is crucial to assess the advantages and disadvantages of various technologies. Table 1 showcases the operating limits of each reformer under atmospheric pressure.

**Table 1:** Types and operating conditions of biogas reforming technologies at atmospheric pressure [5-8].

Steam reforming (Ni/Mg $\text{Al}_2\text{O}_4$ spinel)	Temperature: 600-900°C $\text{H}_2\text{O}/\text{CH}_4$ : 1-3
Dry reforming (Rh/ $\text{Al}_2\text{O}_3$ )	Temperature: 600-1,000°C $\text{CO}_2/\text{CH}_4$ : 0.4-1
Auto-thermal reforming (Ni/Mg $\text{Al}_2\text{O}_4$ )	Temperature: 600-1,000°C $\text{H}_2\text{O}/\text{CH}_4$ : 1-5 $\text{O}_2/\text{CH}_4$ : 0-0.5
Tri-reforming (Ni/Mg $\text{Al}_2\text{O}_4$ )	Temperature: 600-1,000°C $\text{H}_2\text{O}/\text{CH}_4$ : 1-2 $\text{O}_2/\text{CH}_4$ : 0.1-1



## Process Analysis Method

The process was assessed using multiple criteria, including unit production cost (UPC), Net CO<sub>2</sub>-equivalent emissions (NCE), and energy efficiency (EEF), as demonstrated in Eqs. (1)–(3). In the economic evaluation, capital expenditures (CAPEX) and operating expenses (OPEX) were divided by the quantity of produced hydrogen. The UPC was estimated using an interest rate of 7% over a 25-year plant lifespan.

$$UPC = \frac{CAPEX+OPEX}{Amount\ of\ hydrogen}, (\$/kg_{H_2}) \quad (1)$$

In order to calculate the environmental parameters, we considered the net CO<sub>2</sub>-equivalent emissions (NCE). The term 'direct CO<sub>2eq</sub>' refers to greenhouse gas emissions during process operations, whilst 'indirect CO<sub>2eq</sub>' pertains to the emissions caused by the consumption of utilities for the consumption of electricity and heating sources in the process operation. The 'carbon credit' (CC) represents the decrease in the amount of CO<sub>2</sub> achieved by utilizing CO<sub>2</sub> as a reactant.

$$NCE = \frac{Direct\ CO_{2e}+Indirect\ CO_{2e}-CC}{Amount\ of\ hydrogen}, (kg_{CO_2-eq}/kg_{H_2}) \quad (2)$$

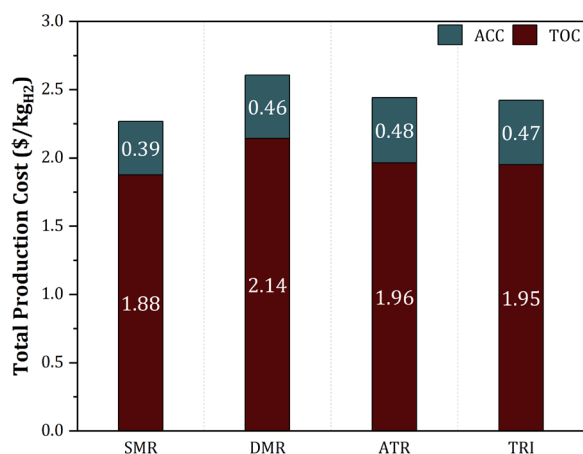
For the assessment of the technical aspects of the biogas-to-hydrogen process, energy efficiency (EEF) serves as an assessment parameter. To determine the EEF of the process, the heat generated by the product (hydrogen) is divided by the energy used by utilities and the heat introduced by the fed feedstock (biogas).

$$EEF = \frac{Heat\ of\ hydrogen}{Heat\ of\ fed\ biogas + E_{utilities}} \times 100\%, (\%) \quad (3)$$

## 3. RESULTS AND DISCUSSION

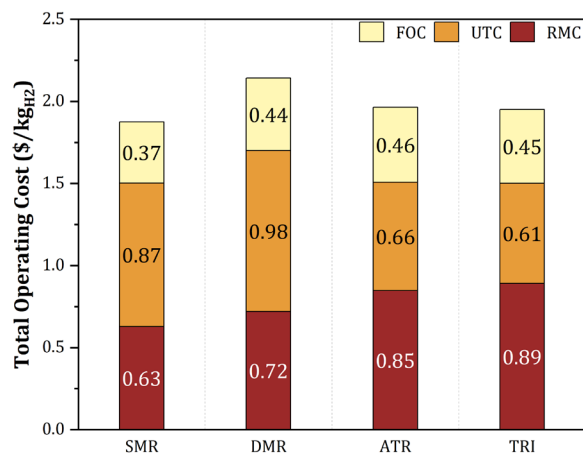
### Techno-economic analysis

For the purpose of economic evaluation, we analysed the breakdown of overall production cost and operational cost for producing hydrogen.



**Figure 2:** Total production cost analysis of biogas-to-hydrogen processes with different reformers.

To address different strategies, we initially studied the optimal operating conditions for each type of reformers. Figure 2 presents the cost breakdown of hydrogen production, indicating that the total operating cost (TOC) represents more than 80% of the total cost for all technologies. To evaluate the impact of OPEX, which is a significant factor in this research, we analysed the specific breakdown of the total operating cost, which is displayed in Figure 3.



**Figure 3:** Analysis of total operating cost of the biogas-to-hydrogen processes.

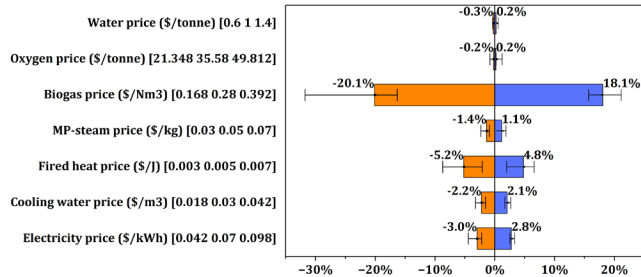
The analysis of total operating costs indicates that utility costs (UTC) dominate the process that employs SMR and DMR for reforming, while raw material costs (RMC) are the most influential parameter for the process using ATR and TRI reactions. It has been determined that highly endothermic reforming reactions yield a relatively larger amount of hydrogen than processes utilizing partial oxidation reactions. Furthermore, we have observed that the process which adopts highly endothermic reforming reactions can produce a greater quantity of hydrogen compared to the partial oxidation reaction process. However, the high heat consumption rate has led to an increased proportion of utility costs in the total operating costs.

### Sensitivity analysis

The sensitivity analysis of the hydrogen UPC was conducted by examining the impact of a 20% change in major economic parameters, as displayed in Figure 4. Optimal conditions for every biogas composition and type of reformer were analyzed, with results presented as a range distribution. The bar graph demonstrates the average values obtained across all cases. Technical abbreviations have been explained within the text.

The results of the sensitivity analysis indicate that the price of biogas (feedstock) has the largest impact on the UPC, decreasing and increasing by 20.1% and 18.1%, respectively. The price of fired heat and electricity are

the second and third most sensitive parameters, resulting in UPC changes of around 5% and 3%, respectively.



**Figure 4:** Sensitivity analysis for the unit production cost of total biogas-to-hydrogen processes.

Other factors have a minor impact on UPC, staying below 2%. Accordingly, we scrutinized that the cost of generating hydrogen is significantly linked to OPEX, particularly the expenses concerning raw materials, fired heat, and electricity. Thus, these costs have a trade-off relation with respect to economic and environmental factors.

### Optimization model

To identify the optimal pathways of biogas-to-H<sub>2</sub>, we developed optimization models using a mixed-integer linear programming (MILP) technique. Here, we considered various evaluation criteria as objective functions to identify the viable biogas utilization strategies. Eq. (1) identifies the minimum unit production cost for a fixed capacity of biogas as a feedstock, which represents the most economic strategy.

$$MinUPC = \sum_j \alpha_j + \sum_j \phi_j + \sum_{i \in I^F} \phi_i F_{ij} + \sum_{i \in I^U} \pi_i U_{ij} \quad (1)$$

where  $\alpha_j$  is the total capital investment cost of pathway  $j$ ,  $\phi_j$  is the fixed operating cost factor of pathway  $j$ , and  $\pi_i$  and  $\phi_i$  are the unit costs for utilities and feedstock, respectively.  $F_{ij}$  and  $U_{ij}$  is the amount of feedstock  $i \in I^F$  and utilities  $i \in I^U$ , respectively.

Eq. (2) seeks for the most eco-friendly process, which is minimum net CO<sub>2</sub> emission strategy for the same amount of process capacity.

$$MinNCE = \sum_j \delta_j + \sum_{i \in I^F} \varepsilon_i U_{ij} - \sum_{i \in I^F} \gamma_i F_{ij} \quad (2)$$

where  $\delta_j$  is the amount of directly emitted CO<sub>2</sub> by technology  $j$ .  $\varepsilon_i$  is the amount of indirect CO<sub>2</sub> emission by using utility  $i \in I^U$ , and  $\gamma_i$  is the CO<sub>2</sub> inventory for feedstock  $i \in I^F$ .

Finally, maximum energy efficient process is identified via Eq. (3), which is maximum energy stored in product with a certain input energy to process.

$$MaxEE = \sum_j \rho_i P_{ij} \quad (3)$$

where  $\rho_i$  is the heating value of final product  $i \in I^P$ .  $P_{ij}$  is the amount of product of pathway  $j$ .

The optimization model with proper corresponding constraints is developed as follows.

**Flow conservation:** Eq. (4) is used to balance the amount of flows between the utilized feedstock and the total amount of processed feedstock in pathway  $j$ .

$$\sum_h F_{ij} = \sum_h Q_{ij}^{in} \quad \forall i \in I^F \quad (4)$$

where  $Q_{ij}^{in}$  is the flow rate of the feedstock input to technology  $j$ .

Similarly, the flow rate of the feedstock input to the technology must be identical to the sum of the three output flow rates for the final product, by-product, and waste, as illustrated in Eq. (5)

$$Q_{ij}^{in} = P_{i'j} + B_{i''j} + W_{i'''j} \quad \forall i \in I^F, i' \in I^P, i'' \in I^W, j \in J \quad (5)$$

where  $P_{i'j}$ ,  $B_{i''j}$ , and  $W_{i'''j}$  are the amount of product, byproduct and waste from pathway  $j$ , respectively.

**Feed availability:** The amount of feedstock should be bounded by an upper limit (i.e., availability) and a lower limit (i.e., minimum purchase for realistic technology operation) as shown in Eq. (6)

$$r_i \leq \sum_j F_{ij} \leq \omega_i \quad \forall i \in I^F \quad (6)$$

where  $r_i$  and  $\omega_i$  are the minimum purchase and feed availability, respectively.

**Logistic constraint:** In case an input can be processed by various technologies, only one technology that can satisfy the objective function should be selected using binary variables.

$$\sum_{j_n} X_{j_n} \leq 1 \quad \forall \{j_1, \dots, j_n\} \in J \quad (7)$$

where  $X_{j_n}$  is a binary variable that represents the pathway selection.

**Technology capacity:** The involved technology is limited by its capacity ( $\psi_j$ ).

$$Q_{j_n}^{in} \leq \psi_j X_{j_n} \quad \forall j_n \in J \quad (8)$$

### Optimal strategies of biogas-to-hydrogen process

The most efficient hydrogen production methods from diverse biogas compositions are evaluated using Figure 5, based on three different criteria. The first column of Figure 5, representing the optimization results, takes into account black utility and levelized cost of biogas (LCOB). According to the estimated process cost by UPC, SMR utilization strategies show the most cost-effective pathways for all the examined biogas

Optimization results: Black utility + LCOB					Optimization: Alternative Scenarios								
Percentage of methane in biogas: (●) 75% (●) 65% (○) 55%					Case1: Green utility			Case2: Biogas price = 0			Case3: Green utility + Biogas price = 0		
Design objective	Reformer type	Value (unit)	Molar flowrate of raw materials (kmol/h)	Operating temperatures (°C)	Design objective	Reformer type	Value (unit)	Design objective	Reformer type	Value (unit)	Design objective	Reformer type	Value (unit)
Min UPC	SMR	2.18 (\$/kg)	H <sub>2</sub> O(R): 550/CO <sub>2</sub> : 0/O <sub>2</sub> : 0/H <sub>2</sub> O(W): 400	T1: 900/T2: 300/T3: 200	Min UPC	SMR	2.78 (\$/kg)	Min UPC	ATR	1.49 (\$/kg)	Min UPC	TRI	1.95 (\$/kg)
Min NCE	TRI	5.43 (kg <sub>CO2</sub> /kg <sub>H2</sub> )	H <sub>2</sub> O(R): 150/CO <sub>2</sub> : 0/O <sub>2</sub> : 287/H <sub>2</sub> O(W): 400	T1: 1,000/T2: 333/T3: 233	Min NCE	SMR	0.41 (kg <sub>CO2</sub> /kg <sub>H2</sub> )	Min NCE	TRI	5.43 (kg <sub>CO2</sub> /kg <sub>H2</sub> )	Min NCE	SMR	0.41 (kg <sub>CO2</sub> /kg <sub>H2</sub> )
Max EE	SMR	53.89 (%)	H <sub>2</sub> O(R): 550/CO <sub>2</sub> : 0/O <sub>2</sub> : 0/H <sub>2</sub> O(W): 1,300	T1: 900/T2: 300/T3: 300	Max EE	SMR	53.89 (%)	Max EE	SMR	53.89 (%)	Max EE	SMR	53.89 (%)
Min UPC	SMR	2.27 (\$/kg)	H <sub>2</sub> O(R): 550/CO <sub>2</sub> : 0/O <sub>2</sub> : 0/H <sub>2</sub> O(W): 400	T1: 825/T2: 300/T3: 200	Min UPC	SMR	2.97 (\$/kg)	Min UPC	TRI	1.64 (\$/kg)	Min UPC	TRI	2.22 (\$/kg)
Min NCE	TRI	6.43 (kg <sub>CO2</sub> /kg <sub>H2</sub> )	H <sub>2</sub> O(R): 150/CO <sub>2</sub> : 0/O <sub>2</sub> : 287/H <sub>2</sub> O(W): 400	T1: 867/T2: 367/T3: 233	Min NCE	DMR	0.56 (kg <sub>CO2</sub> /kg <sub>H2</sub> )	Min NCE	TRI	6.43 (kg <sub>CO2</sub> /kg <sub>H2</sub> )	Min NCE	DMR	0.56 (kg <sub>CO2</sub> /kg <sub>H2</sub> )
Max EE	SMR	53.04 (%)	H <sub>2</sub> O(R): 550/CO <sub>2</sub> : 0/O <sub>2</sub> : 0/H <sub>2</sub> O(W): 1,300	T1: 900/T2: 300/T3: 200	Max EE	SMR	53.04 (%)	Max EE	SMR	53.04 (%)	Max EE	SMR	53.04 (%)
Min UPC	SMR	2.44 (\$/kg)	H <sub>2</sub> O(R): 550/CO <sub>2</sub> : 0/O <sub>2</sub> : 0/H <sub>2</sub> O(W): 400	T1: 750/T2: 300/T3: 200	Min UPC	ATR	3.27 (\$/kg)	Min UPC	SMR	1.84 (\$/kg)	Min UPC	ATR	2.57 (\$/kg)
Min NCE	TRI	7.89 (kg <sub>CO2</sub> /kg <sub>H2</sub> )	H <sub>2</sub> O(R): 150/CO <sub>2</sub> : 0/O <sub>2</sub> : 287/H <sub>2</sub> O(W): 400	T1: 600/T2: 400/T3: 300	Min NCE	DMR	0.65 (kg <sub>CO2</sub> /kg <sub>H2</sub> )	Min NCE	TRI	7.89 (kg <sub>CO2</sub> /kg <sub>H2</sub> )	Min NCE	DMR	0.65 (kg <sub>CO2</sub> /kg <sub>H2</sub> )
Max EE	SMR	50.00 (%)	H <sub>2</sub> O(R): 550/CO <sub>2</sub> : 0/O <sub>2</sub> : 0/H <sub>2</sub> O(W): 1,300	T1: 900/T2: 300/T3: 200	Max EE	SMR	50.00 (%)	Max EE	SMR	50.00 (%)	Max EE	SMR	50.00 (%)

Figure 5: Optimal strategies for different evaluation scenario by considering different biogas compositions.

compositions. The utilization process for low-concentration methane within biogas resulted in an increase in UPC due to the limited production of hydrogen caused by the low levels of methane. By comparing the operating conditions of strategies aimed at minimizing UPC, we can determine that a relatively low-temperature reforming process was utilized for low-concentration methane within biogas. When considering the environmental impact, the TRI utilization strategies exhibit the lowest NCE for all biogas compositions evaluated. The TRI reaction has the potential, in theory, to use CO<sub>2</sub> as a reactant via partial oxidation, thereby supporting the highly endothermic steam methane reforming reaction. The process has the potential to reduce utility consumption through partial oxidation while also reducing CO<sub>2</sub> in the feedstock.

Based on the results of the sensitivity analysis, three parameters that potentially present a trade-off between UPC and NCE when modified can be viewed as constraints for the optimization process. In Case 1, we opted for the use of a green utility with a low CO<sub>2</sub>-eq value despite its higher utilization cost instead of a black utility. Remarkably, the optimal pathway for hydrogen production from biogas has been altered. For all evaluation values (UPC, NCE, and EEF), the utilization of SMR process displays the most viable option for converting biogas-to-hydrogen. The SMR utilization strategy generates significant hydrogen production with minimal CO<sub>2</sub> emissions from the utilities, which is the best pathway for minimizing NCE when the biogas contains 75% concentrated methane. Optimal pathway changes when biogas contains less than 65% methane. The DMR strategies were selected for their eco-friendly nature, reducing significant CO<sub>2</sub> emissions by using CO<sub>2</sub> as a feedstock. Consumption of sustainable, 'green' utilities can reduce the significant CO<sub>2</sub> emissions generated from traditional utilities. In case 2, we assumed a biogas price of 0, which can be obtained

from upcycling waste. There is no significant change in the optimization results, but the utilization processes of ATR and TRI could be more cost-effective pathways compared to the SMR process for biogas with methane concentrations of 55% and 65%, respectively.

Finally, if we consider both assumption that are used in Case 1 and Case 2, the optimization results can be drastically altered. In terms of cost-effectiveness, strategies utilizing partial oxidation reactions to reduce energy consumption are preferred, while SMR and DMR may be the most environmentally friendly processes by reducing CO<sub>2</sub> emissions from utility consumption and utilizing CO<sub>2</sub> as a feedstock. For technology, the SMR-based process remains the most energy-efficient option.

## 4. CONCLUSIONS

In this study, we analysed the process of converting biogas directly to hydrogen and optimised it for three evaluation criteria: UPC, NCE, and EEF. This was achieved by constructing a superstructure that considered multiple variables and evaluation parameters. We determined the economic, environmental, and technical feasibility and suggested optimal pathways for the process. The major findings and contributions of this research are as follows:

- To solve the problem of multi variables and evaluation parameters, the superstructure based data driven-modeling was conducted.
- Before the optimization process, we conducted techno-economic evaluation and sensitivity analysis that identified major cost drivers (i.e., price of feedstock, price of electricity and fired heat).
- For the case of process which utilizes black utilities and LCOB, the SMR based process is economically and technically optimal while TRI

based process shows the best results for environmental parameters for every biogas compositions.

- The analysis for different assumptions which showed trade-off relationship between economical and environmental parameters were conducted.
- By considering every assumptions for alternative scenarios which was shown in Case 3, we can analyze that ATR and TRI based process is cost-effective process for each biogas compositions while SMR and DMR can be the technically effective and eco-friendly process.

and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



## ACKNOWLEDGEMENTS

This work was supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry & Energy (MOTIE) of the Republic of Korea (No.20224C10300040).

## REFERENCES

1. Muradov NZ, Veziroğlu TN. From hydrocarbon to hydrogen carbon to hydrogen economy. *Int J Hydrogen Energy* 30:225-37 (2005).
2. Braga L.B, et al. Hydrogen production by biogas steam reforming: A technical, economic and ecological analysis. *Renewable and Sustainable Energy Reviews*, 28:166-173 (2013).
3. Gao Y, et al. A review of recent developments in hydrogen production via biogas dry reforming. *Energy Conversion and Management*, 171: 133-155 (2018).
4. Zhao X, Joseph B, et al. Biogas Reforming to Syngas: A Review. *iScience* 23: 101082 (2020)
5. Chouhan K, et al. Simulation of steam reforming of biogas in an industrial reformer for hydrogen production. *Int J Hydrogen Energy* 46: 26809-26824 (2021).
6. Takami D, et al. Dry reforming of methane over alumina-supported rhodium catalysts at low temperatures under visible and near-infrared light. *Catal, Sci. Technol*, 10: 5811 (2020).
7. Leandro de Souza T, et al. Thermodynamic analysis of autothermal reforming of methane via entropy maximization: Hydrogen production. *Int J Hydrogen Energy* 39: 8257-8270 (2014).
8. Kenkel P, et al. Biogas Reforming as a Precursor for Integrated Algae Biorefineries: Simulation and Techno-Economic Analysis. *Processes*, 9: 1348 (2021).

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator

# Optimization of Solid Oxide Electrolysis Cell Systems Accounting for Long-Term Performance and Health Degradation

Nishant V. Giridhar<sup>a</sup>, Debangsu Bhattacharyya<sup>a\*</sup>, Douglas A. Allan<sup>b</sup>, Stephen E. Zitney<sup>c</sup>, Mingrui Li<sup>d</sup>, Lorenz T. Biegler<sup>d</sup>

<sup>a</sup> Department of Chemical and Biomedical Engineering, West Virginia University, Morgantown, WV 26506, USA

<sup>b</sup> NETL Support Contractor, Pittsburgh, PA 15236, USA

<sup>c</sup> National Energy Technology Laboratory, Morgantown, WV 26507, USA

<sup>d</sup> Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA

\* Corresponding Author: [debangsu.bhattacharyya@mail.wvu.edu](mailto:debangsu.bhattacharyya@mail.wvu.edu)

## ABSTRACT

This study focuses on optimizing solid oxide electrolysis cell (SOEC) systems for efficient and durable long-term hydrogen (H<sub>2</sub>) production. While the elevated operating temperatures of SOECs offer advantages in terms of efficiency, they also lead to chemical degradation, which shortens cell lifespan. To address this challenge, dynamic degradation models are coupled with a steady-state, two-dimensional, non-isothermal SOEC model and steady-state auxiliary balance of plant equipment models, within the IDAES modeling and optimization framework. A quasi-steady state approach is presented to reduce model size and computational complexity. Long-term dynamic simulations at constant H<sub>2</sub> production rate illustrate the thermal effects of chemical degradation. Dynamic optimization is used to minimize the lifetime cost of H<sub>2</sub> production, accounting for SOEC replacement, operating, and energy expenses. Several optimized operating profiles are compared by calculating the Levelized Cost of Hydrogen (LCOH).

**Keywords:** Fuel Cells, Dynamic Degradation Modelling, Hydrogen, Optimization, Solid Oxide Cells

## INTRODUCTION

Solid oxide cells (SOCs) are flexible energy conversion and H<sub>2</sub> generation devices that have many benefits including modularity, high theoretical efficiency, and low operating greenhouse gas emissions. They can be operated reversibly in either power generation mode as a solid oxide fuel cell (SOFC) or H<sub>2</sub> production mode as a solid oxide electrolysis cell (SOEC). Steam electrolysis in SOECs has a variety of benefits over a conventional steam-methane reforming process such as no direct emission of CO<sub>2</sub> when renewable sources of electricity are used. High-temperature SOC offer high electrical efficiency especially at high temperatures.

However, the high temperature needed for achieving high efficiency in SOC also leads to chemical degradation of the microstructure in the electrodes and electrolytes. Chemical degradation results in changes in the composition and microstructure of the triple-phase

boundary. The dynamics of these changes are very slow, but the changes can steadily build up during the operating lifetime of a cell. Chemical degradation causes an increase in the voltage losses that reduce the effective voltage utilized for electrolysis.

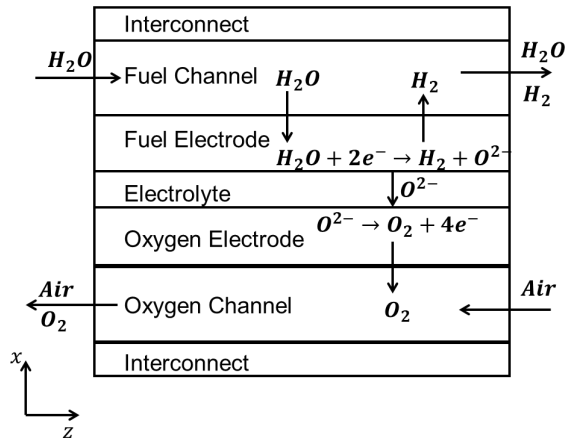
$V_{cell}$ , the cell voltage that contributes to electrolysis, is given by Equation 1. In Equation 1,  $V_{Nernst}$  is the Nernst potential and  $V_{act}$ ,  $V_{Ohmic}$ , and  $V_{conc}$  are the activation, Ohmic, and concentration overpotentials. The magnitudes of these overpotentials depend on the operating conditions of the cell, but these losses are inevitable even in fresh, undegraded cells. The total overpotential due to all long-term chemical degradation mechanisms is given by  $V_{degradation}$ . In a new (i.e., undegraded) cell, the  $V_{degradation}$  term is 0, but this term increases with time.

$$V_{cell} = V_{Nernst} - V_{act} - V_{Ohmic} - V_{degradation} \quad (1)$$

Figure 1 shows a typical SOEC and the reactions



occurring in the fuel and oxygen electrodes. In the oxygen electrode (anode during electrolysis), high partial pressure of oxygen contributes to the oxidation of the active material LSM-YSZ. In the fuel electrode, nickel (Ni) agglomeration occurs due to Ostwald ripening.



**Figure 1.** SOEC Schematic and Reactions

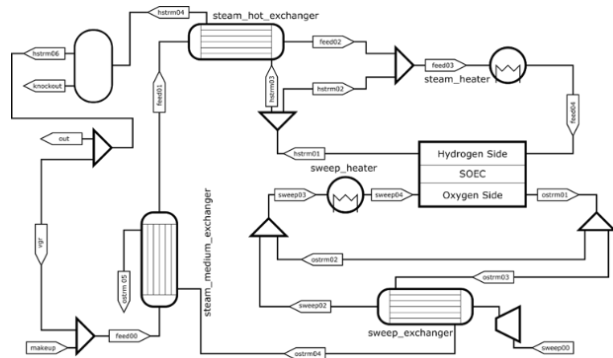
In the literature, many approaches have been proposed to include degradation models for making SOC design and operation decisions. A few relevant optimization studies are reviewed here. For detailed review, we refer readers to [1]. Parhizkar et al. [2] combined microstructure degradation models with a simplified input-output SOFC performance model to perform steady-state optimization to identify operating conditions for minimizing the cost of electricity. Naeini et al. [3] coupled a data-driven model for SOEC degradation with a steady-state model for SOEC operation. In their work, the Levelized Cost of Hydrogen (LCOH) is minimized over a 20-year operating horizon.

In this work, models of key degradation mechanisms are coupled with a first principles two-dimensional (2D) non-isothermal SOEC model. This work seeks to fill two gaps in the existing literature. First, most modeling studies in the literature use isothermal 0D models; therefore, spatial variation and thermal characteristics of a degraded cell are neglected. The 2D non-isothermal SOC model used in this work provides detailed insights into the spatial distribution of degradation in each cell. A second gap, which this work aims to fill, is that most SOC degradation simulation studies do not consider the effects of degradation on the balance of plant. While some studies incorporate simplified heating and compression duties in their calculations for total energy consumption, this work extends this to capture the impact of cell level degradation on the balance of plant by considering a fully modeled SOEC flowsheet. The flowsheet model is used in dynamic optimization to identify optimal operating profiles for system- and cell-level decision variables as a function of time.

## MODELING AND METHODS

### Flowsheet modeling details

As shown in Figure 2, the SOEC hydrogen production system is modeled as a collection of linked unit models in the open-source IDAES, equation-oriented IDEAS (Institute for the Design of Advanced Energy Systems) platform. The system consists of an SOEC stack (modeled as a single cell), and electric heaters for the feed and sweep input streams along with heat exchangers for heat recovery. For a detailed description of SOEC flowsheet modeling details, we refer readers to [4-5]. The models used in this study are available in the open-source IDAES modeling framework [9].



**Figure 2.** SOEC system flowsheet

### Health modeling

Degradation mechanisms in SOCs vary depending on the materials of construction of the cell. For this study, we consider fuel electrode of a Ni-YSZ composite, an oxygen electrode of LSM-YSZ separated by a YSZ electrolyte as a part of a planar fuel electrode supported SOC.

One of the dominant degradation mechanisms that take place in the fuel electrode is the agglomeration of Ni particles due to surface diffusion. The dynamics of the Ni particle growth are described by Equation 2 [6]. Here,  $\bar{d}_{Ni}$  is the average Ni particle diameter,  $m = 0.5$  and  $n = 8$ .

$$\frac{d(\bar{d}_{Ni})}{dt} = \frac{a'}{d_{Ni}^{n-1}} \left( \frac{P_{H_2O}}{P_{H_2}^{0.5}} \right)^m e^{-\frac{E_a}{RT}} \quad (2)$$

The oxygen electrode undergoes degradation due to the growth of oxide scales under high oxygen partial pressure. Under these conditions, Wagner's law for parabolic oxidation is used as a semi-empirical model to describe growth of oxide scale. In LSM-YSZ electrodes, the scales formed are of lanthanum zirconate (LZO) and chromium oxide (COS). The chromium oxide scale is localized to the interface between the oxygen electrode and interconnect. Equation (3) describes the temperature-dependent Wagner's law for a general oxide scale growth length  $l_{sc}$  [7]. Here,  $X_{0,sc}$ ,  $\rho_{sc}$ ,  $K_{sc}$ , and  $E_{sc}$  are the weight fraction, density, weight gain rate, and activation

energies of oxide scale growth, respectively.

$$\frac{dl_{sc}}{dt} = \frac{K_{sc}}{2(X_{0,sc}\rho_{sc})^2} e^{-\frac{E_{sc}}{RT}} \quad \forall sc \in \{COS, LZO\} \quad (3)$$

Other degradation mechanisms that are modeled include coarsening of the LSM-YSZ oxygen electrode and phase transformation of the YSZ electrolyte. Additionally, we make use of property models to relate microstructure parameters to degradation overpotentials [8].

## Solution methodology

Incorporating degradation models into first principles unit models requires solving a differential algebraic equation (DAE) system with multiple timescales. For the SOEC system, these timescales can be broadly divided into two categories: one faster timescale for process operations and SOEC dynamics (minutes to hours) and another slower timescale for degradation (hundreds of hours). To simplify the problem, we assume quasi-steady state for the faster timescale processes. This allows us to discretize and dynamically integrate the slower degradation timescale. With this approach, we can simulate SOEC operation with degradation effects for extended periods, spanning tens of thousands of operational hours, providing insights into long-term performance and design strategies.

We employ the implicit Euler method to discretize the degradation dynamic equations. The discretized system of equations can be solved simultaneously and efficiently using a nonlinear solver such as IPOPT [10]. Because of the quasi-steady state assumption, the only active differential equations in the model are those pertaining to the slow degradation dynamics. The advantage of a simultaneous solution method is the ability to optimize across the entire operational horizon, ensuring a comprehensive assessment of system performance. Additionally, this approach enables the incorporation of constraints spanning multiple time points, thus providing enhanced modeling versatility. In contrast, the sequential time-stepping approach restricts optimization to instantaneous performance metrics like instantaneous efficiency.

## Hydrogen Production Profiles

The discretized SOEC system model is used to assess different long-term operating scenarios by first specifying the nature of the hydrogen production profile.

**Constant H<sub>2</sub> Production:** In this profile, a constant H<sub>2</sub> production rate is maintained throughout the operating lifetime. As shown in Equation 4, an equality constraint indexed over time ensures that the average current density  $J_{avg}$  remains constant. Here,  $j_{t,z}$  is the local current density at a point  $z$  along the length of the cell and  $N_z$  is the number of finite elements. As degradation proceeds,

it is expected that increased degradation losses will result in an increase in operating voltage to sustain the fixed H<sub>2</sub> production rate.

$$\frac{1}{N_z} \sum_z j_{t,z} = J_{avg,0} \quad \forall t \quad (4)$$

**Constant Potential Operation:** In this strategy, a constant cell potential is maintained over time [3]. This keeps cell-level energy consumption constant. However, as degradation progresses, current density decreases. This leads to a gradual reduction in H<sub>2</sub> production at a constant operating temperature. However, by varying the cell operating temperature over time, more complex operating scenarios can be investigated. Cell voltage is held constant by fixing operating voltage for all time points (Equation 5).

$$V_{cell,t} = V_{cell,0} \quad \forall t \quad (5)$$

## Dynamic optimization

Typically, steady-state optimization is used to determine the operational setpoints of a system. In this case, minimization of the system power requirement for a given H<sub>2</sub> production rate [5]. This methodology can be used to determine steady-state optimal operating conditions corresponding to different load conditions. In the absence of degradation, operational efficiency, power consumption, and local cell temperature profile would remain at the steady-state optimal value. In this work, we provide a methodology to optimize the operational profile over 20,000 hours for long-term stability and efficient operation, while accounting for performance degradation.

For long-term dynamic optimization, we propose to maximize the integral average efficiency over the operating lifetime (Equation 6). As a surrogate for system efficiency, we use a simplified version as described in [11]. Under constant H<sub>2</sub> load, the numerator is invariant, and the objective simplifies to the minimization of system power consumption. Obviously, under constant potential operation, the numerator is not constant. Here,  $P_{system}$  is the total electric power consumption of the system, including stack work, electric heater, and blower duties (Equation 7).

$$\max \sum_{t \in \text{time}} \frac{HHV(\dot{m}_{H_2,t})}{P_{system,t}} \quad (6)$$

$$P_{system} = P_{SOEC} + P_{sweep\ heater} + P_{feed\ heater} + P_{blower} \quad (7)$$

A second operational strategy involves minimization of the degradation rate. To include all degradation effects into one combined term, it is desirable to use the degradation overpotential  $V_{degradation}$  in Equation 1. However, since the overpotential at any given point in time

**Table 1:** Summary of Optimization Results

H <sub>2</sub> Production Profile	Objective Function	Avg. H <sub>2</sub> Production Rate (× 10 <sup>7</sup> kg/yr)	Sp. Energy Consumption (kWh/ kg H <sub>2</sub> )	Voltage Degradation Rate (%/ khr)	Replacement Time (years)	LCOH (\$ /kWh)
Constant H <sub>2</sub> Production Rate	Maximize Integral Efficiency	4.74	36.6	4.4	1.5	0.31
	Minimize Degradation Rate	4.74	40.4	1.5	4.0	0.33
Constant Potential Operation	Maximize Integral Efficiency	3.39	35.6	2.7	2.0	0.29
	Minimize Degradation Rate	1.31	41.7	0.7	8.0	0.37
	Maximize H <sub>2</sub> Production	3.93	38.9	5.0	1.5	0.34

depends on the operating temperature and composition following an Arrhenius-type temperature dependence, we consider the voltage at a reference temperature  $\tilde{V}$ . This ensures that the voltage degradation rate for optimization only depends on the operating trajectory. With this we minimize the average voltage degradation rate from  $t=0$  to  $t=t_f$ , the final time point (Equation 8).

$$\min \frac{\tilde{V}_{t_f} - \tilde{V}_0}{\tilde{V}_0} \quad (8)$$

Finally, under constant potential operation, market conditions may require the system to operate at the highest possible H<sub>2</sub> production rate ( $\dot{m}_{H_2}$ ). This can be achieved by maximizing the H<sub>2</sub> production rate (Equation 9).

$$\max \sum_{t \in \text{time}} \dot{m}_{H_2,t} \Delta t \quad (9)$$

The decision variables for optimization include feed and trim heater duties, recycle ratios, blower flowrate, sweep and steam inlet and outlet temperatures. For details on bounds and operational constraints refer to [5].

### Levelized Cost of H<sub>2</sub> (LCOH) Calculation

The different optimized cases are compared based on the resultant LCOH, as shown in Equation 10. We compute the LCOH by modifying the method of [12] to incorporate stack replacement costs. The LCOH includes the capital recovery factor (CRF) and capital costs (CC) for the stack and balance of the plant (BOP).  $OC$ ,  $EC$ , and  $m_{H_2, \text{lifetime}}$  are the operating costs, energy costs, and lifetime H<sub>2</sub> production, respectively. Costing parameters are obtained from [12].

$$LCOH = \frac{CRF_{BOP} CC_{BOP} + \sum_{i=1}^R CRF_{\text{stack},i} CC_{\text{stack}} + OC + EC}{m_{H_2, \text{lifetime}}} \quad (10)$$

The number of stack replacements ( $n_{rep}$ ) over the plant lifetime of 30 years is computed based on the extent of degradation at the end of the 20,000-hour optimization

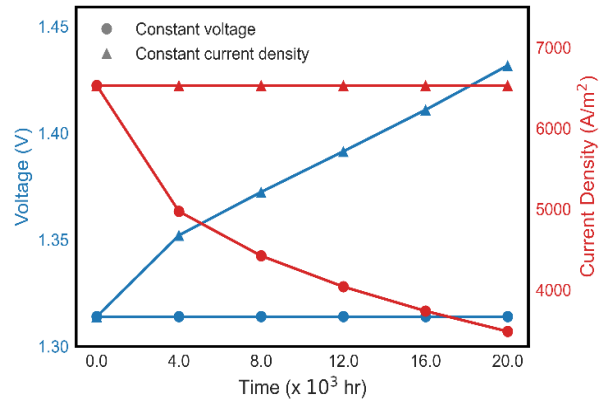
horizon as per Equations 11 and 12.

$$\text{Replacement time} = \frac{\Delta \tilde{V}_{limit}}{\left(\frac{\Delta \tilde{V}}{\Delta t}\right)_{avg}} \quad (11)$$

$$n_{rep} = \frac{30 \text{ years}}{\text{replacement time}} \quad (12)$$

## RESULTS

Figure 3 compares 20,000 hours of operation, at constant current density and constant potential through simulations. As the degradation overpotential increases over time, we observe that constant current density operation results in a higher voltage degradation rate than constant potential operation. Due to the increase in Ohmic resistance, the thermal characteristics along the length of the cell change with time.

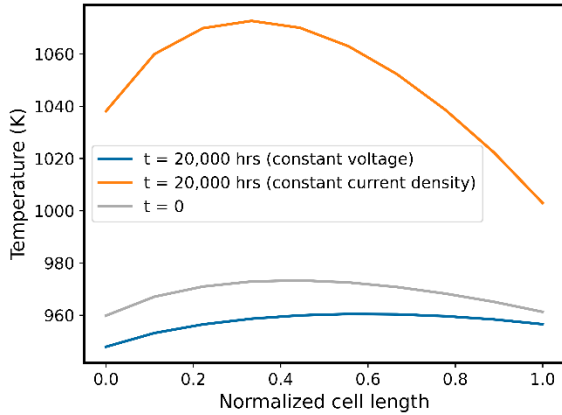


**Figure 3.** Effect of degradation on voltage and current density

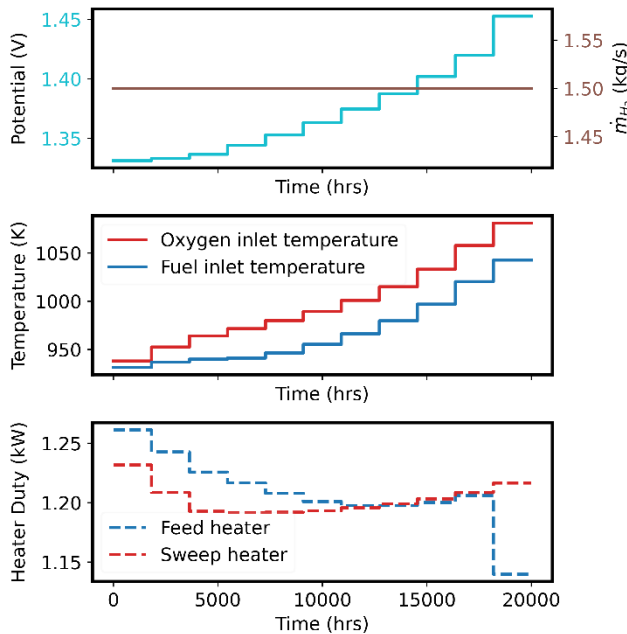
As shown in Figure 4, there is a significant increase in the average temperature along the length of the cell for constant current density operation. Here the inlet temperatures were held constant over the 20,000 hours

of operation.

The spatial temperature profile shown is at the center of the cell along the electrolyte layer. During constant current density operation, we observe a large increase in local cell temperature. The increase in degradation over-potential results in an increase in operating potential, which corresponds to operation above the thermoneutral voltage. Furthermore, the larger Ohmic resistance due to degradation also results in higher Ohmic heating. Under constant potential operation, a drop in current density results in lower cell temperatures.



**Figure 4.** Change in temperature profile due to degradation



**Figure 5.** Optimal Constant H<sub>2</sub> Production Operation (Maximize Integral Efficiency)

Table 1 compares key performance measures from the different dynamic optimization case studies. These include the average H<sub>2</sub> production rate, specific energy

consumption, and average degradation rate as calculated in Equation 13.

$$\text{Voltage Degradation Rate} = \frac{1}{t_f} \frac{\bar{V}_t - \bar{V}_0}{\bar{V}_0} \times 100 \quad (13)$$

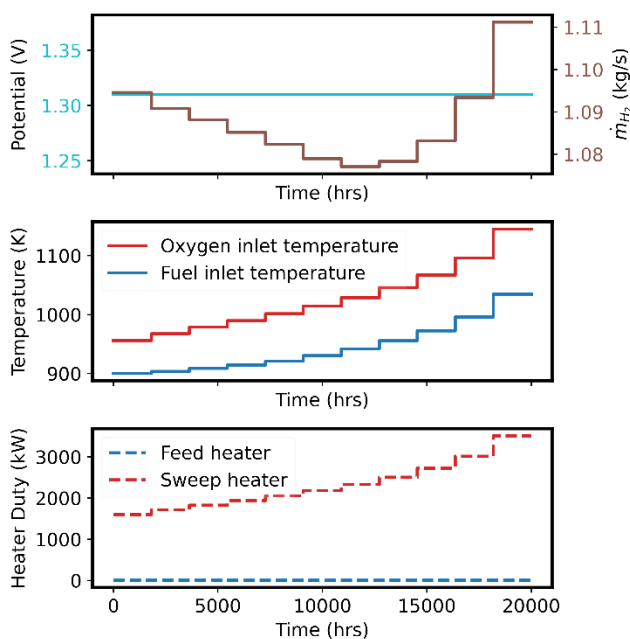
The initial operating point for all cases is identical, and the specific operating conditions were obtained by a steady-state optimization to ensure a H<sub>2</sub> production rate of 1.5 kg/s. This corresponded to an optimal initial voltage of 1.31 V. For more details on the steady-state optimization, refer to [5].

When operating at a constant H<sub>2</sub> production rate, the choice of operating objective can result in significant changes in specific energy consumption and degradation. In this work, we consider the fixed H<sub>2</sub> production rate achieved by constant current density operation to be the upper limit of H<sub>2</sub> production by the system. Based on the costing methods used in this study, a lower LCOH is obtained when integral system efficiency is maximized. Figure 5 describes the temporal profiles of the key decision variables to achieve the maximum integral efficiency. Due to increased voltage losses caused by degradation, operating potential must be increased to maintain the initial H<sub>2</sub> production rate ( $\dot{m}_{H_2}$ ) of 1.5 kg/s. Since inevitable losses depend directly on temperature, the optimal inlet temperatures for both the fuel and oxygen electrodes must be increased to maintain voltage efficiency. Feed and sweep heater duties are kept low due to the presence of recycle streams. Therefore, internal stack resistive heating is used instead of electric heating in the trim heaters. In contrast, degradation rate can be minimized by operating the cell at low temperatures. This results in a significant increase in inevitable losses and, consequently, there is a drop in efficiency as evidenced by the increase in specific energy consumption.

Under a constant potential operating regime (Figure 6), it is observed that the degradation rate depends heavily on the specific operating profile even when operating under constant potential. Under a constant potential over the operating lifetime of the cell, the lowest degradation rate is achieved by operating the cell at low temperatures and, consequently, a very low H<sub>2</sub> production rate (~70% reduction compared to constant H<sub>2</sub> production). Consequently, the specific energy consumption and LCOH when the degradation rate is minimized at constant potential is the highest of all the cases. On the other hand, if operation is optimized for maximum H<sub>2</sub> production under a constant operating potential, we observe a significantly higher degradation rate. At any point in time to maximize H<sub>2</sub> production rate at a constant operating potential, the cell must be operated at the upper bounds of the temperature. Therefore, high temperature ensures a high instantaneous H<sub>2</sub> production rate at all points in time. Consequently, the degradation rate is significantly higher than in all other cases, including those

where  $H_2$  production is held constant.

Figure 6 describes the optimal operating profiles for constant voltage with the objective to maximize integral efficiency. While keeping potential constant, initially it is preferred to operate at a decreasing  $H_2$  production rate to minimize degradation in the early periods of a cell's lifetime. Once degradation in the cell reaches a certain level, it is observed that the  $H_2$  production rate increases again. This is likely because of a phenomenon known as cell break in, resulting from the different rates of degradation of the electrode materials. For example, in Ni-YSZ fuel electrodes, there is a coarsening of Ni particles, while the YSZ particles remain unchanged. The YSZ backbone therefore has a limiting effect on the maximal extent of Ni degradation. Similarly, the parabolic growth rate of oxide scales in the oxygen electrode decreases with the increase in the extent of the oxide scale. To evaluate all optimization formulations and  $H_2$  demand profiles, the LCOH was computed for the five cases. For details on LCOH calculations and costing methodology, refer to [12]. Results presented in the paper correspond to an electricity price of 0.30 \$/kWh. According to this costing methodology, operating at a constant potential and maximizing integral average efficiency leads to the lowest LCOH. The cases with the highest LCOH are those with high energy consumption, where the degradation rate is explicitly minimized. This indicates that system efficiency and  $H_2$  production rate have a higher impact on LCOH than stack replacement costs.



**Figure 6.** Optimal Constant Voltage Operation (Maximize Integral Efficiency)

## CONCLUSION

In this work, we present a dynamic optimization methodology to obtain time dependent operating points for an SOEC system considering the effects of chemical degradation on operating efficiency. The quasi-steady state assumption enables model size reduction to perform dynamic optimization on a fully discretized DAE model.

The 2D non-isothermal SOEC model used in this study captures the thermal implications of chemical degradation. In addition to an increase in electrochemical losses, it is found that chemical degradation results in an overall increase in exothermicity due to an increase in resistive heating. The balance-of-plant model facilitates accurate estimation of SOEC system energy requirements from auxiliary equipment such as heaters and blowers.

Dynamic optimization of long-term operating conditions over 20,000 hours was performed under two  $H_2$  production profiles; namely, constant  $H_2$  production and constant potential operation. This study shows that both operating profiles can result in efficient long-term operation depending on the optimization objective. Strong trade-offs between the integral average operating efficiency and the degradation rate are observed. LCOH is computed for each of the cases and is used to evaluate the different operating trajectories. Operating at a constant potential while maximizing the integral efficiency is found to have the lowest LCOH at 0.29 \$/kWh  $H_2$  with a replacement schedule of 2 years. It must be noted that the LCOH calculation depends on many factors such as the interaction between stack replacement costs and the price of electricity. In a situation where stack replacement costs are higher, it is possible that lower degradation rates would be favorable. Furthermore, this study does not consider the impact of operation on physical degradation. This poses the challenge of incorporating operating profiles from a shorter timescale of daily operation.

## ACKNOWLEDGEMENTS

This work was conducted as part of the Institute for the Design of Advanced Energy Systems (IDAES) with support from the U.S. Department of Energy's Office of Fossil Energy and Carbon Management through the Simulation-based Engineering Research Program.

This project was funded by the Department of Energy, National Energy Technology Laboratory an agency of the United States Government, through a support contract. Neither the United States Government nor any agency thereof, nor any of its employees, nor the support contractor, nor any of their employees, makes any warranty, expressor implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process



disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## REFERENCES

1. Peng, J., Zhao, D., Xu, Y., Wu, X., & Li, X. Comprehensive analysis of solid oxide fuel cell performance degradation mechanism, prediction, and optimization studies. *Energies*, 16(2), 788 (2023).
2. Parhizkar, T, and Roshandel, R. Long term performance degradation analysis and optimization of anode supported solid oxide fuel cell stacks. *Energy conversion and management* 133: 20-30 (2017).
3. Naeini, M., Cotton, J. S., & Adams, T. A. An eco-technoeconomic analysis of hydrogen production using solid oxide electrolysis cells that accounts for long-term degradation. *Frontiers in Energy Research* 10, 1015465 (2022).
4. Bhattacharyya, D., Rengaswamy, R., & Finnerty, C. Dynamic modeling and validation studies of a tubular solid oxide fuel cell. *Chemical Engineering Science*, 64(9), 2158-2172 (2009).
5. Allan, D.A., Dabadghao, V., Li, M., Eslicka, J.C., Ma, J., Bhattacharyya, D., Zitney, S.E. and Biegler, L.T. NMPC for Setpoint Tracking Operation of a Solid Oxide Electrolysis Cell System. National Energy Technology Laboratory (NETL), Pittsburgh, PA, Morgantown, WV, and Albany, OR (United States) (2023).
6. Sehested, J., Gelten, J. A., & Helveg, S. Sintering of nickel catalysts: Effects of time, atmosphere, temperature, nickel-carrier interactions, and dopants. *Applied Catalysis A: General*, 309(2), 237-246 (2006).
7. Kamkeng, A. D., & Wang, M. Long-term performance prediction of solid oxide electrolysis cell (SOEC) for CO<sub>2</sub>/H<sub>2</sub>O co-electrolysis considering structural degradation through modelling and simulation. *Chemical Engineering Journal*, 429, 132158 (2022).
8. Mason, J., Celik, I., Lee, S., Abernathy, H., & Hackett, G. Performance degradation predictions based on microstructural evolution due to grain coarsening effects in solid oxide fuel cell electrodes. *Journal of The Electrochemical Society*, 165(2), F64 (2018).
9. Lee A., Ghouse JH, Eslick JE, et al., The IDAES process modeling framework and model library— Flexibility for process simulation and optimization. *J Adv Manuf Process* 3(3): e10095 (2021)
10. Wächter, A., & Biegler, L. T. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical programming*, 106, 25-57 (2006).
11. Harrison, K. W., Remick, R., Hoskin, A., & Martin, G. D. Hydrogen production: fundamentals and case study summaries (No. NREL/CP-550-47302). National Renewable Energy Lab.(NREL), Golden, CO (United States) (2010).
12. Eslick, J., Noring, A., Susarla, N., Okoli, C., Allan, D., Wang, M., ... & Burgard, A. Technoeconomic Evaluation of Solid Oxide Fuel Cell Hydrogen-Electricity Co-Generation Concepts (No. DOE/NETL-2023/4322). National Energy Technology Laboratory (NETL), Pittsburgh, PA, Morgantown, WV, and Albany, OR (United States) (2023).

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Power System Design and Necessary Changes to Accommodate Future Energy Challenges

Iiro Harjunkoski<sup>a\*</sup>, Katarina Knezovic<sup>b</sup>, and Alexandre Oudalov<sup>c</sup>

<sup>a</sup> Hitachi Energy Research, Mannheim, Germany  
<sup>b</sup> Hitachi Energy Research, Baden-Dättwil, Switzerland  
<sup>c</sup> Hitachi Energy Ltd., Zürich, Switzerland

\* Corresponding Author: [iiro.harjunkoski@hitachienergy.com](mailto:iiro.harjunkoski@hitachienergy.com).

## ABSTRACT

The decarbonization of the society has a very high effect on the power grids as especially the energy generation will be almost completely shifted to CO<sub>2</sub>-neutral sources such as wind and solar. This implies significant design changes to the power grids and power systems, which lie between the electricity producers and consumers. In this paper, we discuss both the generation and consumer side, including the grid changes and required data exchange to support the transition.

**Keywords:** Electricity & Electrical Devices, Energy Conversion, Renewable and Sustainable Energy, Energy Systems, Power Grid

## INTRODUCTION

Traditionally, power systems have been relatively straightforward: Control a set of electricity generation units balancing the demand and transfer the needed

power to the consumer – one way. This has been a very robust design scheme between the energy providers and consumers, also from the controllability point of view. However, with the introduction of new energy technologies (wind, solar, distributed energy resources,

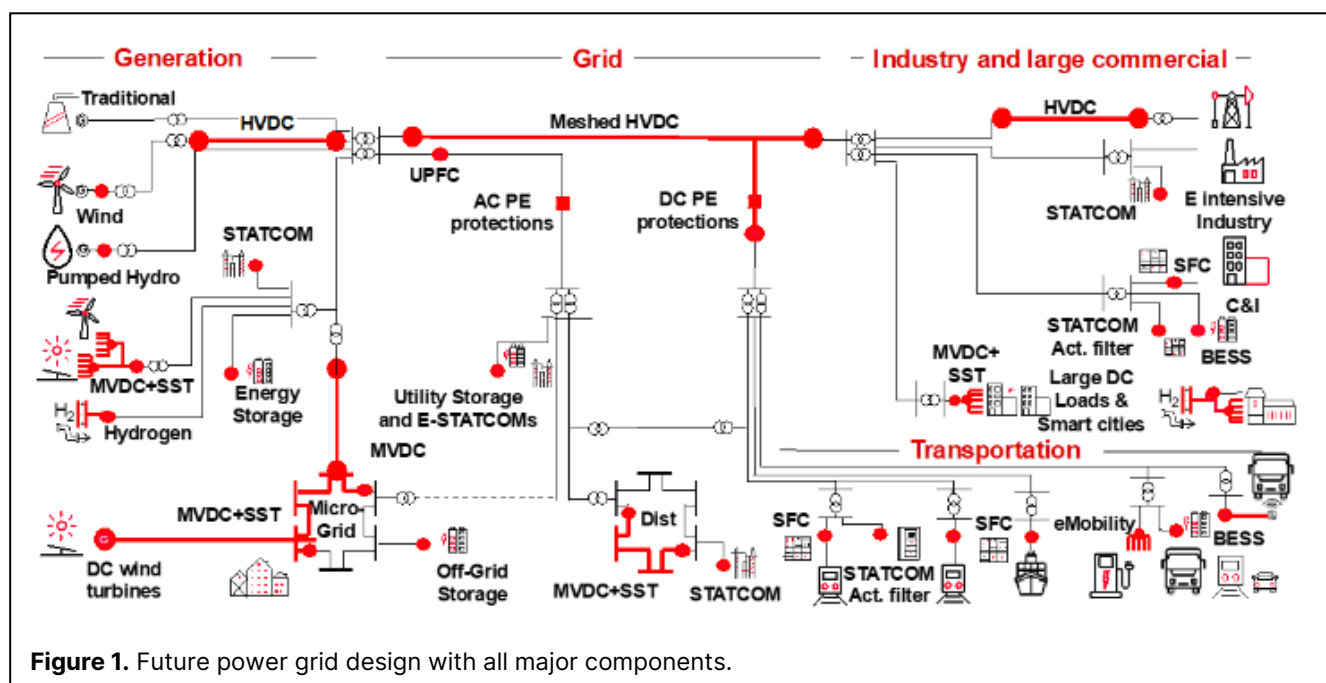


Figure 1. Future power grid design with all major components.

demand-side management, hydrogen, etc.), the traditional design paradigm is no longer sufficient [1][2].

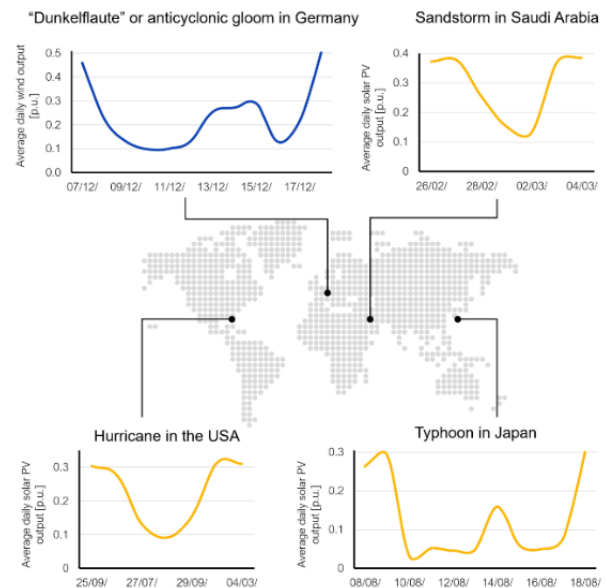
Deregulation of electricity markets and the increased awareness of the climate change have massively increased the share of weather-dependent renewable generation and the need for more flexible power systems. Consequently, due to the volatility of renewable sources, not only are the electricity grids pushed to their boundaries, but there is also a stronger need to create designs where process systems and power systems meet and become strongly integrated. One concrete example is sector coupling, in which various processes are utilized to transfer between electricity and other energy carriers (e.g. ammonia and hydrogen).

**Figure 1** depicts the structure and different components of a future power system. The focus of this paper is on highlighting and discussing potential design changes and challenges involving hardware to enable stronger grids with higher and more intelligent electricity transfer volumes, as well as processes for conversion between electricity and hydrogen. To realize these, the related investments must be both economically and practically feasible to ensure the needed flexibility, and enable secure and sustainable energy supply in all situations, while avoiding stranded infrastructure costs.

In this paper we follow the main structure of **Figure 1** and go into details of the generation and grid aspects. We also discuss the role of industry and large commercial electricity users, as well as the data exchange as an enabler of future collaboration and communication.

## GENERATION SIDE

In short, due to the electrification of industrial processes and transportation sector as well as fast growing share of distributed renewable generation, we can expect that the total global installed electricity generation capacity will increase by a factor of four and the related electricity transfer capacity by a factor of three until 2050. Maybe the most visible changes that have been penetrating the energy landscape in the last decades are on the generation side. To reduce the CO<sub>2</sub>-emissions and due to safety concerns, energy generation has shifted towards variable renewables, mainly utilizing solar and wind power. Many coal-fired and CO<sub>2</sub>-free nuclear power plants have been retired and replaced by significant renewable capacity. In practice, this means going from dispatchable to non-dispatchable energy generation, i.e. we cannot anymore decide and fully control how much and when to generate. The only long-term “traditional” CO<sub>2</sub>-free and flexibly dispatchable electricity generation is hydro power resources, including pump hydro storage. However, as it is both geographically limited and subject to societal constraints,



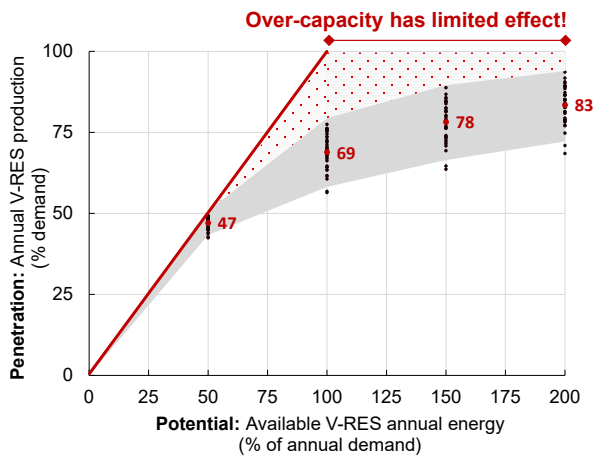
**Figure 2.** Examples of global weather-dependent disturbances on renewable production.

this “green” transition has already presented significant changes in the electricity supply, mainly observable through significant price changes and spikes during energy scarcity.

The imminent change to the power grid is that maintaining the stability (ensuring and controlling the grid frequency and voltage) becomes more challenging as there is less inertia, dynamic voltage support and sufficient short circuit capacity, e.g., from synchronous rotating machinery present in the system. If, for instance 50% of the electricity generation comes from renewable sources, of which solar power is guaranteed to produce 0 MW in the nighttime and wind power production is highly weather-dependent, it is of great importance to design methodologies for how to deal with times where renewable production is low or even non-existing.

**Figure 2** shows examples across the globe where extreme weather events had an unexpected and significant negative impact on the renewable generation. One approach is to increase the capacity, but this will not solve the root cause of the problem. Another approach is to investigate novel electricity generation technologies such as microreactors, wave power, solar power provided in space and many more recently discussed ideas or solutions at a very early prototype or conceptual phase.

The main target is to be able to deal with the high volatility (worst case is a cold and windless winter night where the energy demand is high) and secure energy supply also through those times where renewable energy is temporarily unavailable [3]. As depicted in **Figure 3**, investing in generation overcapacity is not sufficient



**Figure 3.** Actual use of versus potentially available wind and solar generation (dots represent different regions each with an optimal mix of solar and wind generation). Red numbers show the average penetration percentage based on annual generation and demand match with hourly resolution. Static and dynamic system limits will push the grey zone further down.

since it does not solve the underlying problems of spatial and temporal mismatches between the generation and demand. A possible solution could also lie in process design by making processes more resistant to short- or mid-term energy scarcity. Currently, much research is focused on energy storage options [4], where most alternatives can only provide hours or maximally a few days of capacity. Even the largest electrochemical energy storage (e.g. lithium-ion batteries) can only serve short-term demand and therefore alternatives vary between e.g. gravitation-based energy storage, power electronics, compressed air, molten sand batteries and sector coupling.

Sector coupling, i.e. switching to alternative energy carriers such as hydrogen or ammonia is still in its very early stages. There are still challenges to overcome, e.g. the low round-trip efficiency of 35% for hydrogen and the fact that upfront investment costs are currently quite high. However, this technology is attractive from a few points of view.

1. Alternative energy carriers are relatively stable and can in principle with sufficient capacity help bridging across low renewable power periods.
2. They can also be distributed along the power grids to end consumers, naturally with much slower dynamics (pipelines, ships, trucks). However, a compatible infrastructure is still missing.
3. They can be stored at various sites in small amounts reducing the regional risks of energy scarcity. Larger

storage capacities are geographically constrained.

4. Apart from pure energy storage, many of them can also be used as raw-material or produced as by-product of essential industrial processes.

Especially the last point is worth considering as without re-electrification the losses are much smaller. It is still unclear how the related conversion processes should be designed and embedded into an effectively working energy system. The ongoing overall development also implies that generation will become more local (solar rooftop etc.) and therefore the grid structures must also be adapted to meet the ongoing changes.

## GRID SIDE

As already stated in the introduction, the increased generation also results in roughly three times more energy transmission capacity needed by 2050. Due to this, there are many design changes necessary in the power grids, both on the level of transmission grids that are responsible for the long-distance electricity transfer and connected to large-scale power plants, as well as distribution grids, which operate on a lower voltage level and connect to local consumers and prosumers (e.g. households with solar rooftop and energy storage systems) [5]. The increasing power production within the distribution grid must be supported by stronger local grids. Already today, there are situations where the generation or load must be curtailed in order to avoid grid congestion or voltage issues.

Network switching within a distribution grid, i.e. changing the network topology, can also help in avoiding some congestions. Here, the connection points to the transmission grid may also be changed. Thus, in many situations a closer and faster coordination between the distribution and transmission grids with different voltage levels is required. As much of the energy generation is in fact done in direct current (DC) (e.g. solar power), it must be converted to alternating current (AC) through power converters. Here, many discussions on building pure local DC grids and experimental microgrids for this exist [6]. This could also affect the industry electrification as many processes also utilize direct current.

One of the perhaps largest changes in the transmission grids is caused by the long transmission distance needs between remote renewable generation and the major load centers (often > 1000 km, the longest one of 3300 km built in China). Here a large capacity is transferred and, in order to reduce the power losses and keep the voltage drop within acceptable range, this should take place in high voltage and preferably using high voltage direct current (HVDC) technology, which has significantly lower losses compared to traditional AC-

grids. This also enables the transfer of very high capacities of e.g. 12 GW. Investment costs in HVDC grids and converters are not only very high, but together with other grid improvement measures, including more flexible switching and power converters, also open up new dimensions that require a higher level of automation to work seamlessly. The power network should be continuously monitored and seamlessly adapted to changing situations as well as be able to minimize the impact of local disturbances trying to isolate blackouts to as small areas as possible.

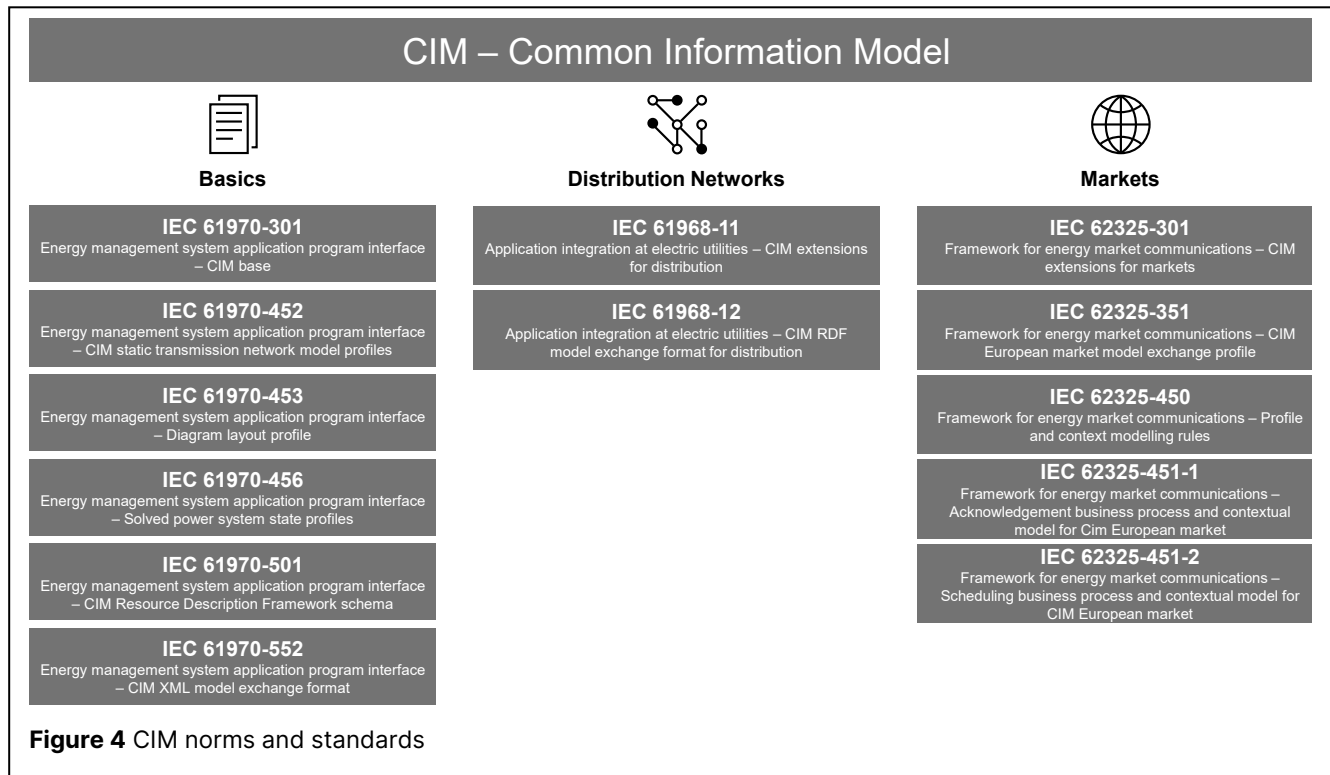
In order to both build more safety buffers for the grid, as well as to maximize the utilization possibility of intermittent renewable generation, we also need energy storage options to support short-term network imbalances and long-term seasonal variations. All of these must naturally also be connected to the power grid (be it transmission or distribution), and may also include sector coupling that calls for simultaneous coordination between multiple energy carriers, e.g. electricity, ammonia, hydrogen and the respective processes managing the conversion.

As the energy supply chain gets increasingly concentrated around electricity, it also becomes a possible target for various threats. Cyber security should be significantly strengthened to ensure a fully functional power grid in all situations – also when the automation systems are under attack.

## INDUSTRY AND LARGE COMMERCIAL

Industry and large commercial are typically consuming around 50% of the total energy (region-dependent) and will play a significant role in the new energy system since we can expect more industrial processes being electrified. With the increasing volatility as well as higher electricity demand, these sectors must provide more flexibility by design, i.e. become prosumers (consumers that can also at least virtually act as producers by temporarily lowering their energy demand). Typically, we refer to this as industrial demand-side management [7]. The main idea is to be able to plan the processes in a way that provides more flexibility towards their electricity need, often with the main objective of reducing the electricity costs. This also gives the power grid operators the opportunity to provide incentives to the market participants through dynamic pricing schemes. Enabling the capability to participate may require significant and expensive changes in process design, as e.g. can be seen in the ongoing electrification of steel industry [8].

Once processes are adapted to provide more flexibility, the next natural step would be to create collaboration opportunities to support large electricity fluctuations with lower impact on the individual processes. Here, we can also expect to see more “isolated” or per need autonomous microgrids that can maintain all critical operations even during a major blackout. These could comprise communities, but also critical functions such as hospitals or industrial sites that





are strongly affected by any power disruptions. There is already much research on how to realize this, e.g. using smart cities concepts with rural areas [9] and university campuses as testing platforms. Nonetheless, it is still an open question how to realize this on a larger scale.

What is unquestionable is that certain autonomy requires local energy generation – even power plants for securing times without renewable generation. To further ensure a CO<sub>2</sub>-free future, these could also be based on sector coupling and various energy storage technologies, either in form of electricity (batteries), hydrogen, ammonia or other chemical forms, as well as number of currently studied technologies, such as gravity batteries [10]. There has been recently also studies looking into either coupled processes with batteries [11] or enabling the process itself being used as energy storage [12][13].

In order to enable and realize these opportunities, there needs to be – apart from novel designs and respective operational updates – also changes to the current power systems market design, which is still not built for enabling sufficient flexibility but rather relies on dispatchable energy generation.

## DATA EXCHANGE

To enable smooth cooperation between various entities in the power system, including system operators, market participants and owners of various assets, standardized data exchange processes are needed. This includes not only standardized data formats to reduce the overhead of data post-processing, but also systematic platforms for data exchange with sufficient cyber-security and ensured data privacy. Information exchange between various entities can be split into three categories based on the time domain:

1. Static data exchange covering general information such as grid topology and installed assets with their characteristics and capabilities.
2. Scheduled data exchange covering information on scheduled functioning of different elements such as day-ahead market bids, generation and load forecasts, planned outages, and information about expected flexibility.
3. Real-time data exchange covering all real-time information such as telemetry measurements, control setpoints, open/close states of switching equipment, node voltages, and any calculated values (line flows, reserves, etc.).

Depending on the nature of data and the role of different entities, data exchange can be either unilateral or bilateral. It is complex to define a unique list of data requirements that are suitable for all entities and all possible cooperation architectures. Still, it is commonly

agreed upon that in addition to the data values themselves, each dataset should be supplemented with metadata including data source, data user, purpose of the data exchange, principles to be applied (e.g. privacy), and KPIs (e.g. data accuracy and latencies) to verify compliance with requirements of different entities.

International Electrotechnical Commission's (IEC) has introduced the Common Information Model (CIM) standards to facilitate data exchange between different entities [14] as shown in **Figure 4**. Despite the great improvement in interoperability that CIM standards provide, they still need to continuously evolve to meet the changing requirements due to the emerging power system landscape. For instance, practical implementation can be challenging since CIM standards do not cover many aspects such as data portability, aggregation and anonymization, data exchange logs and authentication information. How to create working interfaces to other industrial standards and protocols (e.g. OPC UA, ANSI/ISA-95 and S88) is still undefined, as most standards have been created in an isolated business-specific manner. This, however, needs to be done to enable stronger and more straightforward collaboration opportunities and relationships between the power grids and industrial processes.

## CONCLUSIONS

In this paper we have discussed future challenges arising through the energy transition towards zero emission society. This transition has a great impact on the generation and consumption side but also on the often-ignored power grids and -systems. These must be properly expanded w.r.t. capacity and capability to enable a flexible electricity supply and demand. Energy storage and sector coupling play here an important role. The transition will not happen in an isolated way but calls for stronger collaboration as the resulting system will be a co-play between process- and power systems. One of the first steps is to have proper data exchange where the common information models play a central role. Other concrete design-related challenges are:

- Electrification of processes and the related major changes needed in process design.
- Making processes more resilient against short/mid-term electricity scarcity.
- Related to this, evaluating the cost and benefits of microgrid-enabling design alternatives.
- Enabling flexibility by design in order to allow stronger operational collaboration between process and energy markets.
- Optimal designs to support and benefit from sector coupling, using alternative carriers as CO<sub>2</sub>-

free raw-materials, energy storage for trading and to securing own operations.

The next decades will be a golden era for researchers who are interested in focusing on both process and energy aspects – in design, operations and control domains.

## REFERENCES

1. Knezovic K., Marinakis A., Evrenosoglu Y., Oudalov A. Role of grid and bulk storage in the integration of variable renewable energy resources: Framework for optimal operation-driven multi-period infrastructure planning. *Energy* 226 (2021)
2. Brown T., Schlachtberger D., Kies A., Schramm S., Greiner M. Synergies of sector coupling and transmission reinforcement in a cost-optimised, highly renewable European energy system. *Energy* 160:720-739 (2018)
3. Connolly, D., Lund, H., Mathiesen, B. V., & Leahy, M. A review of computer tools for analysing the integration of renewable energy into various energy systems. *Applied Energy*, 87(4): 1059-1082 (2010)
4. Aneke M., Wang, M.. Energy storage technologies and real life applications – A state of the art review. *Applied Energy*, 179: 350-377 (2016)
5. Denholm, P., Hand, M. Grid flexibility and storage required to achieve very high penetration of variable renewable electricity. *Energy Policy*, 39(3): 1817-1830 (2011)
6. Elsayed A.T., Mohamed A.A., Mohammed O.A. DC microgrids and distribution systems: An overview. *Electric Power Systems Research*, 119, 407-417 (2015)
7. Paulus, M., & Borggrefe, F. The potential of demand-side management in energy-intensive industries for electricity markets in Germany. *Applied Energy*, 88(2), 432-441 (2011)
8. Wachs L., McMillan C., Reese S.B. Reviewing flexibility in industrial electrification: Focusing on green ammonia and steel in the United States. *Energy Research and Social Science*, 103, art. no. 103202 (2023)
9. Hubble A.H., Ustun T.S. Composition, placement, and economics of rural microgrids for ensuring sustainable development. *Sustainable Energy, Grids and Networks*, 13, 1-18 (2018)
10. Sahoo S. and Timmann P. Energy Storage Technologies for Modern Power Systems: A Detailed Analysis of Functionalities, Potentials, and Impacts. *IEEE Access*, 11, 49689-49729 (2023)
11. Zhang X., Hug G., Kolter J.Z., Harjunoski I. Demand response of ancillary service from industrial loads coordinated with energy storage. *IEEE Transactions on Power Systems*, 33 (1), 951-961 (2018)
12. Harjunoski I. Process as Energy Storage. *IFAC-PapersOnLine*, 52 (1), 952-957 (2019)
13. Semrau, R., & Engell, S. Process as a battery: Electricity price aware optimal operation of zeolite crystallization in a continuous oscillatory baffled reactor. *Computers and Chemical Engineering*, 171 (2023)
14. ENTSOE-E. Common Information Model. Available at: <https://www.entsoe.eu/digital/common-information-model>

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Towards Energy and Material Transition Integration – A Systematic Multi-scale Modeling and Optimization Framework

Rahul Kakodkar<sup>a,b</sup>, Betsie Montano Flores<sup>a,b</sup>, Marco De Sousa<sup>a,b</sup>, Yilun Lin<sup>a,b</sup>, and Efstratios N. Pistikopoulos<sup>a,b\*</sup>

<sup>a</sup> Texas A&M Energy Institute, Texas AM University, College Station, TX, USA

<sup>b</sup> Artie McFerrin Department of Chemical Engineering, Texas AM University, College Station, TX, USA

\* stratos@tamu.edu

## ABSTRACT

The energy transition is driven both by the motivation to decarbonize as well as the decrease in cost of low carbon technology. Net-carbon neutrality over the lifetime of technology use can neither be quantitatively assessed nor realized without accounting for the flows of carbon comprehensively from cradle to grave. Sources of emission are disparate with contributions from resource procurement, process establishment and function, and material refining. The synergies between the constituent value chains are especially apparent in the mobility transition which involves (i) power generation, storage and dispatch, (ii) synthesis of polymeric materials, (iii) manufacturing of vehicles and establishment of infrastructure. Decision-making frameworks that can coordinate these aspects and provide cooperative sustainable solutions are needed. To this end, we present a multiscale modeling and optimization framework for the simultaneous resolution of the material and energy value chains. A case study focusing on the transition of mobility technology towards electric vehicles in Texas is presented. The key contributions of the proposed framework are (i) integrated network design and operational scheduling, (ii) the tracking of disparate emissions, (iii) simultaneous modeling of the material and energy supply chains, (iv) implementation on `energy`, a python package for the multiscale modeling and optimization of energy systems.

**Keywords:** energy transition, material transition, mixed integer programming, multiscale modeling, carbon accounting

## INTRODUCTION

The ongoing energy transition towards net-carbon neutrality has various challenges, not limited to (i) managing the storage for intermittent renewables, (ii) accounting for emissions from disparate sources, (iii) managing the cost to consumers, (iv) coordinating the transition of technologies for different applications such as the generation of power, transportation and production of dense energy carriers (DECs) (v) meeting the material requirements for establishing infrastructure sustainably [1]. Addressing these challenges requires a holistic view of multiple interdependent supply chains and the synergies between them.

Moreover, decarbonization requires both direct and

indirect emissions to be accounted for. Direct emissions occur from the release of pollutants at the operational level by processes, whereas indirect emissions occur due to the refining and procurement of materials and resources as well as the establishment of processes and infrastructure. Further, indirect emissions occur at the onset of establishing infrastructure, while direct emissions can occur over a protracted temporal horizon throughout the lifetime of technology. Decision-makers planning energy systems of the future need to coordinate both aspects to determine optimal transition pathways.

Materials such as minerals, metals, and polymers are required to establish infrastructure. The refining and processing of these is a contributor to emissions. However, polymers can also act as a carbon sink and, in the

context of the circular economy, provide an opportunity for the valorization of the carbon vector.

Given our contemporaneous dependence on carbon-based feedstock for both power generation and production of polymers, the oil & gas (O&G) industry plays a central role in enabling both the energy and material transition. The repurposing of the O&G industry towards the manufacturing of polymeric materials through captured carbon dioxide will also preserve established production infrastructure and capital. Furthermore, polymers offer a durable alternative to metallic materials which have a high energy demand for refining, produce toxic refuse during mining, and have challenges in terms of recycling and reuse in some cases. For example, passenger vehicle designs have evolved to include a larger amount of polymers in their interior while achieving improvements in efficiency and safety due to lower density and higher impact resistance respectively [2].

The use of hydrocarbon feedstock to produce polymers for electric vehicles (EVs) as opposed to transportation fuels for internal combustion engine (ICE) vehicles can also reduce direct carbon emissions as they can act as a sink for carbon that would be otherwise emitted as greenhouse gases (GHGs) [3]. Readily available polymers such as polypropylene (PP), polyamide (PA), polystyrene (PS), polycarbonate (PC), polyurethane (PUR), polyvinyl chloride (PVC), polyethylene (PE), and polyethylene terephthalate (PET) all find use in automobile production. Moreover, polymers are also needed for insulation in grid infrastructure, power conversion systems, and also for ancillary infrastructure in solar and wind farms.

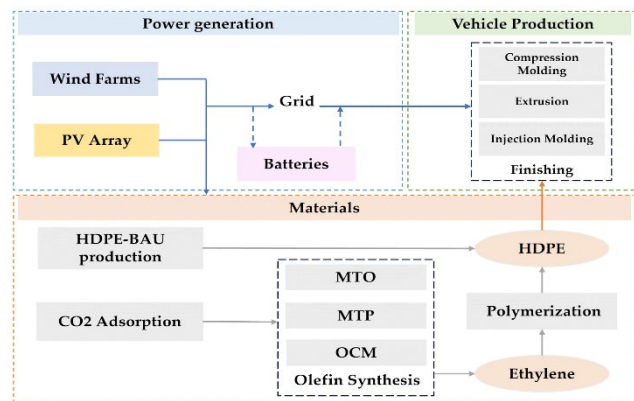
While the modeling and optimization of different aspects of the energy transition has been an active field of research [4,5,6,7], the parallel transition of materials has not received equal attention. In this publication, we introduce a framework for the simultaneous modeling of the material and energy transition, allowing the identification of solutions that reduce emissions across disparate emission scopes. The system is represented through the resource task network (RTN) methodology augmented for the consideration of different material options for process infrastructure. The emissions arising from disparate sources such as material synthesis, resource consumption and utilization post discharge, and production are considered. The mathematical programming framework is modeled as a mixed integer linear program (MILP) with binaries assigned for the establishment of processes and the choice of material modes for said processes. Moreover, the model does simultaneous network design and scheduling which is able to account for renewable intermittency and optimize energy storage.

The publication is organized as follows: first the details of the considered mobility transition supply chain are discussed, next the modeling and solution methodology are elucidated upon and certain aspects of the general

formulation are discussed, followed by the results and ongoing work.

## A MOBILITY TRANSITION SUPPLY CHAIN

The mobility transition with respect to passenger vehicles entails the adoption of EVs to replace ICE vehicles. Using a simple linear forecast to predict EV sales for Texas, more than 900,000 units will be sold in 2050 considering an annual sales growth rate of 16% [8], and about 12 million EVs will be on the road. EVs differ from ICE vehicles in both the energy source for their function, as well as the supply chains involved in making energy available. Fossil fuels utilized in ICE vehicles cause emissions at the point of use. Conversely, EVs have no direct emission but can cause emissions as a result of power generation through fossil fuel-based power generation. Moreover, while power can be generated through renewable means, materials required to establish power generation and distribution infrastructure are still privy to emissions. The transition to EVs will need to be supported by an expansion of the grid and power generation infrastructure.



**Figure 1.** A schematic of the mobility transition supply chain.

The transition of mobility infrastructure requires the transition of technologies for 1) transportation 2) manufacturing, 3) power generation, and 4) polymer production. As a whole, three distinct blocks can be considered as shown in Figure 1, viz. 1) Power generation which include solar photovoltaics (PVs) and wind farms (WFs), 2) Polymer production wherein High Density Polyethylene (HDPE), 3) the production of vehicle parts through a combination of different molding methods. As a whole, we look at two different technology pathways:

1. The business as usual (*BAU*) pathway where in HDPE is produced from fossil feedstock such as natural gas and oil
2. The carbon capture utilization and sequestration (*CCUS*) pathway which necessitates the direct air

capture (DAC) of CO<sub>2</sub> followed by one of three processes for the production of olefins, with polymerization of produced ethylene as the final step

The three pathways for olefin production considered are:

- **Methanol to Olefins (MTO):** In this route, methanol is synthesized from CO<sub>2</sub> hydrogenation and converted into olefins (ethylene and propylene).
- **Methanol to Propylene (MTP):** In this process, the major product is propylene, which is synthesized from CO<sub>2</sub>.
- **Oxidative Coupling of Methane (OCM):** This process produces mainly ethylene through oxidative coupling of methane derived from the hydrogenation of CO<sub>2</sub>

The MTO and MTP processes represent successful initiatives aimed at reducing reliance on petroleum for olefin production. In the MTO process, methanol is transformed into olefins with a carbon selectivity ranging from 78% to 82% at approximately 500 °C and 250 kPa, utilizing a SAPO-34-type zeolite catalyst. The by-products include propylene, C4 fraction, and liquified petroleum gas (LPG).

Similarly, in the MTP process, methanol is converted into propylene with a carbon selectivity of around 71% at temperatures between 400 and 500 °C and a pressure of 150 kPa, employing a zeolite-based catalyst [9]. LPG, propylene, and C5 fraction are by-products.

In the OCM process [10], methane (CH<sub>4</sub>) and oxygen (O<sub>2</sub>) undergo an exothermic reaction on a catalyst bed, resulting in the formation of ethylene (C<sub>2</sub>H<sub>4</sub>), ethane (C<sub>2</sub>H<sub>6</sub>), water (H<sub>2</sub>O), and heat. Traditional reactors, such as packed bed reactors (PBR) and fluidized bed reactors, have been commonly employed for this purpose. However, there is a current exploration of Fluidized Bed Membrane Reactor (FBMR) options to assess the potential improvement in selectivity. The operating conditions for this technology involve temperatures of approximately 800 °C and pressures of 200 kPa. Further, all three processes require the provision of steam, cooling water, oxygen, and hydrogen besides power.

Produced ethylene is converted into HDPE via polymerization through different technologies such as CSTRs, loop slurry, fluidized-bed gas phase, and stirred-bed gas phase reactors. Various processes have been marketed by companies such as LyondellBasell, Mitsui Chemicals, Chevron Phillips Chemicals, Univation, Lummus Novolen, etc. Nevertheless, as data for such processes is not publicly available, the average data from Europe for such processes is considered [11]. Both HDPE production routes, BAU or CCUS, are subject to the finishing processes; a combination of compression molding

(CM) 24.25%, extrusion (EX) 9.25% and injection molding (IM) 66.5% [8].

Currently, the energy demand for EVs and production of resources is met through a mix of fossil fuel (coal and natural gas), and renewable (wind and solar) power generation. In the transition scenario, renewable power will replace fossil fuel-based power generation. Moreover, requisite power generation and management systems can be established through different material alternatives as well. PVs, for example, can be made from monocrystalline or polycrystalline silicon, lithium for lithium-ion batteries can be sourced either through rock or brine lithium, and wind farms can be offshore or land-based each with different material requirements. The other materials required for construction include glass, steel, concrete, aluminum, silicon, copper, and cast iron.

Polymers in EVs account for about 13% of the mass of the vehicle. In the illustrative example, we only consider the need for HDPE which accounts for about 2.7 kg (1.4%) mass of the total amount of polymers used in cars [12], the total car mass considered is 1481 kg. The objective of the considered study is to analyze the network designs of the current HDPE demand for the ~1.3 million passenger vehicles produced annually in Texas through renewable means. Of key interest are 1) the identification of optimal cost and emission technology pathways, 2) the quantitative determination of the trade-offs between cost and emission, 3) identification of optimal material modes for the establishment of infrastructure.

The region of Houston is used as a proxy for the entire state of Texas for the collection of wind and solar data. The GREET model is used to source parameters for natural gas, oil, electricity use, and related process emissions [12], while the Ecoinvent version 3.9.1 database is used to assign a global warming potential (GWP) for every material considered in the study [13].

## METHODOLOGY AND FORMULATION

The framework is multiscale in that it models spatiotemporally disparate phenomena such as renewable intermittency and augmentations in technology costs and efficiency. Moreover, the operational and network planning decisions are modeled simultaneously as well. Emissions are considered from: 1) resource consumption, 2) resource discharge, 3) material procurement for establishing infrastructure. The framework can also consider emissions from construction or manufacturing processes, but this is not modeled in the presented case study given the lack of reliable data. Note, that the framework can also account for different environmental indicators, namely, global warming, ozone depletion, acidification, and eutrophication (marine, terrestrial, freshwater). However, the presented case study only accounts for the global warming aspect of emissions.



For energy technologies, power generation processes are modeled alongside energy storage, and different material modes are considered. The lithium-ion battery (LiI), for example, can use lithium sourced from brine or rock. Each option has different associated emissions due to the process of mining these materials differ significantly. Similarly, different combinations of materials such as steel, cast iron, silicon, concrete, etc. are considered for establishing PVs and WFs.

The system is represented and modeled through an augmentation of the resource task network (RTN) methodology, which considers the materials required to set up processes [14]. This representation is christened the resource task material network (RTMN). Resources, by definition, can be consumed, discharged, stored, transported, or produced. On the other hand, materials are utilized solely for the establishment of infrastructure such as transport linkages and processes. The need for this strict distinction between materials and resources stems from: (i) unlike resources, materials are not converted, (ii) disparate temporal nature of use, wherein resources are used continuously over the scheduling scale and materials are used only in time periods over the network scale when technologies are established, (iii) convenience of reporting emissions in terms of scopes. As a whole, this streamlines the system representation as well as the mathematical modeling. As an example, in the mobility transition supply chain, HDPE is a *resource* produced through the culmination of a set of processes, whereas the concrete required to establish a wind farm is a *material*.

The planning horizon is considered at two distinct discretizations: 1) a scale for network level decisions, 2) a scheduling scale with 8760 discretizations to capture the intermittency of renewables and the flow of resources. Given that a single period is considered for network level decisions, the capital cost or emissions related to material utilization are not annualized. However, the model can be easily expanded to consider multiple network periods along with associated changes in technology costs. Further, while the example presented models a single location, multiple locations along with associated transportation modes between these locations can also be considered.

The framework simultaneously optimizes network design and process scheduling. Mixed integer programming is used with binaries for decisions such as locating processes and determining optimal material modes. Continuous variables capture the mass balance, and monetary aspects. Some important constraints such as the material balance, emission constraints, and network design are discussed here. Network design is achieved through constraints (1 and 2). Material constraints (3-5) allow the model to choose between different available material modes. Each material mode in turn has an

associated consumption of materials ( $\phi$ ) per unit capacity. Notably, emissions can also occur through 1) the consumption or use of produced resources, and 2) direct process emissions as shown in emission constraints (6-9). Inventory balance in every time period in the scheduling horizon is done through constraint 10, wherein each variable is bounded by either a parameter ( $S^{max}$  for S,  $C^{max}$  for C), or by a variable in the network scale ( $Cap^S$  for  $Inv$ ,  $Cap^I$  for  $P$ ).

sets	definition
$l \in L$	location
$i \in I$	process
$r \in R$	resource
$m \in M$	material
$w \in W$	material modes
$e \in E$	emissions
$t \in T^N$ or $t \in T^S$	network or scheduling scale

variables	definition
$Cap_{l,i/r,t}^{I/S}$	process/inventory capacity
$Em_{l,e,r/i/m,t}^{r/i/m}$	resource/process/material emission
$Em_{l,e,t}^{(L)}$	Emission (total at location)
$C_{l,r,t}^{(L)}$	Resource consumption (total at location)
$S_{l,r,t}^{(L)}$	Resource discharge (total at location)
$Mat_{l,i,mt}^I$	Material utilized for process
$Mat_{l,m,t}^L$	Material utilized at location
$Cap_{l,i,m,t}^{I-W}$	Capacity of material mode
$X_{l,i,m,t}^I$	Binary for locating process
$X_{l,i,w,t}^{I-W}$	Binary for choosing process mode
$X_{l,r,t}^S$	Binary for storage facility
$Inv_{l,r,t}$	Inventory level for resource
$P_{l,i,t}$	Production level for process

parameters	definition
$Cap_{l,r/i,t}^{S/I-min/max}$	max/min inventory/process

	capacity
$EP_e^{r/m}$	resource/material emission
$\phi(m, w)$	material consumption
$\eta(r, i)$	conversion

$$Cap_{l,r,t}^{s-min} \cdot X_{l,r,t}^s \leq Cap_{l,r,t}^s \leq Cap_{l,r,t}^{s-max} \cdot X_{l,r,t}^s \quad (1)$$

$$Cap_{l,i,t}^{l-min} \cdot X_{l,i,t}^l \leq Cap_{l,i,t}^l \leq Cap_{l,i,t}^{l-max} \cdot X_{l,i,t}^l \quad (2)$$

$$Mat_{l,i,m,t}^l = \sum_{\forall w \in W} (\phi(m, w) \cdot Cap_{l,i,w,t}^{l-w}) \quad (3)$$

$$Cap_{l,i,t}^{l-min} \cdot X_{l,i,w,t}^{l-w} \leq Cap_{l,i,w,t}^{l-w} \leq Cap_{l,i,t}^{l-max} \cdot X_{l,i,w,t}^{l-w} \quad (4)$$

$$X_{l,i,t}^l = \sum_{\forall w \in W} X_{l,i,w,t}^{l-w} \quad (5)$$

$$Em_{l,e,r,t}^r = EP_e^{r-cons} \cdot C_{l,r,t}^l + EP_e^{r-use} \cdot S_{l,r,t}^l \quad (6)$$

$$Em_{l,e,m,t}^m = EP_e^m \cdot Mat_{l,m,t}^l \quad (7)$$

$$Em_{l,e,i,t}^i = \sum_{\forall i \in I} EP_e^r \cdot \eta(r, i) P_{l,i,t}^l \quad (8)$$

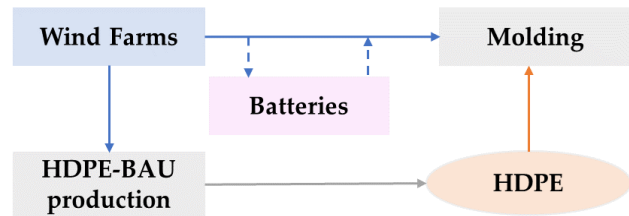
$$Em_{l,e,t}^L = \sum_{\forall r \in R} Em_{l,e,r,t}^r + \sum_{\forall m \in M} Em_{l,e,m,t}^m + \sum_{\forall i \in I} Em_{l,e,i,t}^i \quad (9)$$

$$\sum_{\forall i \in I} \eta(r, i) P_{l,i,t}^l + C_{l,r,t} + Inv_{l,r,t-1} = Inv_{l,r,t} + S_{l,r,t} \quad (10)$$

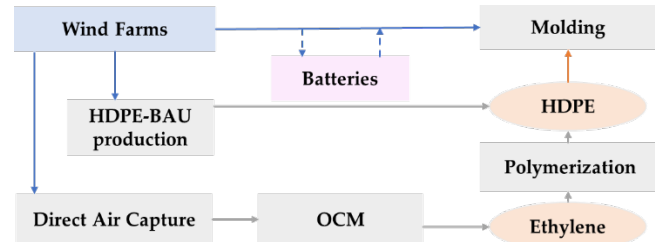
The implementation in *energiapy* [15] utilizes a component based architecture with temporal horizon, resources, materials, processes, locations, and scenarios serving as modeling components. Various material modes can be declared for each process, wherein the material consumption for the establishment of said infrastructure on a per unit basis is provided as attributes. The associated GWP is declared as an attribute of each material and resource in the object-oriented programming framework. A small example is available in the package documentation to guide users on the application.

## RESULTS AND DISCUSSION

The framework can be optimized to various objectives, the two objectives discussed here are to 1) minimize system costs, and 2) minimize system emissions. First a base case is established wherein the system is optimized only to minimize the system cost while not restricting emissions. The network design for the base case as shown in Figure 2 consists of WF, Lil, HDPE production from the BAU pathway and the molding processes (extrusion, compression, injection). Moreover, the land-based material mode is chosen for establishing WF, and brine lithium is chosen for the Lil. The emissions resulting from the 2.7 kg requirement of HDPE per car is found to be 11.08 kg CO<sub>2</sub>-eq. This includes 3.15 and 0.8 kg CO<sub>2</sub>-eq from the purchase of natural gas and oil respectively, 5.12 from the venting of CO<sub>2</sub>, and 2.00 from the material requirements for establishing the land-based WF and brine lithium based Lil.



**Figure 2.** Low cost pathway for HDPE production.

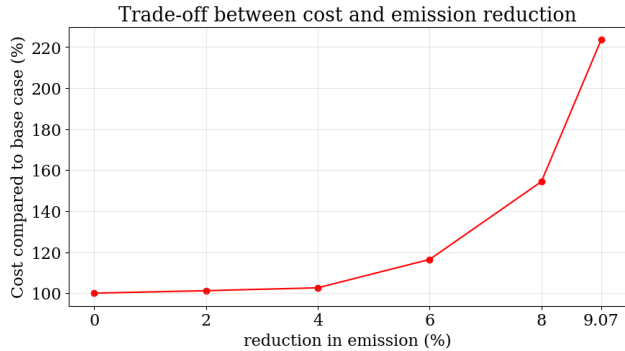


**Figure 3.** Low carbon pathway for HDPE production.

With the base case established, the network is then optimized towards the reduction of GWP. The maximum possible reduction in emissions is found to be ~9.07%. The cost optimal design for a ~9.07% reduction in emission tilts in favor of the CCUS pathway. The reduction entails a 2.2 factor increase in cost compared to the base case. Even in the minimum emission case, the BAU pathway for the production of HDPE is still utilized. This is due to the energy intense nature of the CCUS pathway which requires larger power generation and energy storage capacities, which are subject to both material emissions and cost. Further, the OCM process is considered. Note that the MTP and MTO processes both have LPG as a by-product which results in emissions at the point of utilization. The WF and Lil capacities are 1.86 and 3.74 times larger than the base case scenario.

The system can then be analyzed for the successive reduction of emissions up to 9.07%. Figure 4 shows

a Pareto front which compares the reduction of emissions to the cost as compared to the base case. Notably, the first 6% reduction in emissions comes at a marginal increase (16.44%) in cost. The network design also evolves with the reduction in emissions, this includes the choice of technologies, the choice of materials to establish said processes, and the process capacities.



**Figure 4.** Trade-off between emission cost and emissions reduction.

The initial 4% reduction in emissions is achieved largely by setting up the energy intensive OCM process (refer to Table 1), and managing the capacity of the WF with a larger Lil. Beyond 4%, larger power systems are needed to accommodate the higher power demand. Given the higher cost of power infrastructure, system costs increase considerably.

Further, the consistent choice of land-based wind farms in both the minimum cost and minimum emissions represents a *win-win* scenario that offers the best solution across multiple objectives. Similarly, some options are never chosen such as the olefin production through the MTP and MTO route which are recognized as being both cost and emission intensive. As a reminder, processes can have different efficiencies based on the choice of materials, as well as different emission potentials based on the sourcing of materials (such as rock or brine lithium). Also, the cost of technologies are expected to reduce along with augmentations in efficiency resulting from adoption and research as the transition progresses.

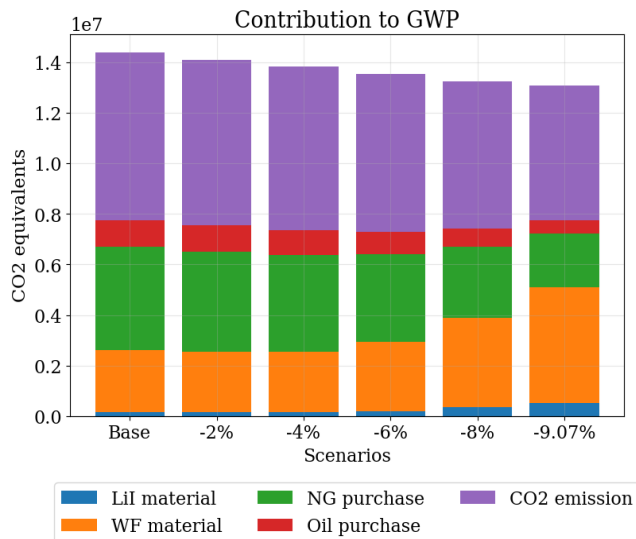
It is also observed that a production capacity of HDPE from the BAU pathway is established even in the low carbon scenarios. This pathway is utilized on days with low wind availability given the low power requirement. If this pathway is not made available, the required power generation and energy storage capacity is significantly larger. In fact, to achieve a 6% reduction in emissions without the BAU pathway, the WF is sized at 28.98 MW and Lil is sized at 17.7 MW, representing an increase of 315% and 514.5% respectively. Meanwhile the cost of the system increases 388%. This highlights the trade-offs between direct and indirect emissions. The CCUS pathway utilizes significantly higher power which results in larger power generation capacities which in turn cause an increase in the emissions resulting from material utilization.

Figure 5 shows the disparate contribution to emissions. The largest source of emissions is direct CO<sub>2</sub> discharge. The lower emission scenarios have a larger contribution from WF and Lil as they are sized at a larger capacity. Natural gas (NG) and oil consumption decreases with the lower reliance on the BAU pathway and hence the emissions resulting from their sourcing and purchase. It can also be noted that, in the given model, none of the sources of emissions can be entirely eliminated.

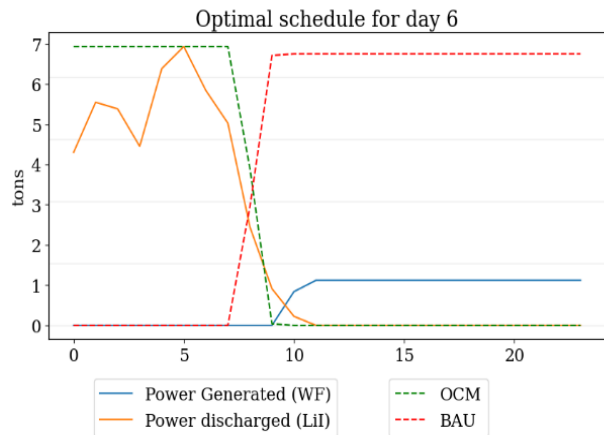
The framework also provides optimal schedules. For example in Figure 6, the schedule for power generation through WF is plotted alongside the power discharged from Lil, and the production through OCM is plotted alongside the production through the BAU pathway for the scenario with a GWP reduction of 6%. It can be seen that the constant demand for HDPE is met either through the OCM or BAU pathway. However, when the wind potential is high the OCM pathway is preferred. During periods of low wind availability, the BAU pathway is preferred with the power demand being met from the discharge of energy stored in the Lil. This is also reflected in the sizing of processes, wherein the BAU process has the same capacity as the OCM process (refer to Table 1).

**Table 1.** Process capacities for different scenarios.

Scenario	DAC (tons)	HDPE-BAU (tons)	HDPE-CCUS (tons)	OCM (tons)	Lil (MW)	WF (MW)
Base	0	0.44	0	0	2.51	8.29
-2%	0.75	0.44	0.1	0.1	2.76	8.03
-4%	2.10	0.44	0.28	0.29	2.75	8.03
-6%	3.23	0.44	0.44	0.45	3.45	9.21
-8%	3.23	0.44	0.44	0.45	6.14	11.98
-9.07%	3.23	0.35	0.44	0.44	9.41	15.45



**Figure 5.** Contribution to GWP from source for each scenario.



**Figure 6.** Schedule for power generation and discharge from energy storage.

Caution should be exercised when interpreting the results due to two limitations: 1) the unavailability of material use data for some processes causing a possible bias in results where values are assumed, and 2) the GWP associated with materials can vary by literature source. Nevertheless, the framework can be interrogated to quantify the sensitivity of the solution to parameter values.

The framework preserves the simultaneous scheduling and network design capability of earlier frameworks in literature [4,5,6,7] while allowing: 1) the estimation of emissions from disparate sources, and 2) identifying optimal material choices for the establishment of processes. Moreover, the framework is applied to a novel case study wherein the material and energy supply chains are modeled simultaneously. This provides a more holistic view of the system, thus enabling decarbonization across the different scopes of emissions. In principle, this framework can be applied towards the optimal

design of systems bearing awareness to life cycle considerations.

## FUTURE WORK

While only GWP is considered in the presented work, other environmental effects such as ozone depletion and acidification potential can also be modeled. The optimization of multiple criteria using multi-objective optimization (MOO) can be reported. In its current form, the emission accounting methodology cannot be considered a comprehensive life cycle assessment (LCA) [7] given the lack of focus on process lifetimes and circularity of materials [16]. A challenge that needs to be addressed in the RTMN methodology is the fact that resources and materials are often not distinct. For example, while HDPE is treated as a resource in the considered mobility example, it may also serve as a material for the establishment of grid infrastructure.

Furthermore, the sensitivity of the model solution to considered parameters can also be assessed. The framework can also be run for longer temporal horizons which will allow the consideration of reductions in technology cost over the temporal horizon. Moreover, only HDPE is analyzed in the presented work, future iterations of the framework will also assess the role of other polymers utilized in the manufacturing of vehicles. Power is also required for charging EVs which can be modeled in tandem. Besides renewable intermittency, the framework can also accommodate variability in the cost of resources such as NG and oil and resource demand. Case studies considering the aforementioned aspects will be presented in future publications.

## ACKNOWLEDGEMENTS

We acknowledge the financial support from Saudi Aramco, Shell Plc., the Texas A&M Energy Institute, and Artie McFerrin Department of Chemical Engineering at Texas A&M University.

## REFERENCES

1. R. Kakodkar, G. He, C. Demirhan, M. Arbabzadeh, S. Baratsas, S. Avraamidou, D. Mallapragada, I. Miller, R. Allen, E. Gençer, et al. (2022). A review of analytical and optimization methodologies for transitions in multi-scale energy systems. *Renew. Sust. Energ. Rev.* 160:112277
2. Cabernard, L., Pfister, S., Oberschelp, C., & Hellweg, S. (2022). Growing environmental footprint of plastics driven by coal combustion. *Nature Sustainability*, 5(2), 139-148.
3. Danilecki, K., Mroziak, M., & Smurawski, P. (2017). Changes in the environmental profile of a popular

passenger car over the last 30 years—Results of a simplified LCA study. *Journal of Cleaner Production*, 141, 208-218.

4. Demirhan, C. D., Tso, W. W., Powell, J. B., Heuberger, C. F., & Pistikopoulos, E. N. (2020). A multiscale energy systems engineering approach for renewable power generation and storage optimization. *Industrial & Engineering Chemistry Research*, 59 (16), 7706-7721.
5. Zhang, Q., Martín, M., & Grossmann, I. E. (2019). Integrated design and operation of renewables-based fuels and power production networks. *Computers & Chemical Engineering*, 122, 80-92.
6. Allman, A., & Daoutidis, P. (2017). Optimal design of synergistic distributed renewable fuel and power systems. *Renewable Energy*, 100, 78-89.
7. Hugo, A., & Pistikopoulos, E. N. (2005). Environmentally conscious long-range planning and design of supply chain networks. *Journal of Cleaner Production*, 13(15), 1471-1491.
8. What cars will we be driving in 2050?. Fuel Freedom Foundation. (n.d.). <https://www.fuelfreedom.org/cars-in-2050/>
9. Z. Zhao, J. Jiang, F. Wang, An economic analysis of twenty light olefin production pathways, *Journal of Energy Chemistry*. 56
10. M. Ali, Y. Tian, D. Kenefake, E.N. Pistikopoulos, Process Design and Intensification of Circulating Catalytic Fluidized Bed Membrane Reactor for Oxidative Coupling of Methane, in: 2023: 2031-2036. <https://doi.org/10.1016/B978-0-443-15274-0.50323-1>.
11. Polymers - ineris. (n.d.). [https://aida.ineris.fr/sites/aida/files/documents-bref/pol\\_bref\\_1006.pdf](https://aida.ineris.fr/sites/aida/files/documents-bref/pol_bref_1006.pdf)
12. GREET 2022 Net software v1.3.0.13991, 2022. Greet Life Cycle Model 2022 Argonne. Argonne National Laboratory. Available online: 10.11578/GREET-Net-2022/dc.20220908.2.
13. Ecoinvent Centre. (2023). Ecoinvent Database. Version 3.9.1. Ecoinvent Centre, St. Gallen.
14. Pantelides, C. C. (1994, July). Unified frameworks for optimal process planning and scheduling. In *Proceedings on the second conference on foundations of computer aided operations* (pp. 253-274). Cache Publications New York.
15. Kakodkar, R., & Pistikopoulos, E. (2023, November). Energiapy-an Open Source Python Package for Multiscale Modeling & Optimization of Energy Systems. In 2023 AIChE Annual Meeting. AIChE.
16. Styliani Avraamidou, Stefanos G. Baratsas, Yuhe Tian, and Efstratios N. Pistikopoulos. (2020). Circular Economy - A Challenge and an Opportunity for Process Systems Engineering. *Comput. Chem. Eng.* 133:106629

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Process and Network Design for Sustainable Hydrogen Economy

Monzure-Khoda Kazi <sup>a,b</sup>, Akhilesh Gandhi <sup>a</sup>, and M.M. Faruque Hasan<sup>a,b,\*</sup>

<sup>a</sup> Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX 77843-3122, USA

<sup>b</sup> Texas A&M Energy Institute, Texas A&M University, College Station, TX 77843-3372, USA

\* Corresponding Author: hasan@tamu.edu.

---

## ABSTRACT

This study presents a comprehensive approach to optimizing hydrogen supply chain network (HSCN), focusing initially on Texas, with potential scalability to national and global regions. Utilizing mixed-integer nonlinear programming (MINLP), the research decomposes into two distinct modeling stages: broad supply chain modeling and detailed hub-specific analysis. The first stage identifies optimal hydrogen hub locations, considering county-level hydrogen demand, renewable energy availability, and grid capacity. It determines the number and placement of hubs, county participation within these hubs, and the optimal sites for hydrogen production plants. The second stage delves into each selected hub, analyzing energy mixes under variable solar, wind, and grid profiles, sizing specific production and storage facilities, and scheduling to match energy availability. Iterative refinement incorporates detailed insights back into the broader model, updating costs and configurations to converge upon an optimal supply chain design. This design encapsulates macro-level network configurations, including centralization versus decentralization strategies, transportation cost analysis, and carbon footprint assessment, as well as micro-level operational specifics like renewable energy contributions, facility scale, and energy portfolio management. The methodology's robustness allows for strategic insights into hydrogen production facility siting, aligning with local energy resources and supply chain economics. This adaptable, multi-scale approach contributes to informed decision-making in the evolution of sustainable hydrogen-based energy systems, offering a roadmap for policy reforms and strategic supply chain development in diverse energy landscapes.

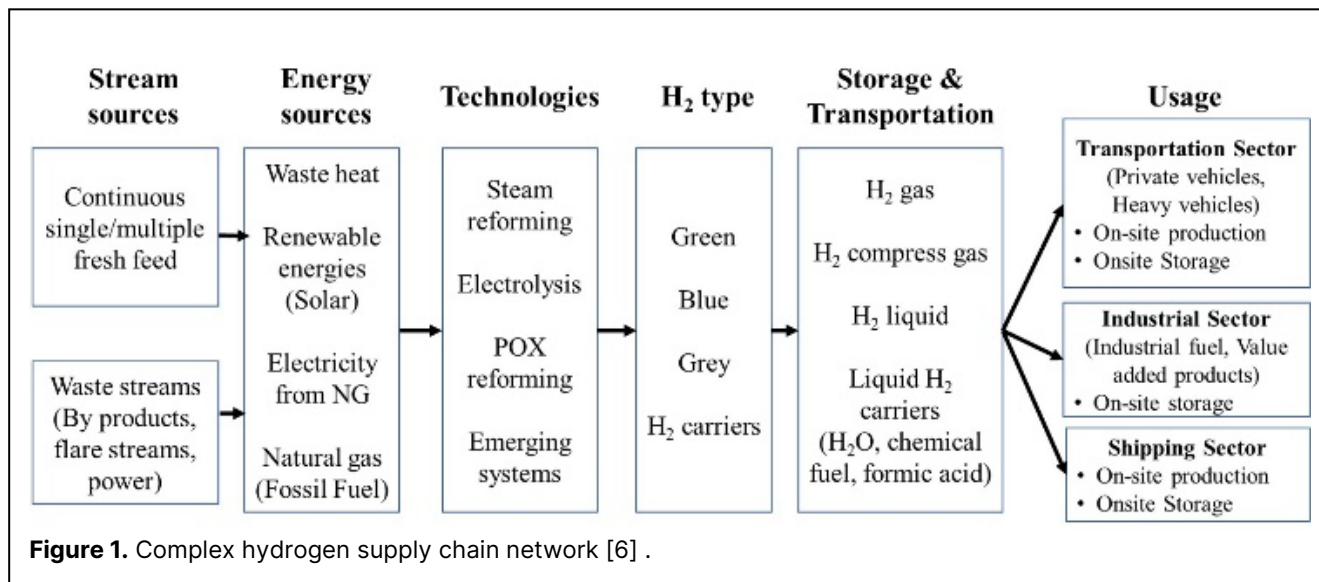
---

**Keywords:** Energy Management, Hydrogen, Optimization, Renewable and Sustainable Energy, Supply Chain, Network Design.

## INTRODUCTION

The global energy landscape is undergoing a paradigm shift towards sustainable and clean energy sources, with hydrogen emerging as a pivotal player in this transition. Hydrogen, particularly green hydrogen produced from renewable energy sources, offers a promising solution to decarbonize various sectors, including transportation, industrial processes, and energy storage [1, 2]. The U.S. Department of Energy's (DOE) investment in hydrogen hubs underscores their pivotal role in advancing the nation's clean energy agenda. With a commitment of \$7 billion towards establishing H2Hubs, alongside \$1 billion to boost clean hydrogen demand and \$1.5 billion to enhance electrolysis technologies, the DOE aims to

significantly reduce the cost of clean hydrogen to \$1 per kilogram within a decade [3]. This initiative is not just an investment in sustainable energy but also a substantial job creator, promising to generate tens of thousands of well-paying jobs across the country. Moreover, the H2Hubs are expected to play a crucial role in environmental conservation by eliminating approximately 25 million metric tons of carbon dioxide emissions annually, equating to the emissions of about 5.5 million gasoline-powered cars [4]. This strategic move marks a significant step towards realizing a more sustainable, low-carbon future, positioning hydrogen hubs as a cornerstone in the transition to cleaner energy sources.



However, establishing an efficient hydrogen supply chain network (HSCN) poses significant challenges due to its complexity and the need for a multi-scale optimization approach encompassing production, storage, transportation, and distribution [5, 6]. The complexity stems from several key aspects: the geographical dispersion of supply and demand centers, the integration of diverse and intermittent renewable energy sources, and the need to align production with fluctuating energy availability. Additionally, the network comprises various interconnected components, including production sites, storage facilities, and distribution hubs, each with its own set of operational constraints and dependencies (see Fig. 1). The variability in renewable energy output, such as solar and wind, adds another layer of complexity, necessitating advanced planning and forecasting methods. This complexity is further amplified by the dynamic nature of market demands, technological advancements, and regulatory landscapes. Solving this multifaceted problem requires not only sophisticated computational models and optimization algorithms but also a deep understanding of the interplay between various elements of the hydrogen supply chain.

Pertinent literature reveals various approaches to modeling and optimizing Hydrogen Supply Chain Networks (HSCNs). Study such as by Vijayakumar et al. [7] have focused on geographic and economic aspects of hub placement. They highlight the importance of long-term planning in mitigating system costs and retail prices, but their deterministic approach overlooks uncertainties in demand and feedstock prices which are crucial for accurate forecasting. In contrast, Li et al. [8] provide an optimization-oriented review of hydrogen supply chain network design, noting gaps such as the treatment of uncertainty. Alkatheri et al. [9] address the intermittency challenges of renewable energies with a multiscale stochastic

programming approach for energy hub design, despite the computational complexity. Moran et al. [10] offer a flexible tool for analyzing regional hydrogen hubs, as exemplified by their Irish case study, but do not fully consider the implications of using grid electricity from non-renewable sources and the variability of the renewable integration. Additionally, Marouani et al. [11] delved into the integration of renewable energy sources into the supply chain. However, the dynamic and variable nature of renewable energy availability, particularly solar and wind, and its impact on hydrogen production and storage scheduling and sizing remains under-addressed.

To address these challenges, this study proposes a novel two-stage mixed-integer nonlinear programming (MINLP) approach. The first stage involves broad supply chain modeling to identify optimal hydrogen hub locations and configurations, considering county-level demand, renewable energy availability, and grid capabilities. The second stage focuses on detailed hub-specific modeling, specifying energy mixes, production and storage capacities, and schedules in alignment with variable energy inputs.

The results of this approach include the identification of optimal hub locations and configurations, tailored energy mixes for each hub, and detailed operational schedules that maximize efficiency and minimize costs. Moreover, the iterative refinement process employed in this study allows for the continuous updating of model parameters, leading to increasingly accurate and optimal solutions. This study not only contributes to the existing body of knowledge on HSCN optimization but also provides a practical and scalable framework for policymakers and industry stakeholders.

## PROBLEM STATEMENT

The central hypothesis of our study posits that by

addressing key research questions within Texas' hydrogen supply chain, the findings could be extrapolated to national or even international scales. McKinsey & Company's sustainability report projects that Texas' demand for clean hydrogen may increase to 21 million tonnes (MT) by 2050, up from the current 3.6 MT produced conventionally [12]. Our assumption is that all 254 counties in Texas will contribute to this demand based on factors such as local energy requirements, population, available land, and the variability of energy sources including the grid, wind, and solar. The target is to answer the following key research questions:

- What are the strategic locations for the hydrogen production plants?
- What constitutes the optimal energy mix for electrolytic hydrogen production, given variable electricity pricing, wind availability, and solar irradiance?
- What are the optimal size of H<sub>2</sub> production facilities, renewable farms, energy storage, considering the temporal variations in renewable energy availability?
- What is the comprehensive cost of the hydrogen supply chain, including production, storage, and transportation?
- What will be the optimal scheduling of the hydrogen production process to match energy availability?

## METHODOLOGICAL APPROACH

We adopt a multi-scale optimization framework that integrates both supply chain optimization (level 1) and process design and energy scheduling (level 2), ensuring convergence towards an optimal supply chain design that encapsulates macro-level network configuration and micro-level operational details.

### Level 1: Supply Chain Optimization

At the macro-level, the network configuration is informed by the optimization of strategic decisions such as site selection, facility sizing, hydrogen distribution, transportation costs to/from other counties, energy portfolio mix and management cost, guided by county-specific roles and requirements. In our supply chain optimization, we employ piecewise linearization to address the economies of scale inherent in hydrogen production. This mathematical technique allows us to model the cost benefits of scaling production facilities accurately. By breaking down the nonlinear cost structure into linear segments, we can analyze scenarios where a single large production facility or multiple smaller ones are more economically viable. This is crucial in evaluating the

feasibility of a hub approach to hydrogen production. Furthermore, this linearization facilitates the use of linear programming techniques, which significantly expedite the optimization process, ensuring a swift and efficient path to finding the optimal supply chain configuration.

$$\min \sum_i \sum_j c_{ij}^{trans,H_2} x_{ij}^{H_2} r_{ij} + \left(\frac{\varphi}{\beta}\right) \sum_i \sum_j c_{ij}^{trans,E} x_{ij}^E r_{ij} + \sum_p c_{i,p}^{prodH_2} \Delta x_{i,p}^{H_2} + \sum_i \sum_{sor} c_{i,sor}^{prodE} x_{i,sor}^E$$

Subject to

$$\begin{aligned} D_i &\leq \sum_j x_{ji}^{H_2} \\ e_i^{intensity} \sum_j x_{ij}^{H_2} &= \sum_{sor} \sum_j x_{j,i,sor}^E \\ \sum_j x_{j,i,sor}^E &\leq e_{i,sor}^{max} \\ \sum_p \lambda_{ip} &= 1 \\ x_{p-1}^L \lambda_{ip} &\leq \Delta x_{ip} \leq x_p^U \lambda_{ip} \\ \sum_j x_{ij}^{H_2} &= \sum_p \Delta x_{ip} \end{aligned}$$

### Level 2: Process Design and Energy Scheduling

The micro-level details focus on operational intricacies within individual counties. This includes determining the roles counties play within the network, scaling production and storage facilities, configuring the energy portfolio, aligning production timing with renewable energy availability, devising energy storage solutions, sizing renewable energy farms for grid independence, and formulating strategies to meet emissions reduction goals.

$$\min TC = C^{iv,tot} + C^{of,tot} + C^{ov,tot} + C^{ov,um}$$

Subject to,

$$\begin{aligned} LCOH &= \frac{C^{iv,tot} + C^{of,tot} + C^{ov,tot}}{\sum_t m_t^{out}} \\ P_t^g + \sum_{ren} P_{ren,t}^{sor} + \sum_i P_{i,t}^S &= P_t^{el} + P_t^{comp} \\ m_t^{dem} &= m_t^{out} + m_t^{um} - m_t^{excess} \\ C^{ov,grid} &= \sum_t \pi_t P_t^g \Delta t \\ C^{ov,um} &= \sum_t \lambda m_t^{um} \Delta t \\ C^{ov,co2} &= \sum_t P_t^g \gamma \varphi \Delta t \\ C^{iv,tot} &= C_{ren}^{iv,sor} + C^{iv,el} + C^{iv,comp} + C^{iv,tank} \\ &\quad + \sum_i C_i^{iv,stor} \\ C^{of,tot} &= C^{of,el} + C^{of,comp} + C^{of,tank} + \sum_i C_i^{of,stor} \\ C^{ov,tot} &= C^{ov,grid} + C^{ov,um} + C^{ov,co2} + \sum_i C_i^{ov,stortot} \end{aligned}$$

Renewable farm constraints,

$$\begin{aligned}
 P_{ren,t}^{sor} &\leq eff_{ren} A_{ren} \\
 Avail_{solar,t} &= ghi_t A_{total} \left( 1 - \max \left( \frac{PD_{th} - PD_{min}}{PD_{max} - PD_{min}}, \frac{PD - PD_{min}}{PD_{max} - PD_{min}} \right) \right) \\
 Avail_{wind,t} &= 0.5 \pi \rho_{air} v_t^3 A_{total} \left( 1 - \max \left( \frac{PD_{th} - PD_{min}}{PD_{max} - PD_{min}}, \frac{PD - PD_{min}}{PD_{max} - PD_{min}} \right) \right) \\
 C_{ren}^{iv,sor} &= CO_{ren} A_{ren} CRF \frac{T}{8760}
 \end{aligned}$$

Electrolyzer constraints,

$$\begin{aligned}
 N^{cell} &= \frac{1000 \Lambda^{el}}{nc_{h2}^{po,cell}} \\
 m_t^{norm} &= \frac{m_t^{el}}{N^{cell} m_{cell,max}^{el}} \\
 P_t^{el} &= (-8.5231 m_t^{norm2} + 23.995 m_t^{norm} + 47.752) m_t^{el} 10^{-3} \\
 C^{iv,el} &= CO^{iv,el} \Lambda^{el} CRF \frac{T}{8760} \\
 C^{of,el} &= CO^{of,el} C^{iv,el}
 \end{aligned}$$

Hydrogen compressor constraints,

$$\begin{aligned}
 P_t^{comp} &= m_t^{el} c_{h2}^{po,comp} \\
 P_t^{comp} &\leq \Lambda^{comp} \\
 C^{iv,comp} &= CO^{iv,comp} \Lambda^{comp} CRF \frac{T}{8760} \\
 C^{of,comp} &= CO^{of,comp} C^{iv,comp}
 \end{aligned}$$

Hydrogen storage constraints,

$$\begin{aligned}
 m_{t+1}^{tank} &= m_t^{tank} + (m_t^{el} - m_t^{out}) \Delta t \\
 0 &\leq m_t^{stor} \leq m^{tank,max} \\
 C^{iv,tank} &= CO^{iv,tank} m^{tank,max} CRF \frac{T}{8760} \\
 C^{of,tank} &= CO^{of,tank} C^{iv,tank}
 \end{aligned}$$

Energy storage model,

$$\begin{aligned}
 \sum_b z_{i,t,b}^{op} &= 1 \\
 \epsilon y_i &\leq x_i \leq E_i^{ub} y_i \\
 0 &\leq E_{i,t} \leq x_i \\
 E_{i,t+1} &= E_{i,t} - (\eta_{i,t}^S z_{i,t,b=c}^{op} + z_{i,t,b=d}^{op}) P_{i,t}^S \Delta t \\
 \eta_{i,t}^S &= n0_i s_{i,t}^{n1_i} \\
 -z_{i,t,b=c}^{op} P_{i,t}^S &\leq P_{i,t}^{c,max} \leq P_{i,t}^{c,ub} y_i \\
 z_{i,t,b=d}^{op} P_{i,t}^S &\leq P_{i,t}^{d,max} \leq P_{i,t}^{d,ub} y_i \\
 \left| \frac{E_{i,t=NT+1} - E_{i,t=1}}{E_{i,t=1}} \right| &\leq cyclot \\
 \left| (-z_{i,t+1,b=c}^{op} + z_{i,t+1,b=d}^{op}) P_{i,t+1}^S - (-z_{i,t,b=c}^{op} + z_{i,t,b=d}^{op}) P_{i,t}^S \right| &\leq rr_i^{stor} \\
 P_{i,t}^S &= f1_{i,t} \\
 E_{i,t} &= f2_{i,t} \\
 s_i^{lb} &\leq s_{i,t} \leq s_i^{ub} \\
 l_i^{lb} &\leq l_{i,t} \leq l_i^{ub} \\
 tr^{lb} &\leq tr_i \leq tr^{ub}
 \end{aligned}$$

$$\begin{aligned}
 C_i^{iv,stor} &= (c11_i x_i^{\alpha11_i} + c12_i P d_i^{\max \alpha12_i} + c13_i P c_i^{\max \alpha13_i}) CRF_i \frac{T}{8760} \\
 C_i^{of,stor} &= (c21_i x_i^{\alpha21_i} + c22_i P d_i^{\max \alpha22_i} + c23_i P c_i^{\max \alpha23_i}) \frac{T}{8760} \\
 C_{i,t}^{ov,stor} &= (c31_i z_{i,t,b=d}^{op} - c32_i z_{i,t,b=c}^{op}) P_{i,t}^S \Delta t \\
 C_i^{ov,stor,tot} &= \sum_t C_{i,t}^{ov,stor}
 \end{aligned}$$

At its core, the model seeks to balance energy production and hydrogen generation across temporal and spatial dimensions, taking into account the variable nature of renewable energy sources and grid electricity prices. The model incorporates decision variables for energy management, renewable energy farms, energy storage, and hydrogen production and storage. These variables are optimized within a system of constraints that ensure energy balance, hydrogen balance, and operational feasibility (see Fig. 2). For example, the power purchased from the electricity grid at any given time is matched against the power consumed by electrolyzers and compressors, ensuring an overall energy balance. The constraints also enforce the physical and operational limitations of the system, such as the maximum hydrogen storage capacity and the power output limits of storage technologies. This detailed formulation allows for the examination of the economic and environmental implications of the supply chain, with the ultimate goal of minimizing costs and emissions while meeting the hydrogen demand.

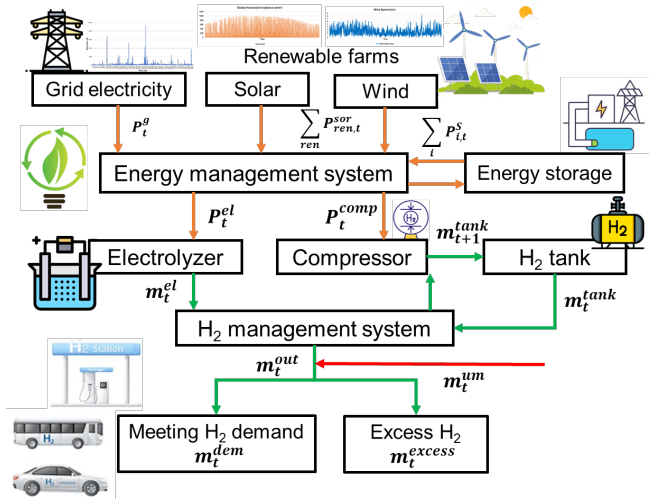
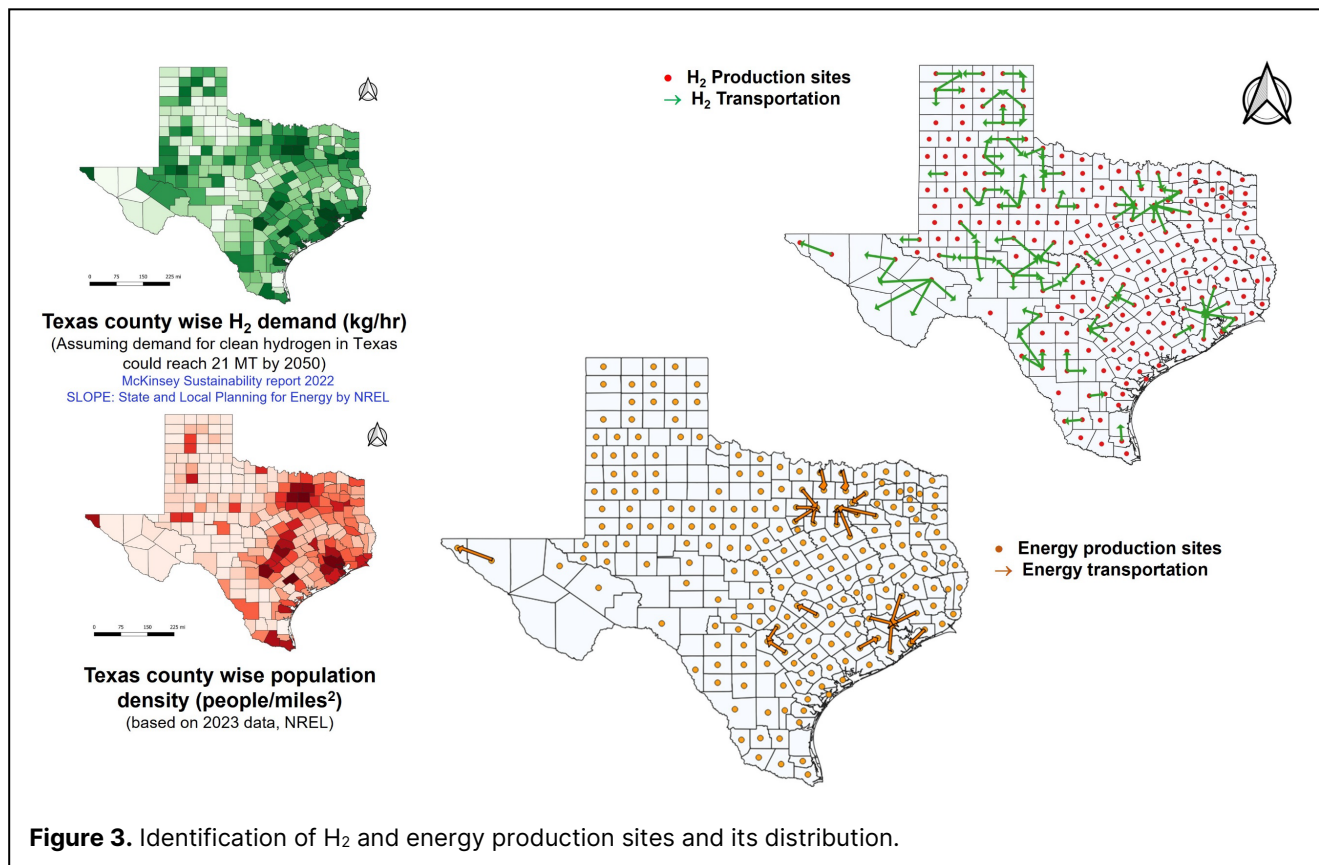


Figure 2. Green hydrogen production system.

This decomposition into two optimization levels allows for a structured breakdown of the complex supply chain problem. The output of level 1, which encapsulates



county-wise energy and hydrogen demand forecasts, informs level 2 decisions. This includes detailed cost assessments for hydrogen production at the county level, which are then fed back into level 1. The iterative feedback loop between the two levels propels the optimization process toward convergence, refining the network configuration with each iteration.

Recognizing the variable nature of energy availability across counties, our model incorporates a robust mechanism for inter-county energy flow. This mechanism dynamically channels surplus energy from counties with excess to those with deficits, thereby maintaining the balance necessary to meet each county's hydrogen production demands. This energy management strategy is integral to our comprehensive approach, affirming that all counties can achieve their hydrogen demand targets through cooperative energy sharing and sophisticated scheduling.

## RESULTS AND DISCUSSION

The broader supply chain optimization achieves strategic positioning of hydrogen production facilities across Texas, with significant concentrations in energy-rich counties. For instance, the production capacities ranged from 0.21 to 21 MT per year, aligning with the variable solar and wind profiles. In Fig. 3, we can see the location of the hydrogen production sites and the

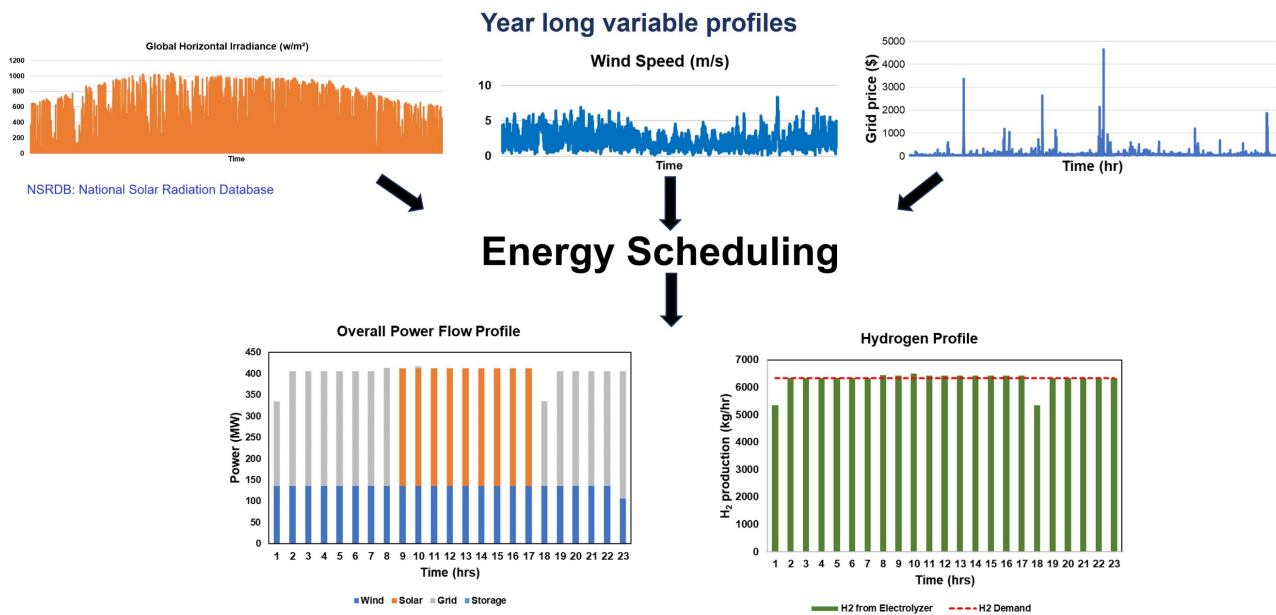
distribution of produced hydrogen to other counties based on the given projected energy demand and population density of each county. We can observe that not every county is producing their H<sub>2</sub> rather few counties are taking the leverage of economics of scale by collaborating with the neighbouring counties. This also gives us the indication that if management or policy makers decided to build up the specific number of hubs for the hydrogen for Texas. Our supply chain optimization can find those locations with some additional constraints. The optimal energy mix was achieved with 35.46% wind, 34.3% solar, and 30.23% grid energy, illustrating a significant reliance on renewable sources.

The simulation results, derived from a year-long variability profile for solar irradiance and wind speed, indicate a consistent alignment between the overall power flow and the hydrogen production profile, which is crucial for maintaining a sustainable energy supply for hydrogen production (see Fig. 4). The hydrogen profile, compared with the demand, reveals that the production from electrolyzers is well-aligned with the demand pattern, suggesting an efficient design of the electrolyzer capacity and operational scheduling. Notably, peak production periods do not always coincide with peak demand times, indicating the necessity for robust storage solutions within the supply chain to balance the temporal discrepancies. In terms of energy supply, the integration of solar and wind energy contributes significantly to the overall



**Table 1:** Selected information on hydrogen production process design.

County	Max installed wind capacity (MW)	Max installed solar capacity (MW)	Installed electrolyzer capacity (MW)	LCOH (\$/kg)	Penalty on unmet H <sub>2</sub> demand (\$)	Total system cost (MM \$)	Emission penalty on grid bought electricity (MM \$)
Austin	528.831	2563.51	1602.3	4.58	-	992	218
Angelina	1.93	286.06	159.73	4.40	-	92	21.2
Archer	194.43	0	68.86	3.39	-	31.5	0.5
Austin	168.38	4.01	61.28	3.51	-	28.8	0.69
Bailey	0	26.38	14.54	4.10	-	6.32	1.18
Bastrop	0	182.41	105.68	4.40	-	61	14.7
Bee	788.58	24.35	290.05	3.73	-	145	8.17
Bell	915.15	55.73	351.74	3.84	-	181	12.4
Blanco	57.03	0	20	3.27	-	8.91	0.02
Borden	0	0	1.80	5.46	-	0.86	0.36



**Figure 4:** Simultaneous design and scheduling model.

power flow, with grid energy supplementing the shortfall. However, the reliance on grid energy varies throughout the day, suggesting potential areas for further optimization of renewable energy sources or storage solutions to minimize grid dependence and enhance sustainability. Furthermore, our optimization model has successfully identified strategic locations for hydrogen production,

factoring in county-specific variables such as land area and energy profiles. This strategic placement, alongside an optimized energy mix, effectively minimizes transport costs and maximizes the use of local renewable energy, supporting the overarching goal of a resilient and sustainable hydrogen economy. Numerical insights obtained from the optimization highlight the potential for a

reduction in carbon footprint through optimized renewable energy use and strategic site placement. The proposed supply chain configuration promises to meet the projected 21 MT hydrogen demand by 2050, with a comprehensive cost analysis indicating a favorable comparison to current conventional hydrogen production costs.

The selective overview of Texas counties reveals a diversified approach to hydrogen production process design (see Table 1). For instance, Austin County showcases a substantial investment in renewable energy sources with impressive wind and solar capacities, facilitating a large-scale electrolyzer capacity that could cater to future hydrogen demands. However, the associated LCOH of \$4.58/kg suggests a higher production cost, potentially due to the scale of installed capacities and emission penalties. Conversely, counties like Archer, with a focus on wind energy, and Blanco, with a conservative renewable approach, indicate a more cost-effective production with their lower LCOH. Borden County's minimal figures might reflect an opportunity for growth or a strategic decision to maintain a small-scale operation. These data points indicate that while some counties are positioning themselves as potential leaders in hydrogen production, others may opt for a scaled approach or are in the early stages of infrastructure development. The diverse strategies underscore the need for a multifaceted, tailored approach in optimizing hydrogen production that balances cost, demand, and environmental impact.

## CONCLUSION

In our study, the placement of strategic hydrogen production sites was pivotal, with locations selected to align local demand with the availability of energy resources. The design of the network integrated insights from county-level contributions and renewable sources, setting a robust foundation for the configuration and scale of these production sites. Our analysis into the energy mix probed the feasibility of utilizing solar, wind, and grid sources to create a flexible energy portfolio for hydrogen production. By adopting an iterative, two-level optimization approach, we enhanced the supply chain model, ensuring economic feasibility and environmental sustainability. Our findings point to a future-adapted hydrogen supply network, resilient and scalable to meet the burgeoning demand and shifts in the energy sector. This synthesis of theoretical insights and numerical analysis underscores the viability of the proposed supply chain configuration, offering a viable pathway to achieving the DOE's goal of \$1 per kilogram of clean hydrogen within a decade.

## ACKNOWLEDGEMENTS

The authors acknowledge partial support for this

work from the Texas A&M Energy Institute.

## REFERENCES

1. Squadrito G, Maggio G, Nicita A. The green hydrogen revolution. *Renewable Energy*. 2023;216:119041. doi:<https://doi.org/10.1016/j.renene.2023.119041>.
2. Oliveira AM, Beswick RR, Yan Y. A green hydrogen economy for a renewable energy society. *Current Opinion in Chemical Engineering*. 2021;33:100701. doi:<https://doi.org/10.1016/j.coche.2021.100701>.
3. The White House. Biden-Harris Administration Announces Regional Clean Hydrogen Hubs to Drive Clean Manufacturing and Jobs. Statements and releases. 2023;<https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/13/biden-harris-administration-announces-regional-clean-hydrogen-hubs-to-drive-clean-manufacturing-and-jobs/> (accessed on 28 November, 2023).
4. DoE. Regional Clean Hydrogen Hubs Selections for Award Negotiations. Office of Clean Energy Demonstrations. 2023;<https://www.energy.gov/oced/regional-clean-hydrogen-hubs-selections-award-negotiations> (accessed on 28 November, 2023).
5. Li L, Manier H, Manier M-A. Integrated optimization model for hydrogen supply chain network design and hydrogen fueling station planning. *Computers & Chemical Engineering*. 2020;134:106683. doi:<https://doi.org/10.1016/j.compchemeng.2019.106683>.
6. Eljack F, Kazi M-K. Prospects and challenges of green hydrogen economy via multi-sector global symbiosis in Qatar. *Frontiers in Sustainability*. 2021;1:612762.
7. Vijayakumar V, Jenn A, Ogden J. Modeling future hydrogen supply chains in the western United States under uncertainties: an optimization-based approach focusing on California as a hydrogen hub. *Sustainable Energy & Fuels*. 2023;7(5):1223-44. doi:10.1039/D3SE00043E.
8. Li L, Manier H, Manier M-A. Hydrogen supply chain network design: An optimization-oriented review. *Renewable and Sustainable Energy Reviews*. 2019;103:342-60. doi:<https://doi.org/10.1016/j.rser.2018.12.060>.
9. Alkatheri M, Alhameli F, Betancourt-Torcat A, Almansoori A, Elkamel A. Clustering Approach for the Efficient Solution of Multiscale Stochastic Programming Problems: Application to Energy Hub Design and Operation under Uncertainty. *Processes* 2023 doi:10.3390/pr11041046.
10. Moran C, Moylan E, Reardon J, Gunawan TA,

Deane P, Yousefian S et al. A flexible techno-economic analysis tool for regional hydrogen hubs – A case study for Ireland. *International Journal of Hydrogen Energy*. 2023;48(74):28649-67. doi:<https://doi.org/10.1016/j.ijhydene.2023.04.100>.

11. Marouani I, Guesmi T, Alshammari BM, Alqunun K, Alzamil A, Alturki M et al. Integration of Renewable-Energy-Based Green Hydrogen into the Energy Future. *Processes* 2023 doi:10.3390/pr11092685.
12. McKinsey & Company. Houston as the epicenter of a global clean-hydrogen hub. McKinsey Sustainability Report. 2022;[https://www.mckinsey.com/capabilities/sustainability/our-insights/houston-as-the-epicenter-of-a-global-clean-hydrogen-hub#/\(accessed on 29 November 2023\)](https://www.mckinsey.com/capabilities/sustainability/our-insights/houston-as-the-epicenter-of-a-global-clean-hydrogen-hub#/).

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Integrated Temporal Planning for Design and Operation of the International Green Ammonia Supply Chain

Sunwoo Kim<sup>a,b</sup>, Jounggho Park<sup>b,c</sup>, Jay H. Lee<sup>a\*</sup>

<sup>a</sup> University of Southern California, Mork Family Department of Chemical Engineering and Material Sciences, Los Angeles, California, USA

<sup>b</sup> Korea Advanced Institute of Science and Technology, Department of Chemical and Biomolecular Engineering, Daejeon, Republic of Korea

<sup>c</sup> Korea Institute of Energy Research, Energy AI & Computational Science Laboratory, Daejeon, Republic of Korea

\* Corresponding Author: [jlee4140@usc.edu](mailto:jlee4140@usc.edu)

## ABSTRACT

This research is dedicated to designing and economically evaluating the green ammonia supply chain, considering the fluctuating nature of renewable energy sources and energy demand across both hourly and seasonal variations. It also explores the impact of economies of scale and the delays associated with long-distance shipping to meet energy demands in a timely manner. These considerations require the formulation of a Mixed-Integer Nonlinear Programming model, further complicated by the necessity for a two-stage stochastic programming approach. We introduce a hierarchical optimization framework that utilizes a decomposition method to differentiate between one-time design decisions and subsequent operational choices. At the upper level, potential design solutions are identified through the Bayesian Optimization and Hyperband algorithm, which effectively navigates the non-linear challenges posed by economies of scale. The lower level then addresses a Mixed-Integer Linear Programming problem to independently assess the feasibility of each scenario. Our empirical analysis includes case studies of three potential international routes for transporting green ammonia to Korea. We contrasted our methodology with a hypothetical scenario that presupposes a constant supply of power and a stable demand for energy. Additionally, techno-economic analyses were conducted to evaluate the implications of the minimum operational limits for electrolyzers.

**Keywords:** Green ammonia supply chain, Multi-timescale decision-making, Integrated temporal approach, MINLP, Decomposition approach

## INTRODUCTION

Amidst the escalating global demand for the sustainable energy solutions, the quest for efficient energy transport materials has taken center stage in research and development. Hydrogen and ammonia have emerged as particularly promising candidates, with ammonia gaining traction in international supply chains due to its advantages in long-distance transportation and storage [1].

Nevertheless, the deployment of ammonia as an energy carrier faces challenges associated with the traditional Harbor-Bosch approach, which exhibits limited flexibility and imposes constraints on ramp-up and ramp-down rates [2]. This lack of adaptability poses a major

obstacle to planning a green ammonia supply chain capable of adeptly responding to dynamic renewable energy fluctuations and ensuring timely transport to meet demand. Furthermore, the seasonality and inherent uncertainty in energy production and demand [3], combined with the extended duration of vessel transportation, mandate operational decisions that span multiple time zones.

In the face of these challenges, existing research lacks a comprehensive integrated temporal approach that considers both hourly profiles and weekly shipping schedules, hindering the determination of reliable and economic capacity design. Addressing this deficiency in the literature, our study introduces an innovative meth-

odology for bi-level decision-making, combining Bayesian optimization and hyperband (BOHB) with mathematical programming. The primary objective of this study is to analyze the impact of incorporating uncertainty and intermittency in the environment, as along with techno-economic parameters related to hourly operations, on the economic viability and design of a green ammonia supply chain. By adopting this integrated approach, we significantly lower the complexities associated with energy management and ship scheduling under multi-timescale uncertainties. This, in turn, provides valuable insights into fostering efficient decision-making within the burgeoning field of green ammonia project.

## PROBLEM DESCRIPTION

As depicted in Figure 1, the focal problem addressed in this paper is to minimize the levelized cost of green ammonia (LCOA) by optimizing the capacity sizes of various components, including turbines, PV panels, batteries, PEM water electrolyzers, Harbor-Bosch plants, hydrogen tanks, ammonia tanks, and ammonia ships. The optimization process takes into account operations spanning multiple timescales. Specifically, ship scheduling is determined on a weekly basis, while operational decisions for the energy management of a green ammonia production facility are made on an hourly basis. The model accommodates time delays arising from long-distance shipping, and it incorporates uncertainty on both the demand and supply sides.

Weather and demand uncertainties are modeled through various scenarios generated based on historical weather patterns and energy consumption data. Given that capacity decisions must be made independently of these scenarios, operational decisions are tailored to respond to specific scenarios, serving as recourse decisions. Thus, the problem is formulated as a two-stage stochastic optimization problem, as illustrated in Figure 2. The first-stage decisions revolve around capacity decisions, while the second-stage decisions focus on operational decisions. The distinct timescales associated with the first- and second-stage decisions contribute to the complexity of solving this problem. Furthermore, the difficulty of solving the problem is compounded by the fact that large-scale capacity design necessitates the consideration of economies of scale. Additionally, shipping-related variables and facility sizes, being discrete entities, introduces integer variables into the mix. Consequently, the mathematical programming is formulated as mixed-integer nonlinear programming (MINLP). To address this complexity, we introduce a novel bi-level decision-making framework, described in the method section.

## SYSTEM DESCRIPTION

As shown in Figure 3, energy is harnessed from both wind turbines and PV panels. Batteries play a role in mitigating the inherent volatility of renewable energy, with 'curtailment' denoting the direct release of excess energy

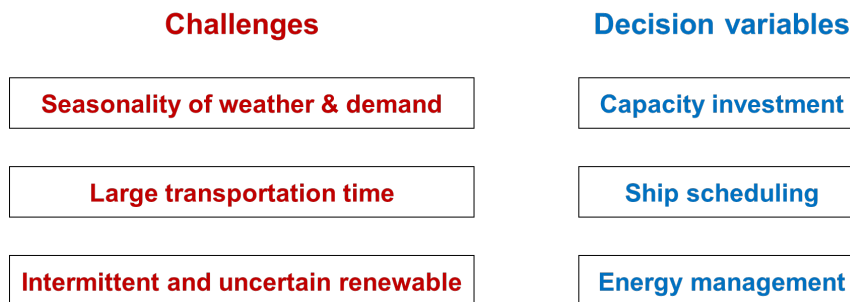
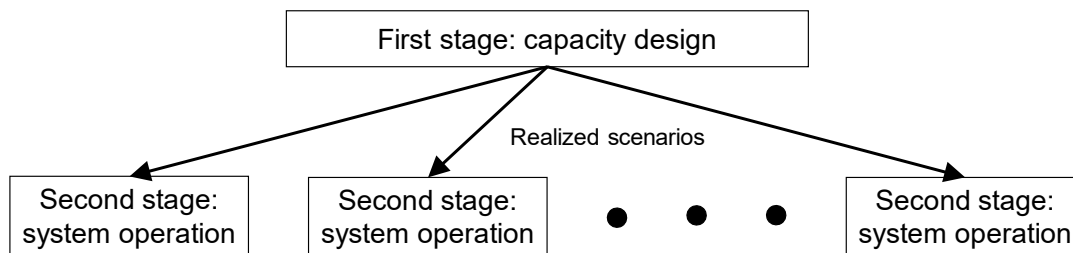


Figure 1. Challenges and multi-timescale decision variables.



### MINLP & stochastic programming

Figure 2. Illustration of two-stage stochastic programming for the problem of this study.



and the management of operations such as idling turbines. Gaseous hydrogen is produced through a water electrolyzer, a process integral to synthesizing ammonia. Any surplus hydrogen generated is stored in a dedicated hydrogen tank. To achieve the minimum load of electrolyzer, the system occasionally relies on the reconversion of hydrogen via a fuel cell. The synthesized ammonia is then stored in tanks before being transported from the exporting country to the importing country via shipping.

For efficient regulation of various current types generated and required by diverse equipment, appropriate converters and current buses are needed. Given the prevalence of DC-type equipment, this study employs a DC bus topology to effectively accommodate the distinctive electrical characteristics.

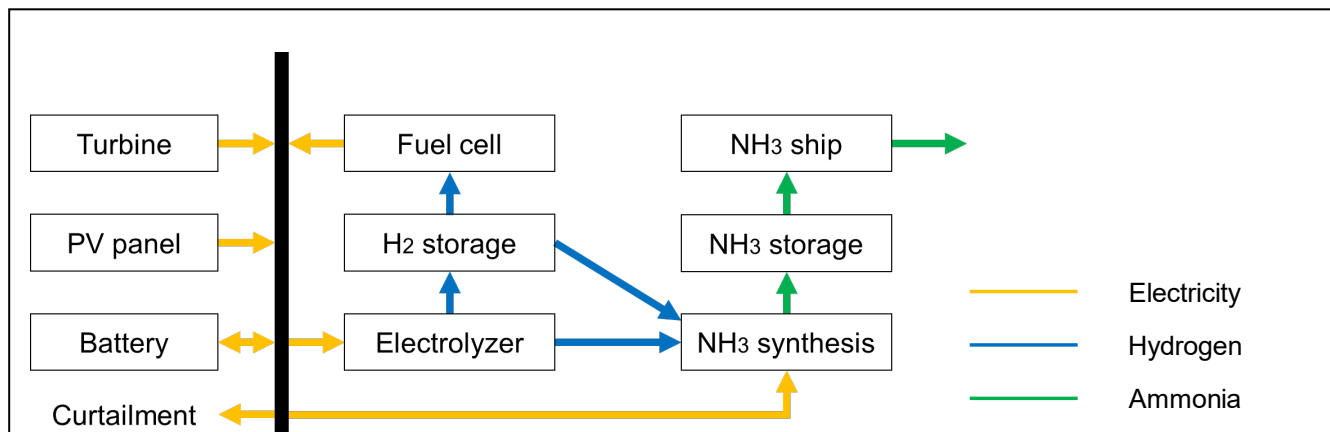
## METHOD

A bi-level decision-making framework is introduced, as illustrated in Figure 4. The initial design decision, facilitated by the BOHB, transforms the remaining optimization into mixed-integer linear programming (MILP). Thus, system operation is addressed through a set of independently manageable sub-problems, markedly lowering

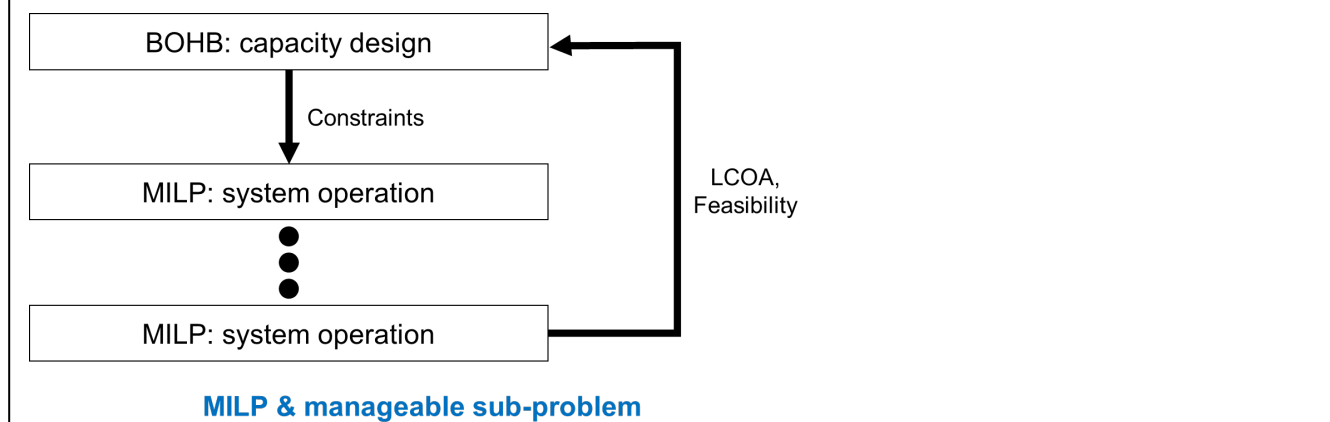
the complexity of the original two-stage stochastic programming.

BOHB, an advanced hyperparameter optimization algorithm, blends the strengths of Bayesian optimization (BO) and hyperband [4]. Known for its ability to achieve both robust and efficient performance in finding the optimal configuration, BOHB overcomes the limitations of traditional methods. While hyperband excels at smaller sampling budgets, it struggles with larger budgets due to its reliance on random sampling. In contrast, BO initially progresses slowly but, given enough time, outperforms hyperband. Combining these approaches, BOHB delivers superior performance and rapid convergence across varying timeframes.

In BO, a surrogate model, often based on Gaussian processes, approximates the distribution of the objective function across decision variables. However, the complexity of Gaussian models grows with increasing samples. BOHB addresses this by using a tree-structured Parzen Estimator (TPE) surrogate model. Leveraging kernel density estimation (KDE), TPE models good and bad hyperparameters separately, providing probability density estimates that account for observed performance [5].



**Figure 3.** Supply system description.



**Figure 4.** Bi-level decision-making framework

Within BO, the acquisition function entails a trade-off between exploration and exploitation, commonly addressed by the expected improvement (EI) function. In TPE, EI relies on estimated probability densities of good and bad performance. Poor-performing configurations are actively avoided in favor of those with high probabilities of good performance. BOHB's TPE model employs a multivariate KDE for probability density estimation, defined by a smoothing function and bandwidth matrix controlling the smoothness of the resulting density curve [6]. BOHB efficiently allocates resources to promising configurations and discards insufficient ones early in the process. Hyperband divides the configuration space into brackets, utilizing consecutive halves within each bracket. BOHB extends hyperband by incorporating model-based search using KDE, guiding searches toward promising regions while retaining hyperband's adaptive resource allocation.

The objective function of the MILP is modified from the original nonlinear form into a linear term involving slack variables, as expressed in the following equation:

$$\sum_t [AP_t^E + AP_t^H + AP_t^A + AC_t^{AIT} + AC_t^{AET}] \quad (1)$$

Here,  $AP_t^E$  represents additional electricity power,  $AP_t^H$  indicates additional hydrogen,  $AP_t^A$  signifies additional ammonia,  $AC_t^{AIT}$  indicates additional capacity of ammonia import tank, and  $AC_t^{AET}$  represents additional capacity of ammonia export tank.

It's important to note that a non-zero objective value from the Mixed-Integer Linear Programming (MILP) analysis indicates that the proposed capacity design is not viable. In instances of unviable capacity designs, the Bayesian Optimization and Hyperband (BOHB) algorithm is informed through the integration of the levelized cost of ammonia (LCOA) with a substantial penalty cost, denoted as  $M$ , to direct the search away from such infeasible design areas. This mechanism is encapsulated in the equation below:

$$receive \begin{cases} LCOA & \text{for feasible design} \\ LCOA + M & \text{for infeasible design} \end{cases} \quad (2)$$

The IBM CPLEX solver is employed to address the MILP problem, which encompasses 122,529 variables and 61,311 constraints.

## RESULT

Three countries—Saudi Arabia, Australia, and Chile—were selected as potential international green ammonia promising trading partners with Korea. A comparative analysis of their economic outcomes are presented in Figures 5-7, and their optimal configurations are presented in Tables 1-3.

In first case (the idealistic case), supply and demand are assumed to remain constant throughout the year (an

assumption often made in prior studies). Figure 5 reveals an overall LCOA ranging between 0.39 and 0.43 \$/Ammonia, with storage costs below 0.01 \$/Ammonia. Australia emerges as the most economically option, boasting the shortest transportation time.

In Figure 6, considering seasonality and uncertain supply and demand (the base case), the overall LCOA fluctuates between 0.59 and 0.68 \$/Ammonia, while storage costs exceed 0.1 \$/Ammonia. Furthermore, the conversion cost couples compared to the first case, suggesting an average operating rate of conversion devices around 50%, not 100%. Saudi Arabia stands out as the most economically promising partner due to its consistent ammonia production capability.

Lastly, Figure 7 examines the design and economics of a scenario where the minimum load constraint of the water electrolyzer is relaxed to 3% (considered an optimistic case, compared to 5% for the base case). The overall LCOA decreases from the base case, as accompanied by a reduction in storage costs. This reduction is attributed to the decrease in battery devices required to meet the minimum load of the water electrolyzer. This highlights the intricate interaction between various factors in the supply chain, aspects often overlooked in many prior studies.

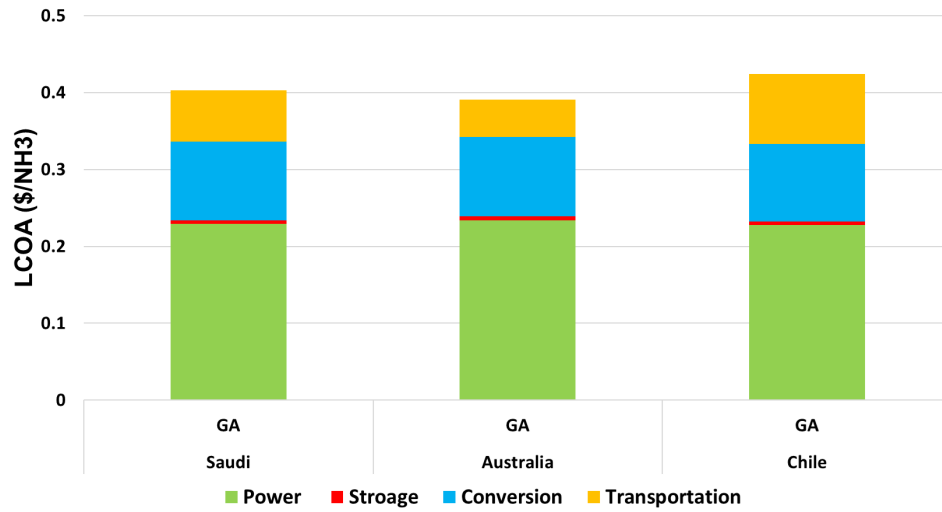


Figure 5. Unrealistic case: constant supply & demand assumption

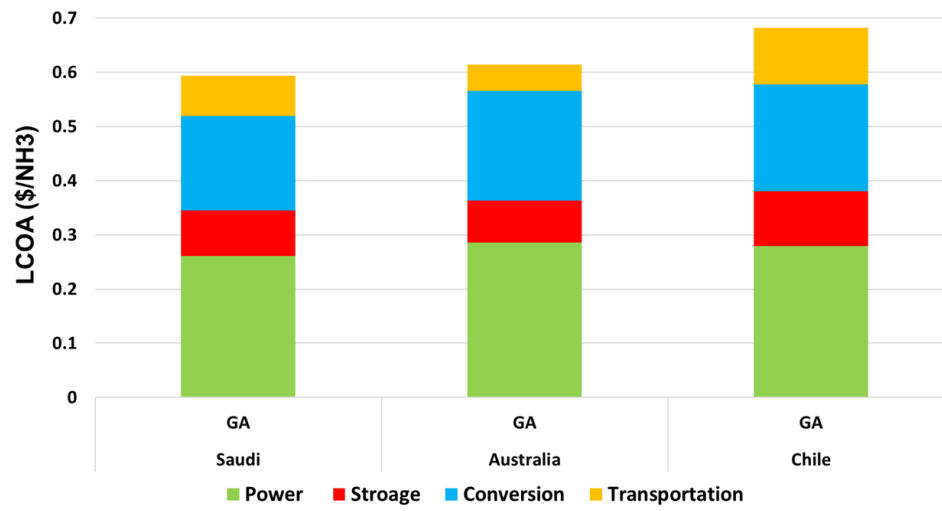


Figure 6. Base case: considering uncertainty & seasonality

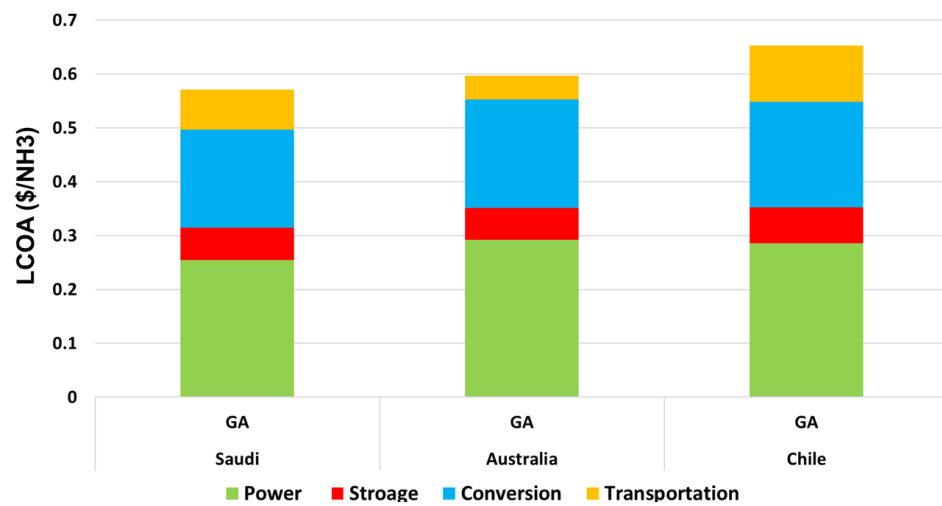


Figure 7. Optimistic case: reduced minimum operating load (=3%) of water electrolyzer

**Table 1.** Idealistic scenario: Optimal configuration of each routes

	T (GW)	P (GW)	W (GW)	B (GWh)	S	HB (Mt A/yr)	H tank (kt)	FC (100 MW)	Ex. A tank (100 kt)	Im. A tank (100 kt)
Saudi	16	0	6	0	8	7	0	0	2	2
Aust.	0	28	6	0	6	7	0	0	2	2
Chile	0	28	6	0	11	7	0	0	2	2

**Table 2.** Base case: Optimal configuration of each routes.

	T (GW)	P (GW)	W (GW)	B (GWh)	S	HB (Mt A/yr)	H tank (kt)	FC (100 MW)	Ex. A tank (100 kt)	Im. A tank (100 kt)
Saudi	13	9	13	11	8	8	5	4	6	4
Aust.	7	22	15	10	6	8	5	5	5	3
Chile	6	24	17	17	11	8	2	3	5	4

**Table 3.** Optimistic case: Optimal configuration of each routes.

	T (GW)	P (GW)	W (GW)	B (GWh)	S	HB (Mt A/yr)	H tank (kt)	FC (100 MW)	Ex. A tank (100 kt)	Im. A tank (100 kt)
Saudi	12	10	14	7	8	8	4	4	6	4
Aust.	8	21	15	7	6	8	4	4	5	3
Chile	7	23	17	10	11	8	2	3	6	3

## CONCLUSION

This paper presents a bi-level optimization framework that integrates BOHB with MILP to streamline temporal planning in the design and management of an international green ammonia supply chain. This method effectively simplifies the complexities found in the original MILP and stochastic programming approaches. The research delves into the impacts of uncertainty and techno-economic parameters via three case studies, highlighting the significance of incorporating fast-time-scale operations for the development of a reliable and cost-efficient supply chain strategy. Additionally, it investigates the influence of the minimum operating load of water electrolyzers on the green ammonia supply chain's efficiency and feasibility.

This study bridges a crucial knowledge gap by revealing the interactions between various supply chain components through an integrated temporal planning approach. It underscores the necessity of considering real-time operations, which involve making decisions without future insights, as crucial for achieving a realistic and practical supply chain implementation. This acknowledgment points toward the importance of adaptability and responsiveness in managing the intricacies of the green ammonia supply chain in the face of operational and market uncertainties.

## ACKNOWLEDGEMENTS

This work was supported by KAIST-Aramco CO<sub>2</sub> management center.

## REFERENCES

- Zhang, H., Wang, L., Maréchal, F., & Desideri, U., Techno-economic comparison of green ammonia production processes. *Applied Energy*, 259, 114135 (2020)
- Armijo, J., & Philibert, C., Flexible production of green hydrogen and ammonia from variable solar and wind energy: Case study of Chile and Argentina. *International Journal of Hydrogen Energy*, 45(3), 1541-1558 (2020)
- Kim, S., Choi, Y., Park, J., Adams, D., Heo, S., & Lee, J. H., Multi-period, multi-timescale stochastic optimization model for simultaneous capacity investment and energy management decisions for hybrid Micro-Grids with green hydrogen production under uncertainty. *Renewable and Sustainable Energy Reviews*, 190, 114049 (2024)
- Falkner, S., Klein, A., & Hutter, F., BOHB: Robust and efficient hyperparameter optimization at scale. In *International conference on machine learning* (pp. 1437-1446). PMLR (2018)
- Bergstra, J., Bardenet, R., Bengio, Y., & Kégl, B., Algorithms for hyper-parameter optimization, *Advances in Neural Information Processing Systems*, 24 (2011)
- Klein, A., Falkner, S., Springenberg, J. T., & Hutter, F., Learning curve prediction with Bayesian neural

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Integration of a Chemical Heat Pump with a Post-combustion Carbon Capture Sorption Unit

Rajalakshmi Krishnadoss and Thomas A. Adams II\*

Norwegian University of Science and Technology, Department of Energy and Process Engineering, Trondheim, Norway

\* Corresponding Author: [thomas.a.adams@ntnu.no](mailto:thomas.a.adams@ntnu.no).

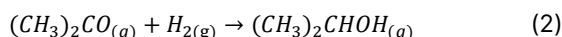
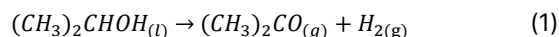
## ABSTRACT

A novel process system which integrates an isopropanol-based chemical heat pump with a post-combustion carbon capture unit was proposed, designed, and analyzed. The system uses low-quality waste heat ( $\sim 80^\circ\text{C}$ ) produced through the  $\text{CO}_2$  adsorption step of a carbon capture process and upgrades that heat to a higher temperature ( $\sim 150^\circ\text{C}$ ) using the chemical heat pump. The chemical heat pump is powered mostly by the waste heat and requires only a small amount of electricity. The higher temperature heat produced can be used in the desorption stage of the  $\text{CO}_2$  capture process, displacing a portion of the existing fossil energy required. The energy and exergy performance characteristics of the chemical heat pump were computed using the results of a steady state simulation in a systems analysis. Using exergy cost correlations, the profitability of the chemical heat pump concept was estimated. It was found that for this particular configuration, the fossil energy load of desorption could be reduced by roughly 2.7% with very little parasitic electric load.

**Keywords:** Energy Efficiency, Exergy Efficiency, Heat integration, Chemical heat pump

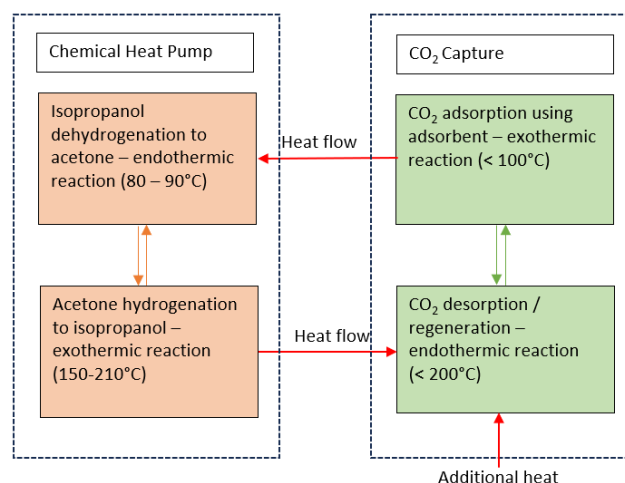
## 1. INTRODUCTION

Chemical heat pumps convert low temperature waste heat into high temperature heat through reversible chemical reactions. Compared to conventional compression-based heat pumps, chemical heat pumps consume far less electricity. One widely researched organic chemical heat pump is the isopropanol acetone system, mainly due to its wide and useful temperature range:



Isopropanol dehydrogenation (Equation 1) is an endothermic reaction operating in the temperature range  $80^\circ\text{C}$ – $90^\circ\text{C}$ . The reverse (acetone hydrogenation, Equation 2) is exothermic and between  $150^\circ\text{C}$ – $210^\circ\text{C}$  [1]. The heat of endothermic reaction is  $100.4 \text{ kJ/mol}$  of isopropanol reacted whereas the heat of exothermic reaction is  $55 \text{ kJ/mol}$  of isopropanol formed [1]. Additional heat is needed for the endothermic reaction since it involves vaporising the isopropanol to the gas phase for the reaction. Previous isopropanol acetone system studies looked at individual aspects, such as catalyst

nuances or the effects of reactor design parameters on performance/efficiency. However there have been very few studies which look at the integration of isopropanol chemical heat pumps with other processes.



**Figure 1.** Chemical heat pump integration with adsorption based  $\text{CO}_2$  capture.

Post-combustion carbon capture involves removing carbon dioxide from flue gas generated during combustion of fuel. The most common technology used for this type of carbon capture is amine absorption. Sorption techniques employing solids materials are in the research & development phase. These materials include polyethyleneimine and dry sorbents. We chose a potassium-based dry sorbent carbon capture processes for this study because the temperature of its waste heat produced and its high-quality heat consumed matches very well with an isopropanol based chemical heat pump.

The carbonation process of the potassium-based sorbent is given the following equation.



The reverse reaction is the regeneration process. In the real system, the heat required for regeneration is more than the heat released during the carbonation process. The regeneration heat comprises of the reaction heat of regeneration as well as the sensible heat which is required to heat the sorbent to the regeneration temperature [5].

Park *et al.* [6] carried out post combustion carbon capture using a potassium-based dry sorbent in a 0.5MW pilot plant. In their work, flue gas from the coal fired power plant was around 80°C, so the carbonation temperature was also maintained the same. The regeneration temperature was in the range 150°C – 200°C. Yi *et al.* [7] carried out carbon capture using a potassium-based sorbent in a bench scale fluidized bed reactor. They have also maintained the carbonation temperature of 80°C and regeneration temperature of 150°C – 230°C.

Here we propose a novel integration of an isopropanol heat pump with a post-combustion sorption-based carbon capture process. In this work we designed the chemical heat pump system, simulated it in Aspen Plus, and analyzed its performance and suitability for this particular application.

## 2. SYSTEM DESCRIPTION

The integrated system is shown in Figure 1. In the carbon capture process, CO<sub>2</sub> from flue gas gets chemisorbed by the sorbent, which is a mild exothermic process (less than 100°C). The regeneration of the sorbent is an endothermic process (less than 200°C) [3].

The sorption energy is generally considered waste as this is low grade heat and is difficult to recover. In the proposed integrated system, the low grade waste heat from the adsorber is upgraded to high temperature heat used for the desorber. Thus, the chemical heat pump can partly supply the energy required for the sorbent regeneration.

There is some literature which reports the simulation of isopropanol chemical heat pumps. There are two

different configurations explored in the literature. The first configuration involves an exothermic reactor, a distillation column, and an endothermic reactor. Here the products (mixture of isopropanol, acetone, and hydrogen) from the exothermic reactor gets separated in the distillation column. The distillate (acetone and hydrogen) goes to the exothermic reactor. The bottoms product (majorly isopropanol) goes back to the endothermic reactor. The second configuration involves the reboiler acting as an endothermic reactor. In this case there is no separate endothermic reactor. Both are represented in Figures 2 and 3.

Configuration 1 requires a greater amount of heat consumption per exothermic heat produced compared to configuration 2. Configuration 2 takes the advantage of the fact that the endothermic reaction temperature and the boiling temperature of isopropanol can be very similar. By synerizing both into the same unit, the total heat load of the system is reduced, because it avoids “boiling the isopropanol twice.”

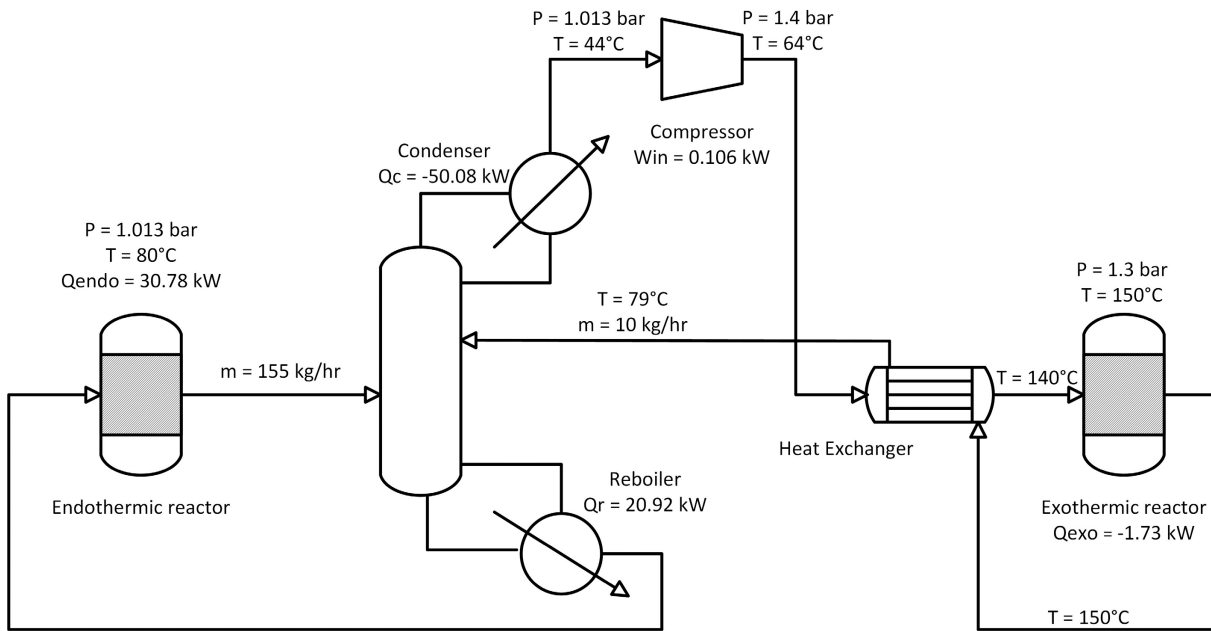
In both designs, the pressure of the exothermic reactor is a little bit higher (1.3 bar) than that of the distillation column (atmospheric). This is because the exothermic reaction equilibrium is more favourable at higher pressure. The optimal pressure, however, is uncertain because of the parasitic electric load of the compressor required. For chemical heat pumps, very low electricity consumption is desirable, as the low-electricity quality is an intentional feature. Related to this, the pressure of the distillation column is atmospheric because the boiling point of isopropanol at atmospheric pressure closely matches the available waste heat temperature.

## 3. METHODOLOGY

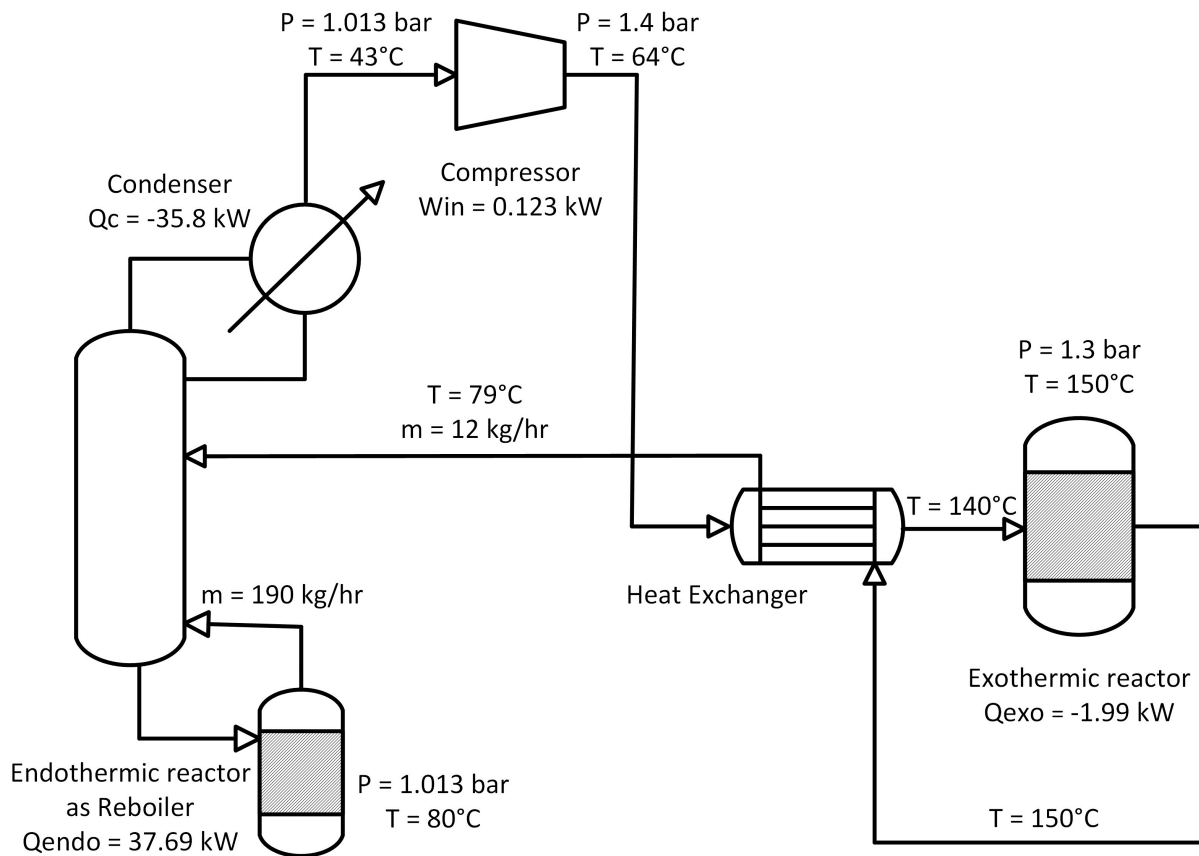
A steady state simulation was created for both configurations using the software Aspen Plus. Configuration 1 was based on the design of [8] but was modified to match the temperature ranges of our particular carbon capture application. The results were similar to those reported in the previous works. Configuration 2 was based on the design of [3] but modified to meet our application-specific requirements. One particularly important change is that we used an equilibrium-based model for the reaction that as will be noted later, results in more conservative reaction yields than those predicted in the reference paper.

### 3.1 Modelling Details

The RADFRAC equilibrium-based model was used for the distillation column. The column design parameters were chosen using the methodology for pre-economic binary distillation column design described in [9]. In this



**Figure 2:** Configuration 1 of isopropanol chemical heat pump



**Figure 3:** Configuration 2 of isopropanol chemical heat pump

methodology, the reflux ratio is adjusted to match distillate purity specifications, and for Configuration 1 the distillate-to-feed ratio is fixed based on obvious desired

recoveries. The initial number of stages above and below the feed was chosen based on the results of the prior work, and then adjusted upward or downward based on

the separate efficiency of each stage such that unnecessary or low-impact stages were eliminated. Finally, the distillate purity specification, which is a degree of freedom, was modified in a sensitivity analysis to understand the impact on the performance of the overall system, repeating the above procedure with each modification. The trade-offs are that high distillate purities yield higher exothermic reaction yields, but at a higher cost of reboiler duty, with diminishing returns. The final selected values were chosen heuristically to balance the distillation heat load with the resulting Coefficient of Performance (COP) of the system as a whole. This method gives very good results and is appropriate for this scope of conceptual design and analysis. A more rigorous optimization approach could be taken in the future given more detailed economic criteria, but the results are not expected to deviate much.

**Table 1:** Specification of the simulation for configuration 1 and 2

Description	Value
Distillation column pressure	Atmospheric pressure
No of stages	7
Feed location	4
Isopropanol mol fraction in distillate (by varying reflux ratio)	0.08
Exothermic reactor temperature	150°C
Exothermic reactor pressure	1.3 bar
Endothermic reactor pressure for configuration 1	Atmospheric pressure
Endothermic reactor temperature for configuration 1	80°C
Distillate to feed ratio for configuration 1	0.11
Compressor discharge pressure	1.4 bar
Heat exchanger: Hot inlet – cold outlet temperature difference	10°C

The *RGIBBS* reactor model was chosen for the exothermic reactor in both configurations and the endothermic reactor in Configuration 1. For Configuration 2, the *RGIBBS* model was used for the reboiler, and the *RADFRAC* model used no reboiler. All *RGIBBS* models used assumed isothermal (at a specified temperatures matching the integrated carbon capture absorber and desorber) and isobaric (no pressure drop) conditions.

The compressors used the *COMP* model, assumed default isentropic efficiencies of 0.72, and required only a single stage. The compressor discharge pressure was chosen such that after upstream pressure drop, the

exothermic reactor pressure was the value recommended in [10] of 1.3 bar.

The *HEATX* model was used for the heat exchanger, with the specification that the cold outlet / hot inlet temperature difference is 10°C, a common specification for pre/post reactor economizers.

Because this study does not consider economics, and all distillation and reaction models are based on equilibrium assumptions, the scale of the simulation is arbitrary, and performance criteria such as efficiency or coefficient of performance does not change with scale. Therefore, all system analysis results are reported on a normalized basis. Furthermore, because phase and reaction equilibrium is assumed, these results are the “best” possible outcome for the reported configuration. A summary of the main specifications are given in Table 1.

## 4. RESULTS

The key results of both simulations are shown in Figures 2 and 3.

### 4.1 Performance Analysis

The COP and exergy efficiency are used to evaluate the performance of the isopropanol chemical heat pump. The COP of this chemical heat pump is calculated according to the formula:

$$COP = \frac{Q_{exo}}{Q_{endo}} \quad (4)$$

where  $Q_{exo}$  is the rate of high temperature heat produced (the exothermic reactor heat duty) in kW and  $Q_{endo}$  is the rate of low temperature heat consumed (the sum of the reboiler and if it exists, the endothermic reactor heat duty) in kW. To compare the COP of the chemical heat pump with a conventional electric heat pump, the COP of the heat pump based on electrical work input is defined:

$$COP_W = \frac{Q_{exo}}{W_{in}} \quad (5)$$

where  $W_{in}$  is the electrical work input in kW. For the isopropanol chemical heat pump this is the compressor power input.

The COP considers only the quantity of heat. Exergy includes both the quantity and the quality of the heat. System exergy input is defined by the below formula:

$$Exergy_{in} = Q_{endo} \left(1 - \frac{T_{ref}}{T_{endo}}\right) + W_{in} \quad (6)$$

where  $T_{ref}$  is the reference temperature in K and  $T_{endo}$  is the endothermic reaction temperature in K. System exergy output is defined by the below formula:

$$Exergy_{out} = Q_{exo} \left(1 - \frac{T_{ref}}{T_{exo}}\right) \quad (7)$$

where  $T_{exo}$  is the exothermic reaction temperature in K. System exergy efficiency is given by the formula:

$$\eta = \frac{Exergy_{out}}{Exergy_{in}} \quad (8)$$

The input values required for COP and exergy calculations are taken from the simulation. The results are presented in Table 2:

**Table 2:** Performance parameters

Description	Config. 1	Config. 2
COP	0.033	0.0525
COP <sub>w</sub>	16.26	16.08
Exergy in (ref temp 25°C)	8.16	6.00
Exergy out (ref temp 25°C)	0.51	0.58
Exergy efficiency	0.063	0.0975
Net profit, \$ per GJ low temp. heat	0.11	0.18

The calculated COP of Configuration 2 is less than what is reported in the literature for a similar case [2] since the equilibrium conversion of the endothermic reaction in our simulation is around 4%, whereas in the literature it is around 7%. Unfortunately, it is unclear whether how that number was obtained or if it was assumed as a parameter in that work. Both works used very similar temperature, pressure, and reagent stoichiometric ratio conditions in the exothermic reactor, so this difference cannot be explained by differences in feed conditions. It should be noted that our model uses an equilibrium assumption that in theory should result in the upper bound on reactor conversion in a real situation without reactor modifications, such as a membrane to remove reaction products in-situ as the reaction progresses.

Increasing the isopropanol purity in the distillate or changing the distillation column pressure beyond the final case reported in Table 1 did not have much effect on COP. The main parameter that governs the COP and the exergy efficiency of the isopropanol chemical pump is the endothermic reaction conversion. The COP<sub>w</sub> of this chemical heat pump is around 16 whereas it is typically around 2 for traditional electric heat pumps. The main advantage of the chemical heat pump is that it utilises significantly less electrical power to produce the temperature lift. Thus chemical heat pumps can play an important role for a future low carbon economy.

## 4.2 Economic Approximation Using Exergy

Exergy is used as a key metric to estimate the profitability of the isopropanol chemical heat pump. A recent analysis has shown that the price of heating utilities is linearly correlated to the exergy of that utility when the reference temperature chosen is "plant ambient" conditions, or about 78°C. In other words, in a typical chemical plant, the value of waste heat at 78°C is on average zero, and above that is valued at \$29.9 (USD<sub>2022</sub>) per GJ of exergy

of heat relative to 78°C [4]. Therefore, we can estimate the net profit from this chemical heat pump by the below formula.

$$net\ profit = value\ of\ product - electrical\ cost \quad (9)$$

where net profit is expressed as \$/GJ of low temperature heat, value of product is \$/GJ of low temperature heat and electrical cost is \$/GJ of low temperature heat. The net profit does not include the value of low temperature waste heat input as it is currently not recovered at all. The value of the high temperature heat product is estimated [4]:

$$value\ of\ product = \frac{Exergy_{out}}{Q_{endo}} \times 29.9 \quad (10)$$

where the constant 29.9 is in \$/GJ of exergy.  $Exergy_{out}$  is calculated using the reference temperature 78°C. The cost associated with the electrical input is calculated as:

$$electrical\ cost = \frac{W_{in}}{Q_{endo}} \times 27.77 \quad (11)$$

where the constant 27.77 is in \$/GJ of electricity. Substituting the simulation results in the above equations, we get the net profit of this chemical heat pump as shown in Table 2.

## 4.3 Economic assessment of the integrated system

As illustrated in the system description section, the endothermic energy required for the isopropanol chemical heat pump is given by the carbon capture carbonation process. The heat required for the carbon capture regeneration process is given by the exothermic heat of the isopropanol chemical heat pump.

Heat generated by the carbonation process of potassium based sorbent is,  $H_{carb} = 141.16$  kJ/mol [5]. The exothermic heat generated by the isopropanol chemical heat pump is given by the below formula.

$$Q_{exo} = H_{carb} \times COP \times F \quad (12)$$

where  $F$  is the carbon capture flow rate in mol/s. The endothermic heat required for carbon capture regeneration is given by the following formula.

$$Q_{regen} = (H_{carb} \times F) + (T_{regen} - T_{carb}) \times \frac{C_p}{1000} \times \frac{F}{A} \quad (13)$$

where  $Q_{regen}$  is the heat required for regeneration of sorbent in kW,  $T_{regen}$  is the sorbent regeneration temperature in °C,  $T_{carb}$  is the sorbent carbonation temperature in °C,  $C_p$  is the specific heat capacity of sorbent with a value of 830 J/kg K [11],  $A$  is the absorption capacity of sorbent with a value of 0.43 mol CO<sub>2</sub>/kg sorbent [5]. The purpose of this integration is to supply part of the regeneration energy for the carbon capture process. The percentage energy savings resulting from this integration is given by the following formula.



$$\% \text{ energy savings} = \frac{Q_{\text{exo}}}{Q_{\text{regen}}} \times 100 \quad (14)$$

When  $T_{\text{carb}} = T_{\text{endo}} = 80^\circ\text{C}$  and  $T_{\text{regen}} = T_{\text{exo}} = 150^\circ\text{C}$ ,  $\% \text{ energy savings} = 2.68\%$ . Here the temperature difference between the chemical heat pump and the carbon capture is kept zero and the resulting energy savings is the theoretical maximum. Altering the chemical heat pump's endothermic and exothermic temperature to maintain some temperature difference would result in reduction of energy savings. The estimation of the carbon capture rate of the potassium-based sorbent at  $80^\circ\text{C}$  and the regeneration efficiency at  $150^\circ\text{C}$  is beyond the scope of this work.

## 5 CONCLUSIONS

The performance of a thermochemical heat pump and its integration with potassium-sorbent-based post-combustion carbon capture was studied. The limiting step in the isopropanol chemical heat pump is the isopropanol dehydrogenation to acetone and hydrogen. Any process improvements to increase the isopropanol conversion to acetone and hydrogen would increase the COP and exergy efficiency. Potassium-sorbent-based carbon capture was selected for integration with this chemical heat pump due to the matching operating temperature. Around 2.7% of thermal regeneration energy savings is possible when both the systems are operated in the temperature range  $80^\circ\text{C} - 150^\circ\text{C}$ . Ultimately, we can recover about 10% of the value (exergy) of the low temperature heat as high temperature heat. The dollar savings would increase when the capacity of the carbon capture unit increases and hence this integration could be economical for large scale systems.

## DIGITAL SUPPLEMENTARY MATERIAL

The Aspen Plus simulation files of the isopropanol chemical heat pump is available in LAPSE at <http://PSEcommunity.org/LAPSE:2023.36832>.

## ACKNOWLEDGEMENTS

The authors would like to thank the Energy and Process Engineering Department at the Norwegian University and Science Technology for financial support.

## REFERENCES

1. KlinSoda I, Piumsomboon P. Isopropanol–acetone–hydrogen chemical heat pump: A demonstration unit. *Energy Conversion and Management* 48(4): 1200-1207 (2007)
2. Spoelstra S, Haije WG, Dijkstra JW. Techno-economic feasibility of high-temperature high-lift chemical heat pumps for upgrading industrial

waste heat. *Applied Thermal Engineering* 22(14): 1619-1630(2002)

3. Bonaventura D, Chacartegui R, Valverde JM, Becerra JA, Ortiz C, Lizana J. Dry carbonate process for CO<sub>2</sub> capture and storage: Integration with solar thermal power. *Renewable and Sustainable Energy Reviews* 82:1796-1812 (2018)
4. Deng L, Adams TA II, Gundersen T. *Exergy Tables*. McGraw-Hill (2023). ISBN 9781264715725
5. Wu Y, Chen X, Ma J, Wu Y, Liu D, Mi J *et al*. System integration for coal-fired power plant with post combustion CO<sub>2</sub> capture: Comparative study for different solid dry sorbents. *Fuel* 280: 118561 (2020)
6. Park YC, Jo SH, Ryu CK, Yi CK. Demonstration of pilot scale carbon dioxide capture system using dry regenerable sorbents to the real coal-fired power plant in Korea. *Energy Procedia* 4: 1508-1512 (2011)
7. Yi CK, Jo SH, Seo Y, Lee JB, Ryu CK. Continuous operation of the potassium-based dry sorbent CO<sub>2</sub> capture process with two fluidized-bed reactors. *International Journal of Greenhouse Gas Control* 1(1): 31-36 (2007)
8. Guo J, Huai X, Xu M. Thermodynamic analysis of an isopropanol–acetone–hydrogen chemical heat pump. *International Journal of Energy Research* 39(1): 140-146 (2015)
9. Adams TA II. *Learn Aspen Plus in 24 Hours*. McGraw-Hill (2022). ISBN 9781264266654
10. Xu M, Cai J, Guo J, Huai X, Liu Z, Zhang H. Technical and economic feasibility of the Isopropanol-Acetone-Hydrogen chemical heat pump based on a lab-scale prototype. *Energy* 139: 1030-1039 (2017)
11. Kant K, Shukla A, David MJS, Rindt CM. Performance analysis of a K<sub>2</sub>CO<sub>3</sub>-based thermochemical energy storage system using a honeycomb structured heat exchanger. *Journal of Energy Storage* 38: 102563 (2021)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# IDAES-PSE Software Tools for Optimizing Energy Systems and Market Interactions

Daniel J. Laky<sup>a</sup>, Radhakrishna Tumbalam Gooty<sup>b,c</sup>, Tyler Jaffe<sup>b,c</sup>, Marcus Holly<sup>b,c</sup>, Adam Atia<sup>b,c</sup>, Xinhe Chen<sup>a</sup>, and Alexander W. Dowling<sup>a\*</sup>

<sup>a</sup> University of Notre Dame, Department of Chemical and Biomolecular Engineering, South Bend, IN 46556, United States

<sup>b</sup> National Energy Technology Laboratory (NETL), Pittsburgh, PA 15236, United States

<sup>c</sup> NETL Support Contractor, Pittsburgh, PA 15236, United States

\* Corresponding Author: [adowling@nd.edu](mailto:adowling@nd.edu)

---

## ABSTRACT

Modern power grids coordinate electricity production and consumption via multi-scale wholesale energy markets. Historically, levelized cost metrics were the *de facto* standard for techno-economic analyses of energy systems and comparison of technology options. However, these metrics neglect the complexity of energy infrastructure including the time-varying value of electricity. An emerging alternative is multi-period optimization, which considers the locational marginal price of electricity as input data (parameters). In this work, we present a general interface for multi-period optimization with time-varying energy prices to facilitate rapid analysis and comparison of potential energy systems models. The PriceTakerModel class is written in the IDAES®-PSE platform and allows users to generate a multi-period, price-taker model instance, as well as automatically generate common operational constraints for their model, such as start-up and shutdown. We show this interface successfully generates multi-period price-taker models, facilitates model discrimination, and aids in analyzing various technologies for deployment in unique energy markets.

---

**Keywords:** Integrated Energy Systems, Software Design, Optimization, Process Operations, Process Design, Electricity Markets

## INTRODUCTION

Modern power grids coordinate electricity production and consumption via multi-scale wholesale energy markets. These multi-faceted markets set the time-varying value of electricity and other products. Mathematical programming has been used to varying degrees of success depending on model complexity and considerations to maximize the profitability of new and existing energy systems. Complexity is garnered by a few factors:

1. the extent to which the grid is impacted by energy systems behavior,
2. the flexibility of the energy system to respond to grid/market signals, and
3. the extent to which the energy system considers grid behavior in design and operation, e.g., production cost models to capture energy system/grid interactions

This short paper begins by comparing modeling approaches to optimize integrated energy systems in the context of energy markets. Next, we present a generalized interface, the PriceTakerModel class, to streamline these analyses in the IDAES-PSE platform.

### Levelized Cost of Electricity

Levelized cost analysis is the predominant strategy for techno-economic analysis (TEA) [1]. In this method, levelized cost of electricity (LCOE) is used as a market model when determining optimal energy system design and operation. Design decisions consider the size of the equipment needed to construct a new energy system or retrofit an existing one. Operational decisions consider the day-to-day production of electricity and other products provided to the grid. Below is the general mathematical model for TEA using LCOE that simultaneously considers design and operating decisions:

$$\min_{d,u} \text{LCOE} = f_{\text{LCOE}}(d,u) \quad (1)$$

$$\text{s. t. } h_j^{\text{des}}(d) = 0, \quad \forall j \in \mathcal{J}^{\text{des}} \quad (2)$$

$$g_k^{\text{des}}(d) \leq 0, \quad \forall k \in \mathcal{K}^{\text{des}} \quad (3)$$

$$h_j^{\text{op}}(d,u) = 0, \quad \forall j \in \mathcal{J}^{\text{op}} \quad (4)$$

$$g_k^{\text{op}}(d,u) \leq 0, \quad \forall k \in \mathcal{K}^{\text{op}} \quad (5)$$

Where  $f_{\text{LCOE}}$  is the mathematical expression for leveled cost of electricity based on design decisions,  $d$ , and operational decisions,  $u$ . Equation sets  $h_j^{\text{des}}$  and  $h_j^{\text{op}}$  represent model equations of design and operational systems, respectively. Here  $j$  is the index of equations in set  $\mathcal{J}^{\text{des}}$  and  $\mathcal{J}^{\text{op}}$  for design and operational equations, respectively. Inequality constraints  $g_k^{\text{des}}$  and  $g_k^{\text{op}}$  represent model and system inequalities for design and operational systems, respectively. Similarly,  $k$  is the index of inequality constraints for design and operational systems in sets  $\mathcal{K}^{\text{des}}$  and  $\mathcal{K}^{\text{op}}$ , respectively.

The major drawback of this formulation is that LCOE does not consider the time-varying nature of the electricity market. As a result, such analyses often lead to incorrect or misleading comparisons of potential energy systems technologies [2].

### Price-taker

Moving beyond LCOE, state-of-the-art TEA uses multi-period optimization to determine the best design and operating policies while considering dynamic market signals (e.g., prices). Multi-period optimization using the pseudo steady-state, price-taker, and self-schedule assumptions is the most popular approach [2-6]. First, time is discretized based on the market, e.g., 1-hour timesteps to match prices from a day-ahead market. During each time period, the energy system is modeled as operating at steady state with constraints for ramping and energy storage, as shown below:

$$\max_{d,u_t,y_t} \text{NPV} = f_{\text{NPV}}(d,u_t,y_t) \quad (6)$$

$$\text{s. t. } h_j^{\text{des}}(d) = 0, \quad \forall j \in \mathcal{J}^{\text{des}} \quad (7)$$

$$g_k^{\text{des}}(d) \leq 0, \quad \forall k \in \mathcal{K}^{\text{des}} \quad (8)$$

$$h_j^{\text{op}}(d,u_t,y_t) = 0, \quad \forall j \in \mathcal{J}^{\text{op}}, t \in \mathcal{T} \quad (9)$$

$$g_k^{\text{op}}(d,u_t,y_t) \leq 0, \quad \forall k \in \mathcal{K}^{\text{op}}, t \in \mathcal{T} \quad (10)$$

$$r_t(u_t, u_{t-1}) \leq 0, \quad \forall t \in \mathcal{T} \quad (11)$$

$$s_t(y_t, y_{t-1}) \leq 0, \quad \forall t \in \mathcal{T} \quad (12)$$

$$y_t \in \{0, 1\} \quad (13)$$

Where operational variables,  $u_t$ , have now been indexed by time along the entire time horizon,  $t \in \mathcal{T}$ , to account for operational decisions that take advantage of the time-varying market. Note that design and

operational equations and inequalities (Eqs. (7) through (10)) are the same functional form as in the leveled cost analysis with the adjustment for time-varying inputs. Accompanying these dynamic decisions are binary variables  $y_t$  which dictate startup, shutdown, and technological decisions. Since the system is now dynamic, constraints  $r_t$  enforce ramping rate limitations and constraints  $s_t$  enforce startup and shutdown of units. Also, more sophisticated metrics such as net present value (NPV) are used to evaluate the economic viability of an optimal design and operating strategy. For example, Dowling and colleagues [2] discuss revenue opportunities from real-time market and ancillary service products.

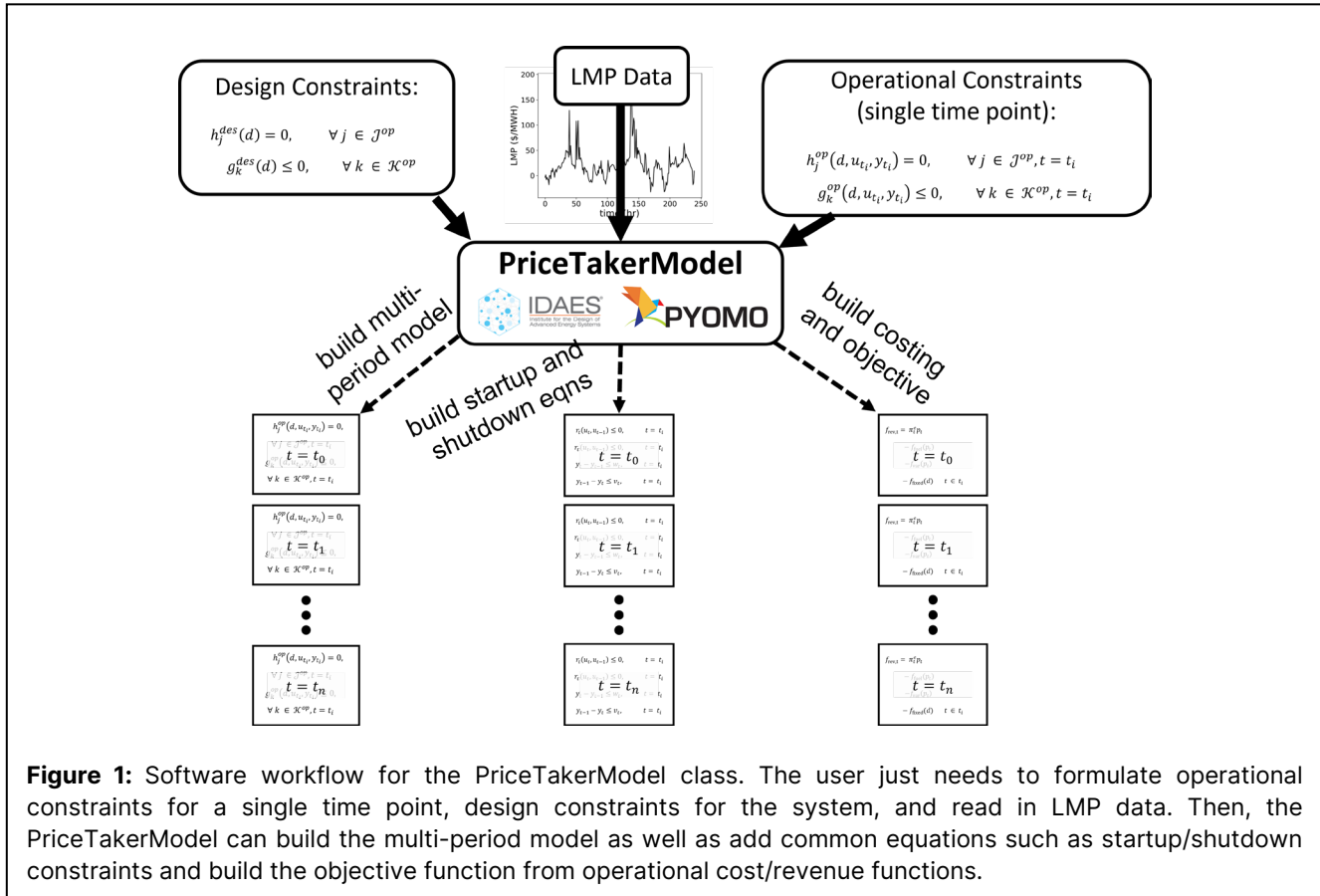
The price taker assumption treats the time-varying energy prices obtained from historical data or forecasts as constant. Likewise, the self-schedule approach ignores the market-clearing processes. Although the price-taker approach is much more informed than leveled cost analysis, the ignorance of the market-clearing processes (e.g., scheduling, dispatching, price settling) can lead to misleading comparisons and incorrect conclusions.

### Beyond Price-taker

Recent multiscale optimization frameworks have emerged to refine these assumptions. Specifically, Gao et al. [7] couple detailed energy system (process) models and grid production cost models (PCMs) to simulate resource bidding (instead of self-schedule), market clearing (instead of price-taker), and tracking (which generalizes beyond steady-state). They demonstrate how the price-taker and self-schedule assumptions break down. This process requires solving multiple optimization problems sequentially. Recent work has highlighted the drawbacks of naïve price-taker approaches through methods using PCM models for solar power plants [8], analyzing IES value in coproduction systems [9], and using PCM for wind energy and storage systems [10]. More information on utilizing PCMs can be found in references [7-10], where methods beyond price taker are shown to perform better.

In the absence of detailed models or to include PCMs in an algebraic formulation, market surrogates can be used to include market behavior without solving market clearing problems [11]. To this end, Jalving et al. [12] propose incorporating market interactions directly into the energy system design and operation co-optimization problem. Jalving et al. [12] use a simple thermal generator case study to illustrate the limitations of the price taker assumption. Nevertheless, the price taker assumption is convenient and has significant advantages over LCOE. It is an excellent starting-point for optimization-based TEA.

### Need Faster Analysis Workflows



**Figure 1:** Software workflow for the PriceTakerModel class. The user just needs to formulate operational constraints for a single time point, design constraints for the system, and read in LMP data. Then, the PriceTakerModel can build the multi-period model as well as add common equations such as startup/shutdown constraints and build the objective function from operational cost/revenue functions.

Unfortunately, most of the optimization-based energy system TEA studies are one-off implementations. This work presents new generalized capabilities in the IDAES®-PSE platform [13] to simplify and standardize optimizing energy systems and market interactions. Specifically, a PriceTakerModel class, which utilizes the Multi-PeriodModel class in IDAES-PSE, facilitates rapid price-taker and self-schedule analyses using either rigorous equation-oriented (EO) process models or surrogates trained for data. Also, the price taker model generally can be solved quickly to provide quick analysis. A generalized interface between process models and a PCM enables rigorous verification by modeling energy system and market interactions.

This work generalizes previous analyses in the IDAES-PSE ecosystem into a software interface for the multi-period price-taker model. Thus, the rest of this work is organized to describe: (i) the software interface using an example to demonstrate the workflow, and (ii) a case study using the software interface and example given in a self-name section.

## SOFTWARE INTERFACE

Typical multi-period price taker models derive complexity from their consideration of time-varying markets.

For instance, let us consider a simple scenario that requires a power plant to operate at a minimum load without shutting down.

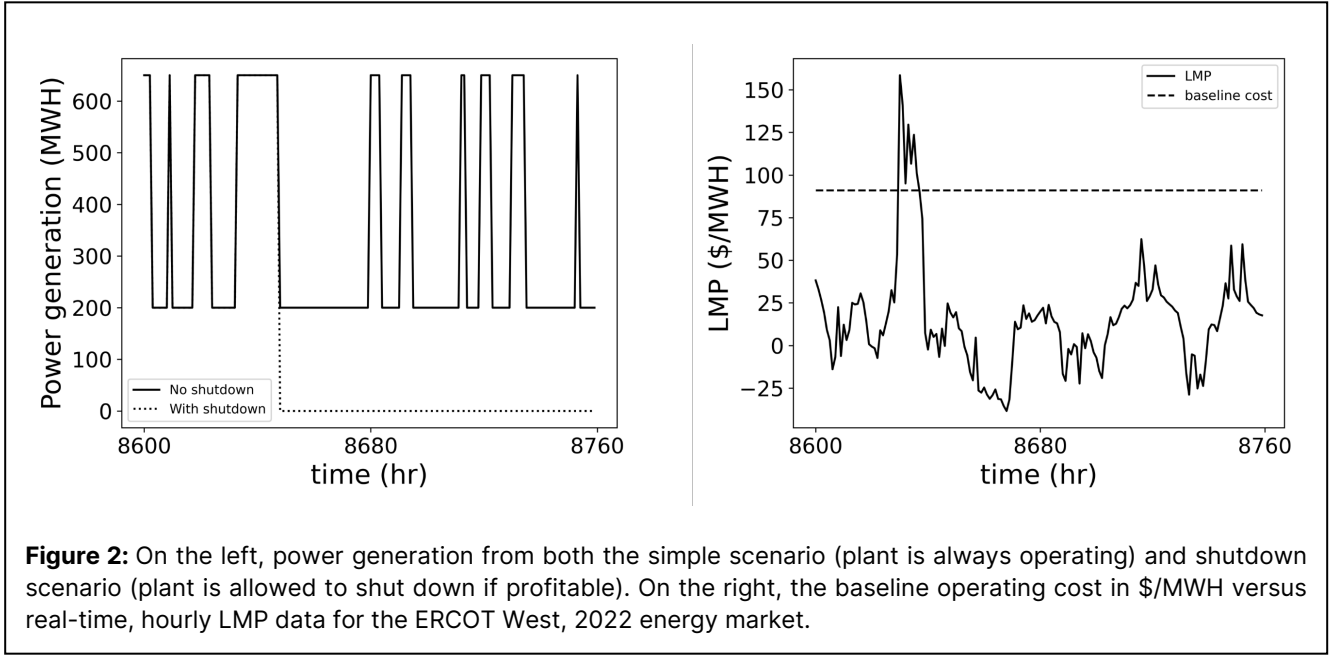
$$\max_{p_t} \text{Revenue} = \sum_{t \in \mathcal{T}} f_{rev,t} \quad (14)$$

$$\text{s. t. } f_{rev,t} = \pi_t^e p_t - f_{fuel}(p_t) - f_{var}(p_t) - f_{fixed}(d) \quad \forall t \in \mathcal{T} \quad (15)$$

$$p_t \in [p_{min}, p_{max}] \quad (16)$$

Here,  $p_t$  is the net power generation at time  $t$ . Total revenue is computed using the locational marginal price data (LMP) at time  $t$ ,  $\pi_t^e$ , with net power generation at that time less the fuel costs, variable costs, and fixed costs of the system, or  $f_{fuel}$ ,  $f_{var}$ , and  $f_{fixed}$ , respectively. Since the plant is always on, net power generation must be between  $p_{min}$  and  $p_{max}$ , or the minimum and maximum load of the plant, at all times.

In the IDAES®-PSE platform, new implementations for grid integration have been developed to facilitate multiperiod, price-taker models using the PriceTakerModel class, which makes use of the MultiPeriodModel class for time-varying models. The MultiPeriodModel class is used to extend steady state models to dynamic models by adding time-varying constraints with appropriate model superstructure (defined as Pyomo blocks of the steady-state model at each time interval). The



**Figure 2:** On the left, power generation from both the simple scenario (plant is always operating) and shutdown scenario (plant is allowed to shut down if profitable). On the right, the baseline operating cost in \$/MWH versus real-time, hourly LMP data for the ERCOT West, 2022 energy market.

PriceTakerModel class encapsulates either a steady state model, or a dynamic model as an instance of the MultiPeriodModel class, to facilitate price-taker analysis of an IES. The PriceTakerModel class also allows automated conversion of steady state models to time-varying multi-period models.

As shown in Figure 1, the modeler only needs to provide a function that defines the single time-point models to build the time-varying model. Also, the user can easily specify the LMP data by providing a .csv, .xls data file, or similar data format. As shown, methods for building the multi-period model, generating start-up and shutdown constraints, and building costing functions to aid in objective function development are all core functionalities of the new software interface. Not all features are required to be used for each model. For instance, the model shown in Eqs. (14) through (16) has no start-up or shutdown constraints, so the user can skip the development of those constraints and other parts of the interface as needed.

Also, more complex objective formulations such as evaluation net present value (NPV) with carbon tax, general corporate tax, and other common process design economic evaluators are readily available while building an objective function. Also, algebraic surrogate models can easily be implemented in this framework, as will be shown in the case study in the following section.

## CASE STUDY

A case study similar to that shown in Eqs. (14) through (16) was analyzed with and without start-up/shutdown constraints to demonstrate the utility of the new software interface. These values are based on a

solid oxide fuel cell (SOFC) power generation process, as described in Eslick et al. [14]. As such, the formulation with start-up and shutdown constraints is given below.

$$\max_{p_t, y_t, v_t, w_t} \text{Revenue} = \sum_{t \in \mathcal{T}} f_{\text{rev},t} - \sum_{t \in \mathcal{T}} (\theta_d w_t - \theta_u v_t) \quad (17)$$

$$\text{s. t. } f_{\text{rev},t} = \pi_t^e p_t - f_{\text{fuel}}(p_t) - f_{\text{var}}(p_t) - f_{\text{fixed}}(d) \quad \forall t \in \mathcal{T} \quad (18)$$

$$f_{\text{fuel}}(p_t) + f_{\text{var}}(p_t) = 23.29 p_t + 49.22, \quad \forall t \in \mathcal{T} \quad (19)$$

$$f_{\text{fixed}}(d) = 13500, \quad \forall t \in \mathcal{T} \quad (20)$$

$$y_{t-1} - y_t \leq w_t, \quad \forall t \in \mathcal{T} \quad (21)$$

$$y_t - y_{t-1} \leq v_t, \quad \forall t \in \mathcal{T} \quad (22)$$

$$\sum_{j=t-\tau_d+1}^t w_j \leq 1 - y_t, \quad \forall (t > \tau_d) \in \mathcal{T} \quad (23)$$

$$\sum_{j=t-\tau_u+1}^t v_j \leq y_t, \quad \forall (t > \tau_u) \in \mathcal{T} \quad (24)$$

$$y_t \in \{0, 1\} \quad (25)$$

$$w_t \in \{0, 1\} \quad (26)$$

$$v_t \in \{0, 1\} \quad (27)$$

$$p_t \in [200, 650] \quad (28)$$

The objective function has an additional term to account for the cost of shutting down and starting up the unit,  $\theta_d$  and  $\theta_u$ , respectively. Here, fuel and variable costs are represented linearly by a surrogate model, as shown in Eq. (19). Whether the process is on or off is dictated by binary variable  $y_t$ . Binary variables  $v_t$  and  $w_t$  represent start-up and shutdown indicators, respectively. Subsequently, start-up (Eq. (21)) and shutdown (Eq. (22)) constraints have been added. Similarly, minimum uptime and downtime variables,  $\tau_u$  and  $\tau_d$ , have been added, as well as minimum uptime (Eq. (24)), and minimum downtime



constraints (Eq. (23)) [15]. All cost values are in U.S. dollars per hour. The surrogate model for fuel and variable cost to run the SOFC unit are the truncated linear version of those given in Elick et al. [14]. LMP values,  $\pi_t^e$ , for an entire year were taken from the ERCOT West, 2022 hourly price data. Thus, the dimension of time set  $\mathcal{T}$  is 8760 for an entire year of hourly data.

Since the surrogate for cost is linear, the problem is a mixed-integer, linear programming formulation. Subsequently, these problems were solved using GUROBI [16]. Two scenarios were analyzed in the ERCOT West, 2022 market: (i) a simple scenario where the SOFC unit must remain on for all times (Eqs. (17) through (20) and Eq. (28)), and (ii) a scenario where the SOFC unit is allowed to shut down if it is more profitable (Eqs. (17) through (28)). Each MIP instance solves in less than 10 seconds on a standard workstation.

Results indicated that the SOFC in both scenarios remains on for almost the entire operating year. However, near the end of the year, the second system takes advantage of low prices that close out the year. Figure 2 shows that both scenarios follow power generation trends until scenario two shuts the SOFC unit off around hour 8650. Initially, the cost is large due to the upfront shutdown cost, but the system's low and sometimes negative LMP late in the year (also shown in Figure 2) results in a higher negative revenue stream than the shutdown cost. It should be noted that since the system never turned back on, the second scenario never withstood the large startup cost. The total profit difference due to this shutdown is about \$130,000. Still, the startup cost for this scenario was \$162,015, indicating that if the system must be turned on again after the start of the next year, it may not have been economically feasible if the time horizon did not end. This may indicate a desire to ensure the starting state of the system must match the ending state of the system (creating a continuous loop of yearly operation) with an additional integer constraint:

$$y_1 = y_{8760} \quad (29)$$

To verify this hypothesis, the system was optimized again with Eq. (29) included. The solutions were identical except that system that included Eq. (29) indeed did not turn off because the states need to match at either end of operation. Other analyses are possible, comparing various LMP networks to see which markets are profitable or not for certain technologies. For instance, ERCOT West 2022 indicates a significant profit for operating whereas other regions are less profitable.

## CONCLUSIONS

In this work, we presented a software interface developed in the IDAES®-PSE platform to facilitate implementation of the multi-period, price-taker model for

integrated energy systems. The interface utilizes Pyomo and existing IDAES infrastructure to streamline the creation of time-varying functions. Notably, the interface utilizes a single time point model to build a collection of blocks that elicit time-varying functionality. If the user desires a comparison between methods that are not time-varying, for instance LCOE analysis, the model for this single time point is sufficient for the operating constraints of the system.

Also, rapid comparison of different modeling paradigms is facilitated by the interface. For instance, utilizing other tools in the IDAES platform in conjunction with the interface presented in this work can lead to conclusions beyond power systems, such as multi-systems analysis including water treatment systems [17, 18], co-optimization of electricity and H<sub>2</sub> coproduction, as well as optimizing flexible carbon capture systems.

We showed how the inclusion of start-up/shutdown constraints had little to no impact on the adoption of SOFC in the ERCOT West 2022 market. This is a proof-of-concept comparison but shows modelers can utilize this framework to understand the required model fidelity to adequately capture the behavior of the system. In the future, we plan to integrate more user features that allow reformulations or relaxations of revenue terms and more formats for revenue inclusion in the objective function. We also plan to extend these tools to support beyond price taker approaches, such as surrogate models, to predict how price distributions change as a function of energy system design and operating decisions. This code will be released in the following year within the IDAES®-PSE platform.

## ACKNOWLEDGEMENTS

This work was conducted as part of the Design Integration and Synthesis Platform to Advance Tightly Coupled Hybrid Energy Systems (DISPATCHES) project through the Grid Modernization Lab Consortium with funding from the U.S. Department of Energy's Office of Fossil Energy and Carbon Management, Office of Nuclear Energy, and Hydrogen and Fuel Cell Technologies Office. Additional work was conducted as part of the Institute for the Design of Advanced Energy Systems (IDAES) with support from the U.S. Department of Energy's Office of Fossil Energy and Carbon Management (FECM) through the Simulation-based Engineering Research Program.

**Disclaimer:** This project was funded by the Department of Energy, National Energy Technology Laboratory an agency of the United States Government, through a support contract. Neither the United States Government nor any agency thereof, nor any of its employees, nor the support contractor, nor any of their employees, makes any warranty, expressor implied, or assumes any legal liability or responsibility for the accuracy, completeness,

or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## REFERENCES

- Dowling, A., Zheng, T., and Zavala, V.M. "Economic assessment of concentrated solar power technologies: a review." *Renewable and Sustainable Energy Reviews* 72: 1019-1032 (2017)
- Dowling, A., Kumar, R., and Zavala, V.M. "A multi-scale optimization framework for electricity market participation." *Applied Energy* 190: 147-164 (2017)
- Gooty, R.T., Ghouse, J., Le, Q.M., Thitakamol, B., Rezaei, S., Obiang, D., Gupta, R., Zhou, J., Bhattacharyya, D., and Miller, D.C. "Incorporation of market signals for the optimal design of post combustion carbon capture systems." *Applied Energy* 337: 120880 (2023)
- Sioshani, R., Denholm, P., Jenkin, T., and Weiss, J. "Estimating the value of electricity storage in PJM: arbitrage and some welfare effects." *Energy Economics* 31: 269-277 (2009)
- Walawalkar, R., Apt, J., and Mancini, R. "Economics of electric energy storage for energy arbitrage and regulation in New York." *Energy Policy* 35: 2558-2568 (2007)
- Flares, R.L., Meyers, J.P., and Webber, M.E. "A dynamic model-based estimate of the value of a vanadium redox flow battery for frequency regulation in Texas." *Applied Energy* 113: 189-298 (2014)
- Gao, X., Knueven, B., Sirola, J.D., Miller, D.C., and Dowling, A.W. "Multiscale simulation of integrated energy system and electricity market interactions." *Applied Energy* 316: 119017 (2022)
- Martinek, J., Jorgenson, J., Mehos, M., and Denholm, P. "A comparison of price-taker and production cost models for determining system value, revenue, and scheduling of concentrating solar power plants." *Applied energy* 231 (2018): 854-865.
- Frew, B., Levie, D., Richards, J., Desai, J., and Ruth, M. "Analysis of multi-output hybrid energy systems interacting with the grid: Application of improved price-taker and price-maker approaches to nuclear-hydrogen systems." *Applied Energy* 329 (2023): 120184.
- Sousa, J.A.M., Teixeira, F. and Faias, S. "Impact of a price-maker pumped storage hydro unit on the integration of wind energy in power systems." *Energy* 69 (2014): 3-11.
- Chen, X., Tumbalam-Gooty, R., Guittet, D., Knueven, B., Sirola, J.D., and Dowling, A.D. "Conceptual design of integrated energy systems with market interaction surrogate models" *Foundations of Computer Aided Process Design (FOCAPD 2024)*, Breckenridge, Colorado, USA, July 14-18, 2024.
- Jalving, J., Ghouse, J., Cortes, N., Gao, X., Knueven, B., Agi, D., Martin, S., Chen, X., Guittet, D., Gooty, R.T., Bianchi, L., Beattie, K., Gunter, D., Sirola, J.D., Miller, D.C., and Dowling, A.W. "Beyond price taker: Conceptual design and optimization of integrated energy systems using machine learning market surrogates." *Applied Energy* 351: 121767 (2023)
- Lee, A., Ghouse, J.H., Eslick, J.C., Laird, C.D., Sirola, J.D., Zamarripa, M.A., Gunter, D., Shinn, J.H., Dowling, A.W., Bhattacharyya, D., Biegler, L.T., Burgard, A.P., and Miller, D.C. "The IDAES process modeling framework and model library – Flexibility for process simulation and optimization." *Journal of Advanced Manufacturing Processes* 3:3 (2021)
- Eslick, J., Noring, A., Susarla, N., Okoli, C., Allan, D., Wang, M., Ma, J., Zamarripa-Perez, M.A., Iyengar, A., and Burgard, A. "Technoeconomic Evaluation of Solid Oxide Fuel Cell Hydrogen-Electricity Co-Generation Concepts." United States: N. p., 2023.
- Rajan, D. and Takriti, S. "Minimum Up/Down Polytopes of the Unit Commitment Problem with Start-Up Costs." Tech. Rep. W0506-050, IBM Research Division, Yorktown Heights, NY (United States) (2005).
- Gurobi Optimization, LLC. "Gurobi Optimizer Reference Manual", <https://www.gurobi.com>, 2023
- Drouven, M.G., Caldéron, A.J., Zamarripa, M.A. and Beattie, K. "PARETO: an open-source produced water optimization framework." *Optimization and Engineering* 24: 2229-2249 (2023)
- Atia, A.A., Allen, J., Young, E., Knueven, B., Bartholomew, T.V. "Cost optimization of low-salt-rejection reverse osmosis." *Desalination* 551 (2023)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Design and Optimization of Processes for Recovering Rare Earth Elements from End-of-Life Hard Disk Drives

Chris Laliwala, Ana I. Torres\*

Carnegie Mellon University, Department of Chemical Engineering, Pittsburgh, PA, USA.

\* Corresponding Author: [aitorres@cmu.edu](mailto:aitorres@cmu.edu).

---

## ABSTRACT

As the United States continues efforts to decarbonize the power and transportation sectors, significant challenges associated with the reliance of clean energy technologies on rare earth elements (REEs) will have to be overcome. One potential approach for increasing the supply of these elements is to extract REEs from end-of-life (EOL) hard disk drives (HDDs). HDDs contain neodymium and praseodymium, which are among the most important REEs for the clean energy transition, as they are crucial to producing the permanent magnets needed for wind turbines and electric vehicles. Here, we propose a superstructure-based approach to find the optimal pathway for recovering REEs from EOL HDDs. The superstructure was optimized by maximizing the net present value (NPV) over 15 years. Projected prices for commercial rare earth oxides and the projected amount of EOL HDDs in the U.S. were estimated and used in the model. These projections were used to establish the base case optimal result, assuming that the plant recycles 60% of personal computers EOL HDDs in the U.S. each year. The model was then expanded to consider the recycling of EOL HDDs generated before the beginning of plant production. Next, a sensitivity analysis was conducted to evaluate the impact of different parameters on the venture's profitability and the optimal processing pathway. Combined, these results offer both valuable insights into the economic viability of REE recycling extraction and a method for performing similar analyses in the future.

---

**Keywords:** Recycling, Rare Earth Elements, Process Design and Optimization.

## INTRODUCTION

The climate crisis presents one of the most significant challenges facing humanity in the twenty-first century, as the Earth has already exceeded pre-industrial temperatures by 1.1°C [1]. It is imperative to take immediate action to prevent further temperature increases, with a preference for staying within 1.5°C above pre-industrial levels [1]. This requires giving priority to the decarbonization of energy and our economy, including a transition to clean energy technologies such as solar panels, wind turbines, and electric vehicles to reduce emissions in the power and transportation sectors. Notably, in the United States, substantial legislative efforts, such as the "Bipartisan Infrastructure Law," entail an unprecedented investment of more than \$430 billion by 2031, aimed at decarbonizing energy among other initiatives [2].

Electrification initiatives are heavily dependent on

rare earth elements (REEs). Presently, the United States mines approximately 15% of the world's REEs but exports all raw materials for separation and refinement, primarily to China, which is responsible for mining around 60% of the world's REEs and maintains near-monopoly control over the entire supply chain [3]. This dependency represents a risk to domestic decarbonization efforts, primarily in terms of establishing domestic supply chains to produce clean energy technologies. Therefore, it is vital for the United States to strengthen its domestic rare earth supply chain as it embarks on the journey toward decarbonization.

The objective of this work is to create an optimization-based framework for designing the most efficient processing pathway for recovering rare earth elements from end-of-life hard disk drives (EOL HDDs). HDDs have received increasing attention in the U.S. as a potential source for REEs as evidenced by the work of agencies and

national laboratories (see for example NREL, EPA, and CMI webpages) [4-6]. The paper is organized as follows; first, an estimation was made regarding the projected quantity of REEs available for recycling from EOL HDDs over a period ranging from 2014 through 2038 [7-13], along with a projection of rare earth oxide (REO) prices over the plant's operational lifetime [14]. Next, a superstructure was formulated containing all potential processing pathways for recovering rare earth elements as rare earth oxides from the incoming HDD-based feedstock. Finally, optimization techniques were employed to find the processing pathway that maximizes the net present value over a 15-year period.

## MODEL ASSUMPTIONS

Rare earth elements (REEs), which encompass the lanthanide elements, scandium, and yttrium, are categorized into two groups based on their atomic weight: light rare earths (LREs) and heavy rare earths (HREs). LREs comprise Ce, La, Pr, Nd, and Sm, while HREs encompass Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb, and Lt. These elements serve various purposes, with magnets (Nd, Pr, Dy, Sm) constituting 29% of their usage, catalysts (La, Ce) contributing to 20%, and polishing agents (La, Ce) making up 13% across the top three global end-use sectors [15]. The varying abundance of individual REEs within deposits results in some elements being in surplus globally while others face shortages. Consequently, their prices exhibit significant discrepancies, with Nd/Pr being almost 100 times more valuable than La and Ce [15, 16]. HDDs contain relatively large amounts of Nd/Pr.

Two scenarios were examined in this paper. For the first scenario, only EOL HDDs generated during the plant's operational phase were considered. In this scenario, we assumed that 60% of all available EOL HDDs in the U.S. each year would be recycled and available to the plant [8]. In the second scenario, the model was expanded to consider the EOL HDDs generated in the U.S. prior to the start of production. In this scenario, it was assumed that 25% of all EOL HDDs generated in the U.S. over the period ranging from 2014 to 2024 would be recycled by the plant in addition to recycling 60% of all EOL HDDs available in the U.S. each year. For both scenarios, it was assumed that the plant's construction would begin in 2024 and be completed by 2025, after which production would start and continue for 14 years through 2038.

Several sources were used to estimate and quantify the amount of feedstock available for recycling as detailed in the subsequent section [7-13]. To estimate the price of the neodymium oxide, the rare earth oxide (REO) product of this process, over the operational lifetime of the plant, projections were taken from [14] and extrapolated through 2038.

We assumed HDDs to have a lifetime of 8 years [8].

The composition by wt. % of the average HDD rare earth permanent magnets (REPM) is 30 wt. % Nd/Pr, 66 wt. % Fe, and 4 wt. % other (Tb, Dy, Fe, B, Co, or Al) [7]. However, in this analysis, the 'other' category was ignored, and the composition of the REPM was assumed to be 30 wt. % Nd/Pr and 70 wt. % Fe. Note that following currently available data, Nd and Pr were considered together; in what follows Nd should be read as Nd/Pr.

## QUANTIFICATION OF FEEDSTOCK AND PRODUCTS

To estimate the amount of REPM available for recycling, the number of EOL HDDs in the U.S. was estimated based on historic and projected sales of laptop and desktop computers in the U.S. and other factors such as HDD type. Although the plant is to run from 2024 through 2038, we wanted to account for all the previously generated EOL HDDs, which we hypothesized some percent of which would still be available for recycling. Thus, the sales of desktops and laptops in the U.S. were estimated over a period of 2006 through 2030.

Sales data for laptops and desktops in the U.S. from 2018 through 2022 was taken from [9, 13]. Additionally, these sources contained projected sales for notebooks and desktops from 2023 through 2028. To estimate sales through 2030, a linear model was fit to the projected sales from 2023 through 2038 and extrapolated through 2038. Estimating sales for years before 2018 proved more difficult due to a lack of data. Thus, the following methodology was employed: First, data for the worldwide sales of personal computers (PCs) and the worldwide sales of laptops from 2006 through 2022 was taken from [10, 11]. Then, the worldwide sales for desktops were calculated by subtracting the sales of notebooks from the sales of PCs. Next, the average percentage of global sales the U.S. accounts for from 2018 through 2022 was calculated. This percentage was assumed to hold from 2006 through 2018. Finally, the annual sales of laptops and desktops in the U.S. from 2006 through 2018 were calculated by multiplying the global sales of desktops and laptops by this percentage.

After estimating the sales of laptops and desktops in the U.S., the next step was to evaluate how many contained HDDs. To accomplish this, we first took data for the global sales of HDDs and SSDs from 2015 through 2022 from [12]. Next, a linear model was fit to the sales data and used to extrapolate sales back to 2006 and up to 2030. In the case of SSDs, this linear model began predicting negative sales in 2012. Thus, the sales of SSDs were assumed to be zero from 2006 through 2012. Using this sales data, the ratio of HDDs to SSDs sold each year was estimated. It was then assumed that this ratio was the same in the U.S. as globally. Additionally, it was assumed to hold for laptops and desktops sold. Thus, the number of HDD-

containing laptops and desktops in the U.S. was estimated by multiplying the sales of laptops and desktops by the percentage of total HDD and SSD sales that HDDs make up. These estimates can be seen in Figure 1.

Finally, using this information, the amount of REPM available for recycling was estimated. HDDs have two sizes: 2.5" and 3.5". We assumed all desktops used 3.5" HDDs and all laptops used 2.5" HDDs. It was also assumed that each 2.5" HDD contains 2.5g of REPM [7]. To estimate the amount of REPM per 3.5" HDD in grams, a linear model taken from [7] was used and shown in Equation 1.

$$17.87 - 0.35t \quad (t = 0 @ 1990) \quad (1)$$

Further, it was assumed that the amount of REPM contained within a 3.5" HDD would stop decreasing once it equaled the mass of REPM contained within a 2.5" HDD (2.5g). The annual amount of REPM and Nd available for recycling from EOL HDDs in the U.S. is in Figure 2.

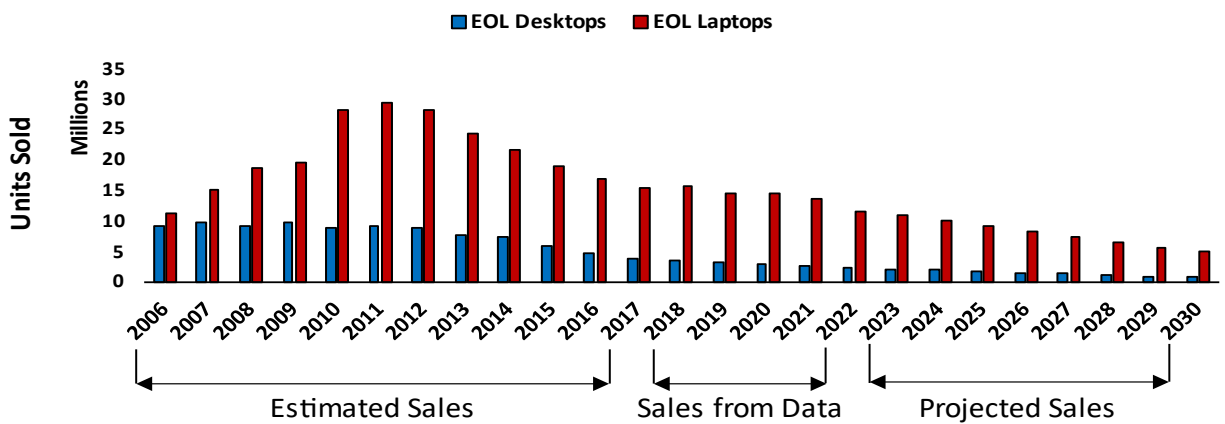
## PROCESSES FOR RECOVERING REES FROM EOL HDDS

Rare earth recycling comprises four primary processing stages: disassembly, demagnetization, leaching, and extraction. Each stage contains several potential

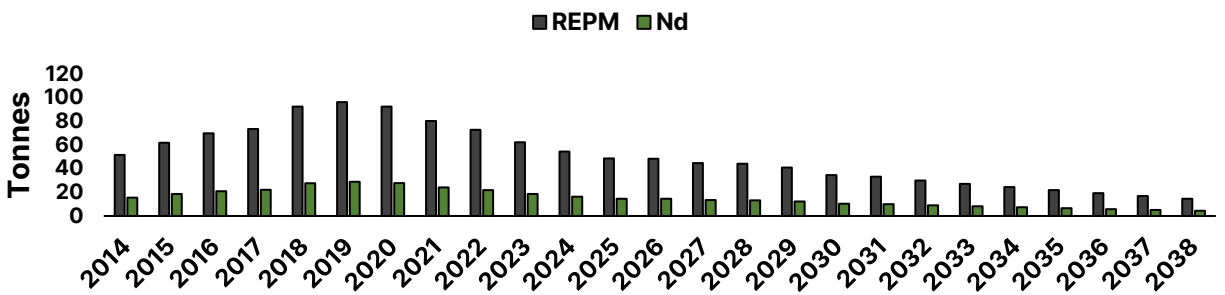
processes, represented as nodes in Figure 3. The naming convention for the nodes is  $\langle \text{stage number}, \text{node number} \rangle$ , where 1 represents the disassembly stage, 2 the demagnetization stage, 3 the leaching stage, and 4 the extraction stage. The superstructure containing all potential processing pathways can be seen in Figure 3.

For the disassembly stage, three processes were considered: Manual Disassembly  $\langle 1,1 \rangle$ , where REPM is removed from EOL HDDs manually [17]; Automatic Disassembly  $\langle 1,2 \rangle$ , which automates the process with robots [18]; and Shredding  $\langle 1,3 \rangle$ , which involves first heating the HDDs to high temperatures to demagnetize them, after which they are shredded using an industrial shredder. The final node is blank and represents skipping the disassembly stage.

In the demagnetization stage, three processes were explored: Hydrogen Decrepitation  $\langle 2,1 \rangle$ , which involves the reaction of the REPM with hydrogen gas at 170°C for 3 hours to produce a demagnetized, friable material [19]; Heating  $\langle 2,2 \rangle$ , which involves heating the magnets to 350°C for 30 minutes followed by milling [20]; and Extreme Heating  $\langle 2,3 \rangle$ , which involves heating the magnets to 950° C for 15 hours to demagnetize and oxidize them [21] followed by milling. This extreme heating results in

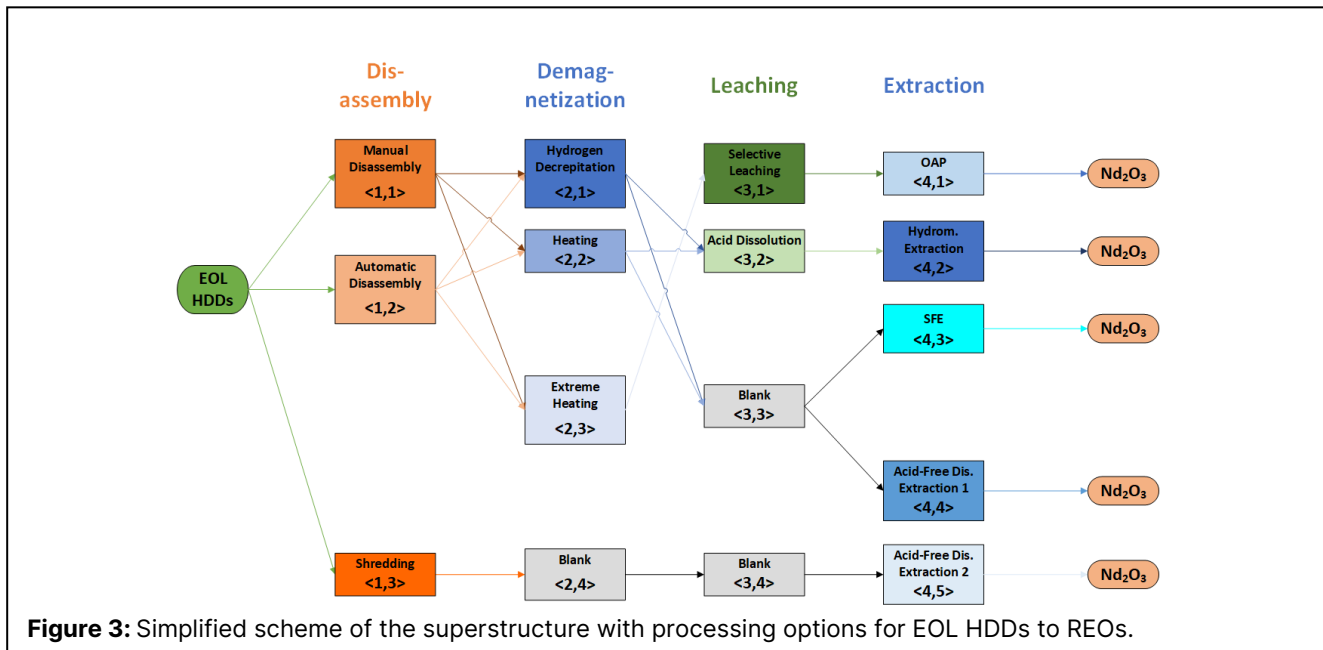


**Figure 1:** Annual estimated sales, sales from data, and projected sales of desktops and laptops containing HDDs in the U.S. from 2006 through 2030.



**Figure 2:** REPM and Nd available for recycling from HDDs in the U.S. from 2014 through 2038.





iron precipitation during the subsequent selective leaching process. The final node  $\langle 2,4 \rangle$  is blank and represents skipping the demagnetization stage.

For the leaching stage, two processes were considered: Selective Leaching  $\langle 3,1 \rangle$  with hydrochloric acid [21], which follows the extreme heating process, and Acid Dissolution  $\langle 3,2 \rangle$  using sulfuric acid [22]. In selective leaching, iron precipitates out of the solution as iron (III) hydroxide. The final two nodes,  $\langle 3,3 \rangle$  and  $\langle 3,4 \rangle$  are blank and represent skipping the leaching stage.

Five processes were examined in the extraction stage. The first process, Oxalic Acid Precipitation (OAP)  $\langle 4,1 \rangle$ , utilizes oxalic acid to directly precipitate neodymium out of solution [23]. The next process, Hydrometallurgical Extraction  $\langle 4,2 \rangle$ , precipitates neodymium as an Nd-sodium double salt to separate it from iron. Subsequent steps involve reactions with oxalic acid and calcination to produce pure  $\text{Nd}_2\text{O}_3$ , with the iron-containing leachate being treated with ammonium sulfate to precipitate it out of solution as iron jarosite [22]. The third potential process, Supercritical Fluid Extraction (SFE)  $\langle 4,3 \rangle$ , utilizes supercritical  $\text{CO}_2$  to separate Nd from iron [24, 25].

In the final two processes, Acid-Free Dissolution Extraction 1  $\langle 4,4 \rangle$  and 2  $\langle 4,5 \rangle$ , REPM powder is dissolved in an aqueous copper(II) nitrate solution [26]. Next, oxalic acid is added to precipitate neodymium out of solution as neodymium oxalate and iron out of solution as iron-ammonium oxalate. The neodymium oxalate precipitate is then filtered from the soluble iron-ammonium oxalate. Finally, the neodymium oxalate is calcined to produce  $\text{Nd}_2\text{O}_3$  [26]. These final two processes are identical, except in their final product yield ( $\text{Nd}_2\text{O}_3$ ). The yield for this process differs depending on whether the incoming feedstock is shredded HDDs or pure REPM. The yield for a pure REPM feedstock is 98.5 wt. %, while the yield for a shredded HDD

feedstock is 73 wt. %. Due to space constraints, details on these processes are not shown in the Figure.

## OPTIMIZATION PROBLEM FORMULATION

Equations (2-8) were formulated to describe the superstructure's configuration. A binary variable  $y_{i,j}$  is introduced to model the selection of node  $\langle i,j \rangle$ . If the node is selected, the binary will equal 1; if not, it will equal 0. Equation (2) enforces that only one node per stage is selected. Equation (3) enforces that if  $\langle 1,1 \rangle$  or  $\langle 1,2 \rangle$  is selected, then either  $\langle 2,1 \rangle$ ,  $\langle 2,2 \rangle$  or  $\langle 2,3 \rangle$  must be chosen. Equation (4) enforces that if  $\langle 1,3 \rangle$  is selected,  $\langle 2,4 \rangle$ ,  $\langle 3,4 \rangle$  and  $\langle 4,5 \rangle$  must all be selected. Equation (5) ensures that if  $\langle 2,1 \rangle$  or  $\langle 2,2 \rangle$  are selected, then either  $\langle 3,2 \rangle$  or  $\langle 3,3 \rangle$  must be selected. Equation (6) ensures that if  $\langle 2,3 \rangle$  is selected, then only  $\langle 3,1 \rangle$  or  $\langle 4,1 \rangle$  can be chosen. Equation (7) ensures that if  $\langle 3,2 \rangle$  is selected then  $\langle 4,2 \rangle$  is also chosen. Equation (8) enforces that either  $\langle 4,3 \rangle$  or  $\langle 4,4 \rangle$  is chosen if  $\langle 3,3 \rangle$  is selected. Finally, big-M constraints were added to relate the flow through a unit to its selection, as shown in Equation (9). Here,  $F_{i,j,c,t}^{in}$  represents the flow of component  $c$  entering node  $\langle i,j \rangle$  in year  $t$ ,  $M^{i,j}$  represents the maximum inlet flow rate for node  $\langle i,j \rangle$ ,  $I$  is the set of all stages in the superstructure,  $J_i$  is the set of all nodes in stage  $i$ ,  $C$  is the set of tracked components, and  $T$  is the set of all years the plant is in operation. The mass balance for a generic node is illustrated in Figure 4.

$$\sum_{j \in J_i} y_{i,j} = 1 \quad \forall i \in I \quad (2)$$

$$\begin{aligned}
y_{1,1} + y_{1,2} &= y_{2,1} + y_{2,2} + y_{2,3} & (3) \\
y_{1,3} &= y_{2,4} = y_{3,4} = y_{4,5} & (4) \\
y_{2,1} + y_{2,2} &= y_{3,2} + y_{3,3} & (5) \\
y_{2,3} &= y_{3,1} = y_{4,1} & (6) \\
y_{3,2} &= y_{4,2} & (7) \\
y_{3,3} &= y_{4,3} + y_{4,4} & (8) \\
F_{i,j,c,t}^{in} &\leq y_{i,j} M^{i,j} \quad i \in I, j \in J_i, c \in C, t \in T & (9)
\end{aligned}$$

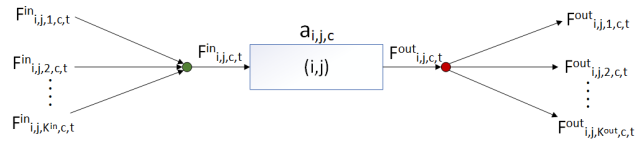
Components in the feedstock, intermediates, and products were tracked. The mass balance at each node includes three sequential steps: convergence of the inlet flows from different upstream nodes, processing, and separation of the outlet flows to other downstream nodes. The specific equations are defined in Equations (10-12).

$$\sum_{k=1}^{K^{in}} F_{i,j,k,c,t}^{in} = F_{i,j,c,t}^{in} \quad \forall i \in I, j \in J_i, c \in C, t \in T \quad (10)$$

$$F_{i,j,c,t}^{out} = a_{i,j,c} F_{i,j,c,t}^{in} \quad \forall i \in I, j \in J_i, c \in C, t \in T \quad (11)$$

$$F_{i,j,c,t}^{out} = \sum_{k=1}^{K^{out}} F_{i,j,k,c,t}^{out} \quad \forall i \in I, j \in J_i, c \in C, t \in T \quad (12)$$

Here,  $F_{i,j,c,t}^{in}$  is the flow of component  $c$  entering node  $\langle i, j \rangle$  in time  $t$ ,  $a_{i,j,c}$  is the yield of component  $c$  for node  $\langle i, j \rangle$ ,  $F_{i,j,c,t}^{out}$  is the flow of component  $c$  leaving node  $\langle i, j \rangle$  in time  $t$ , and  $F_{i,j,k,c,t}^{out}$  is the flow of component  $c$  leaving node  $\langle i, j \rangle$  at time  $t$ .

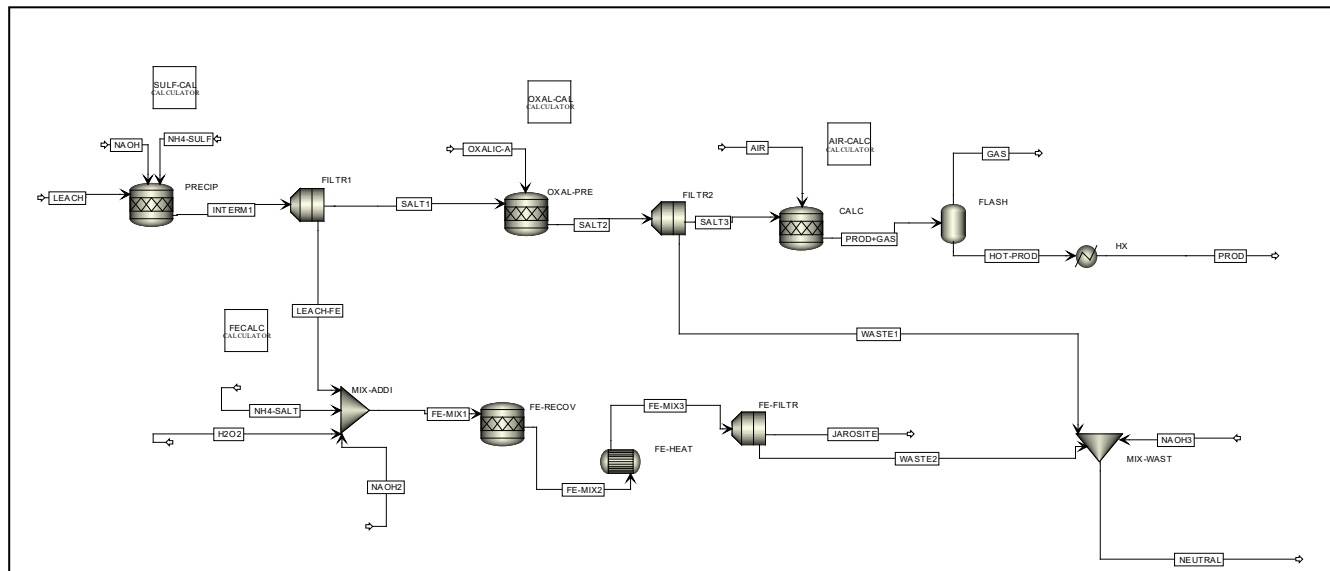


**Figure 4:** Mass balance framework for generic node  $\langle i, j \rangle \forall c \in C, t \in T$ .

## ASPEN FLOWSHEETS AND SIMULATIONS

The study and simulation of REE recycling production pathways is a newer and rapidly evolving area of research. Hence, in several cases, process flowsheets with techno-economic analysis were either not available in the literature or the data were of insufficient detail to effectively leverage for this study. Consequently, process flowsheets for several recycling pathways were designed and simulated in Aspen Plus. These flowsheets included the following unit operations: hydrogen decrepitation, heating/high-temperature heating, acid-free dissolution extraction, selective leaching, hydrometallurgical extraction, solvent extraction, and oxalic acid precipitation.

As an example, Figure 5 depicts the Aspen flowsheet designed for the hydrometallurgical extraction process. In the first step, the incoming leachate from the acid dissolution process is reacted with NaOH and ammonium sulfate at a pH of 1.5 to form Nd-sodium double salt precipitates. In the next step, the precipitate is separated from the iron-containing leachate and is reacted with oxalic acid to form neodymium oxalate. The neodymium oxalate is first filtered and then calcined at 750°C in the presence of oxygen to form Nd<sub>2</sub>O<sub>3</sub>. In the final step, the product is cooled to 100°C using cooling water. However, before the spent leachate can be disposed of, the iron must first be



**Figure 5:** Aspen simulation built for the hydrometallurgical process built using reaction yield data from [24].

precipitated out of solution as jarosite. This is accomplished by heating the leachate to 90°C for 6 hours and reacting it with ammonium sulfate in the presence of an oxidizing agent at a pH of 2 [22]. Finally, the spent leachate is neutralized with NaOH.

The individual operations in these flowsheets were modeled using reaction and separation data from literature. Capital (CAPEX) and operating costs (OPEX) were evaluated using the Aspen Process Economics Analyzer.

## OBJECTIVE FUNCTION

The goal of the optimization problem was to maximize the NPV over the plant's lifetime, as shown in Equations (13-14) [27]. Where  $CF_n$  is the cash flow in year  $n$ ,  $LT$  is the plant's lifetime in years, and  $IR$  is the interest rate.

$$NPV = \sum_{n=1}^{LT} \frac{CF_n}{(1 + IR)^n} \quad (13)$$

$$CF_n = Profit_n - (CAPEX_n + OPEX_n) \quad (14)$$

An interest rate of 10% was used as recommended by [28]. Tax and depreciation were ignored. It was assumed that all capital investments were made in year 1, and that the plant does not begin operation until year 2. Therefore, CAPEX was only considered for year 1, after which only profit and OPEX were considered. The methodology described by [27] was used for each node to calculate the CAPEX and OPEX. For OPEX, only the cost of labor was considered for the disassembly stage, and only the cost of raw materials and utilities were considered for the later stages.

Costing information that was available from the literature was adjusted for capacity using a piecewise-linearized approximation of the six-tenths rule [27]. When not available from the literature, Aspen Plus simulations and the Aspen Process Economic Analyzer were used to estimate CAPEX and OPEX costs for a range of flow rates. A linear regression model was then fit to obtain the CAPEX and OPEX vs incoming flow rates.

## RESULTS

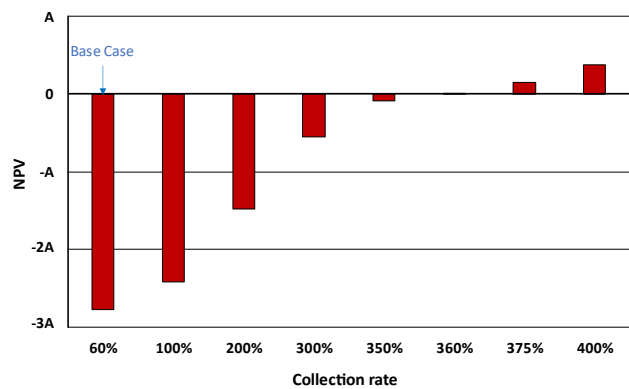
The optimization problem was coded in Pyomo and solved with CPLEX version 22.1.1.0 using the default options. The decision variables consisted of the operations included in the process and their sizing. The formulation for the MILP optimization problem is as follows:

$$\begin{aligned} \max \quad & z = NPV \\ \text{s.t.} \quad & Eq. (2 - 12) \\ & F_{i,j,c,t}^{in} \geq 0, y_{i,j} \in \{0,1\} \forall i \in I, j \in J_i, c \in C, t \in T \end{aligned}$$

The base scenario assumed that the plant recycles

60% of all available EOL HDDs in the U.S. annually [8]. The optimal pathway was found to consist of shredding, followed by acid-free dissolution extraction, and resulted in a negative NPV. The acid-free dissolution extraction processing step was found to be the most significant contributor to CAPEX, accounting for ~90% of the cost. Approximately 162.5 tonnes of  $Nd_2O_3$  were recovered.

Next, we conducted a sensitivity analysis on the percentage of EOL HDDs in the U.S. the plant recycles each year (collection rate, Fig. 6). We found that the plant remained unprofitable even at a 100% collection rate. The NPV breakeven point was found to occur at a collection rate of ~360%, and the optimal process never changed. This suggests that an insufficient number of HDDs are being produced for the plant to leverage economies of scale to be profitable.

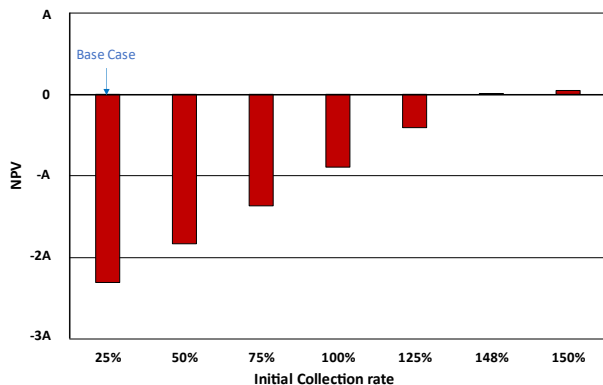


**Figure 6:** NPV for a varying collection rate. The base case is a collection rate of 60% and no recycling of EOL HDDs generated prior to plant production. The NPV break-even point was found to occur at ~360%. Numerical values are not reported to preserve confidentiality.

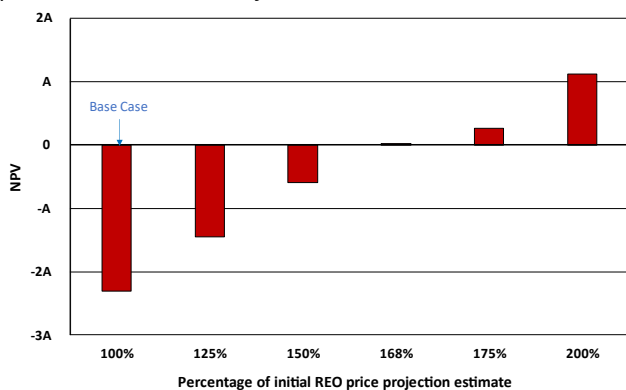
Given the results of our first sensitivity analysis, we hypothesized that the optimal process may prove to be profitable when the recycling of EOL HDDs stockpiled from years prior to the beginning of plant production was considered. Therefore, we expanded our model to recycle 25% of all EOL HDDs generated in the U.S. over the period ranging from 2006 to 2024 (initial collection rate) in addition to recycling 60% of all available EOL HDDs each year of operation. The optimal process was found to be the same as that of the previous scenario, however, the NPV was still found to be negative. We then conducted a sensitivity analysis on the initial collection rate and found the NPV breakeven point to occur at ~148% (Fig. 7). The optimal process never changed.

Finally, we conducted a sensitivity analysis on projected REO prices for the expanded model. To vary the projected REO prices, the new projection was assumed to be some percentage of the initial estimate. This percentage was then varied. The NPV breakeven point was found to occur at ~168% of the initial estimate (Fig. 8). Once

again, the optimal process never changed.



**Figure 7:** NPV for a varying initial collection rate. The base case has a collection rate of 60% and an initial collection rate of 25%. The NPV break-even point was found to occur at ~148%. Numerical values are not reported to preserve confidentiality.



**Figure 8:** NPV for varying percentages of the initial REO price projection estimate. The NPV break-even point was found to occur at ~168%. Numerical values are not reported to preserve confidentiality.

## CONCLUSION

Superstructure optimization was utilized to design a process for the recycling of rare earth elements from EOL HDDs, with the objective of maximizing the NPV. The study determined that the optimal process for both the base case and expanded model consists of shredding, followed by acid-free dissolution extraction. Our results show that the venture considering HDDs from PCs is likely not profitable as a stand-alone plant due to the total amount of REPM available for recycling from EOL HDDs being insufficient given the projected REO prices. However, we hypothesize that by expanding the plant to process multiple feedstocks, or by introducing multiple agents to handle the different processing steps, the optimal pathway may become profitable. Future work will investigate these options.

## ACKNOWLEDGMENTS

This effort was funded by the US Department of Energy’s Process Optimization and Modeling for Minerals Sustainability (PrOMMiS) Initiative, supported by the Office of Fossil Energy and Carbon Management’s Office of Resource Sustainability. Neither the United States Government nor any agency thereof, nor any of its employees, nor the support contractor, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## REFERENCES

- UNFCCC. <https://unfccc.int/documents/310475>
- The White House. <https://www.whitehouse.gov/briefing-room/statements-releases/2023/04/20/fact-sheet-president-biden-to-catalyze-global-climate-action-through-the-major-economies-forum-on-energy-and-climate/>
- Oberhaus D, Rare earths for America’s future. (2023)
- NREL. <https://www.nrel.gov/news/program/2021/in-a-circular-economy-hard-drives-could-have-multiple-lives-in-the-future.html>
- EPA. <https://www.epa.gov/vcs/using-standards-promote-reuse-rare-earth-materials>
- CMI. <https://www.ameslab.gov/news/it-s-all-part-of-the-grind-cmi-s-new-hard-drive-shredder-serves-up-plenty-of-material-for#:~:text=The%20Critical%20Materials%20Institute%20specializes,motors%2C%20and%20wind%20turbine%20generators>
- Sprecher B, Kleijn R, Kramer GJ. The case of computer hard disk drives. *Environ Sci Technol* 48, 16:9506-9513 (2014)
- Maani T, Mathur N, Singh S, Rong C, Sutherland JW, Potential for Nd and Dy recovery from end-of-life products to meet future electric vehicle demand in the U.S. *Procedia CIRP* 98:109-114 (2021)
- Statista. <https://www.statista.com/outlook/cmo/consumer-electronics/computing/desktop-pcs/united->

- [states?currency=USD&locale=en#revenue](#)
10. Statista. <https://www.statista.com/statistics/273495/global-shipments-of-personal-computers-since-2006/>
  11. Statista. <https://www.statista.com/statistics/203691/global-unit-shipments-of-notebooks/>
  12. Statista. <https://www.statista.com/statistics/285474/hdds-and-ssds-in-pcs-global-shipments-2012-2017/>
  13. Statista. <https://www.statista.com/outlook/cmo/consumer-electronics/computing/laptops/united-states>
  14. Hykawy J, Chudnovsky T. <https://www.stormcrow.ca/wp-content/uploads/2021/03/20210308-Stormcrow-UCore-Initiation-Final.pdf>.
  15. Gielen D, Lyons M. Critical materials for the energy transition: rare earth elements (2021)
  16. Smith BJ, Riddle ME, Earlam MR, Iloeje C, Diamond D. Rare earth permanent magnets: supply chain deep dive assessment (2022)
  17. Peeters JR, Bracquene E, Nelen D, Ueberschaar M, Van Acker K, Duflou JR. Forecasting the recycling potential based on waste analysis: a case study for recycling Nd-Fe-B magnets from hard disk drives. *J of Clean Prod* 175:96-108 (2018)
  18. Peiró LT, Girón AC, Durany XG. Examining the feasibility of the urban mining of hard disk drives. *J Clean Prod* 248:119-216 (2020)
  19. Walton A, Yi H, Rowson NA, Speight JD, Mann VSJ, Sheridan RS, Bradshaw A, Harris IR, Williams AJ. The use of hydrogen to separate and recycle neodymium-iron-boron-type magnets from electronic waste. *J Clean Prod* 104: 236-241 (2015)
  20. Klemettinen A, Žak A, Chojnacka I, Matuska S, Leśniewicz A, Wełna M, Adamski Z, Klemettinen L, Rycerz L. Leaching of rare earth elements from NdFeB magnets without mechanical pretreatment by sulfuric (H<sub>2</sub>SO<sub>4</sub>) and hydrochloric (HCl) acids. *Miner* 11: 1374 (2021)
  21. Vander Hoogerstraete T, Blanpain B, Van Gerven T, Binnemans K. From ndfeb magnets towards the rare-earth oxides: a recycling process consuming only oxalic acid. *RSC Adv* 4: 64099-64111 (2014)
  22. Lyman JW, Palmer GR. Recycling of rare earth and iron from NdFeB magnet scrap. *High Temp Mater and Process* 11: 175-187 (1993)
  23. Nawab A, Yang X, Honaker R. Parametric study and speciation analysis of rare earth precipitation using oxalic acid in a chloride solution system. *Miner Eng* 176: 107352 (2022)
  24. Zhang J, Anawati J, Yao Y, Azimi G. Aeriometallurgical extraction of rare earth elements from a NdFeB magnet utilizing supercritical fluids. *ACS Sustain Chem Eng* 6: 16713-16725 (2018)
  25. Azimi G, Sauber ME, Zhang J. Technoeconomic analysis of supercritical fluid extraction processes for recycling rare earth elements from neodymium iron boron magnets and fluorescent lamp phosphors. *J Clean Prod* 422: 138526 (2023)
  26. Chowdhury NA, Deng S, Jin H, Prodius D, Sutherland JW, Nlebedim IC. Sustainable recycling of rare-earth elements from NdFeB magnet swarf: techno-economic and environmental perspectives. *ACS Sustain Chem Eng* 0: 15915-15924 (2021)
  27. Seider WD, Lewin DR, Seader JD, Widagdo S, Gani R, Ng KM. Product and Process Design Principles Synthesis, Analysis and Evaluation. 4<sup>th</sup> Edition, John Wiley & Sons (2017) ISBN 978-1-119-28263-1
  28. Biegler LT, Grossman IE, Westerberg AW. Systematic Methods of Chemical Process Design. Prentice Hall PTR (1997).

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# NMPC for Mode-Switching Operation of Reversible Solid Oxide Cell Systems

Mingrui Li<sup>a</sup>, Douglas A. Allan<sup>ac</sup>, San Dinh<sup>e</sup>, Lorenz T. Biegler<sup>e</sup>, Debangsu Bhattacharyya<sup>d</sup>, Vibhav Dabadghao<sup>e</sup>, Nishant Giridhar<sup>d</sup>, Stephen E. Zitney<sup>b\*</sup>

<sup>a</sup> National Energy Technology Laboratory, Pittsburgh, PA 15236

<sup>b</sup> National Energy Technology Laboratory, Morgantown, WV 26507

<sup>c</sup> NETL Support Contractor, Pittsburgh, PA 15236

<sup>d</sup> West Virginia University, Morgantown, WV 26507

<sup>e</sup> Carnegie Mellon University, Pittsburgh, PA 15213

\* Corresponding Author: [steve.zitney@netl.doe.gov](mailto:steve.zitney@netl.doe.gov).

## ABSTRACT

Solid oxide cells (SOCs) are a promising dual-mode technology that generates hydrogen through high-temperature water electrolysis and generates power through a fuel cell reaction that consumes hydrogen. Reversible operation of SOC requires a transition between these two modes for hydrogen production setpoints as the demand and price of electricity fluctuate. Moreover, a well-functioning control system is important to avoid cell degradation during mode-switching operation. In this work, we apply nonlinear model predictive control (NMPC) to an SOC module and supporting equipment and compare NMPC performance to classical proportional integral (PI) control strategies, while ramping between the modes of hydrogen and power production. While both control methods provide similar performance in many metrics, NMPC significantly reduces cell thermal gradients and curvatures (mixed spatial-temporal partial derivatives) during mode switching. A dynamic process flowsheet of the reversible SOC system was developed in the open-source, equation-based IDAES modeling framework. Our IDAES dynamic simulation results show that NMPC can ramp the SOC system between hydrogen and power production targets within short mode-switching times. Moreover, NMPC can comply with operating limits in the SOC system more effectively than PI, and only NMPC can directly enforce user-specified limits for mixed spatial-temporal partial derivatives of temperature. This allows for management of the trade-off between operating efficiency and cell degradation, which is dependent on these temperature curvatures.

**Keywords:** Sustainability, Implementation, Energy & Environment, Process Optimization & Control, NMPC, Solid Oxide Cells, SOEC, SOFC

## INTRODUCTION

In recent years, a growing share of variable renewable energy generation and ambitious decarbonization targets have spurred a notable shift away from fossil fuels, with hydrogen poised to play a crucial role in this energy transition. By combining hydrogen production and electricity generation, integrated energy systems based on reversible SOC technology (rSOC) can be optimized to provide the operational flexibility required to meet the varying load demand of the modern grid. Hydrogen can function as a valuable energy storage medium and as a

versatile feedstock for other purposes. While most industrial hydrogen today is produced through steam methane reforming, which uses a fossil fuel feedstock, water electrolysis is a promising replacement, producing no direct greenhouse gas emissions when renewable energy is used.

For hydrogen production through water electrolysis, Nernst potential, the minimum potential difference at which electrolysis occurs, decreases with increasing reaction temperature. Because solid oxide electrolysis cells (SOECs) and solid oxide fuel cells (SOFCs) operate at 600°C to 1000°C, much higher temperatures than those

of other electrolysis technologies, they are good candidates for efficient water electrolysis. In power production mode, high temperature operation allows various fuels, including hydrogen, to rapidly undergo oxidation reaction on the fuel electrode. However, high temperature operation comes with significant drawbacks. Besides additional heat exchange equipment and good thermal insulation requirements, transitions between operating points must be carefully controlled to minimize power requirements and avoid thermal stress.

Many SOC systems are reversible and can operate as hydrogen fuel cells when grid demand becomes high. This flexible operation is necessary to stabilize the grid and operate profitably with intermittent renewable energy. Switching between SOEC and SOFC modes while considering both operating performance and equipment longevity can be challenging. Ferrari [1] coupled PI controllers with feed-forward approaches to reduce thermal gradients and limit the peak anode-cathode differential pressure for a tubular SOFC/GT hybrid system. However, modeling of the SOFC was based on "lumped volume" OD models, and the chemical reactions were assumed to be at equilibrium. Abbaker et al. [2] used a discrete-time adaptive terminal sliding-mode control strategy for voltage setpoint tracking for an SOFC. A pseudo-partial derivative technique was used to model the SOFC. Botta et al. [3] considered an rSOFC system and applied a PI controller to prevent dangerous operating conditions at the stack level. They analyzed individual SOFC and SOEC modes, as well as switching between these modes during reversible operation of the stack. Spivey and Edgar [4] developed a dynamic model for a tubular SOFC to capture the dynamics of critical thermal stress drivers and applied it to a MIMO predictive controller. Schotman [5] applied an output feedback adaptive NMPC approach to an rSOC system to control cell temperature and temperature gradients while maintaining desired level of power output. Xing et al. [6] designed a model predictive control (MPC) strategy based on a linear parameter varying model to improve short-term tracking performance and long-term operating efficiency for an rSOC plant. The process model for rSOC was a linear state-space model of the plant.

SOC systems are good candidates for nonlinear MPC (NMPC) since many manipulated variables (MVs) are highly interactive. As mentioned earlier, NMPC [7-8] uses a system model to predict system response to a sequence of MVs and optimizes it with respect to performance metrics. Because the controller can manipulate several degrees of freedom, it affords a quicker response than that of classical control.

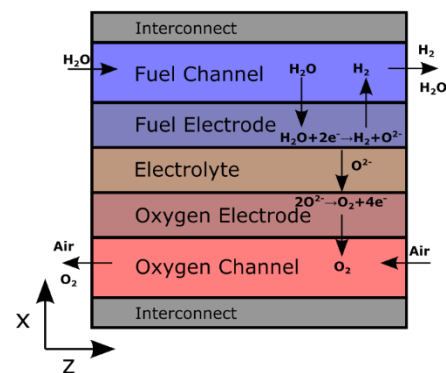
In this work, NMPC is applied to an SOC module and supporting equipment, with its performance compared to that of classical PI control strategies for switching between the modes of hydrogen and power production.

Switching from maximum hydrogen production to power generation and back to hydrogen production is demonstrated for the SOC system. Performance is judged based on the speed of production rate transition, total power usage, and whether unsafe thermal levels occur in the SOC.

## PROCESS MODELING

The SOC system flowsheet is built in the IDAES modeling framework [9]. The SOC model is publicly available in the IDAES GitHub repository [10], and a preliminary version is available in the literature [11]. There is ambiguity in denoting the cell "anode" and "cathode" in a reversible SOC, because the anode in fuel cell mode is the cathode in electrolysis mode and vice-versa. Hence, we refer to the electrode where hydrogen is produced or consumed as the fuel electrode and the electrode where oxygen is produced or consumed as the oxygen electrode. Model parameters were fit to data from the two-cell fuel electrode supported short stack shown in Figure 1 for the SOC system described in Figure 2.

The SOC model shown in Figure 1 is one dimensional in channels and (potentially) two dimensional in the fuel electrode, electrolyte, oxygen electrode, and interconnect. However, because the thickness of the positive electrode-electrolyte-negative electrode (PEN) assembly is only approximately one millimeter (with almost all of that in the fuel electrode), only one finite element is used in the x direction for the fuel electrode, while the electrolyte and oxygen electrode are approximated as thin-film resistors. The interconnect, which is 5 mm thick, also uses one finite element. The length of the cell in the z direction is 23.5 cm and ten finite elements are used in that direction for all subcomponents. The cell is run in a countercurrent configuration. The interconnect is represented by a periodic boundary condition, with one end attached to the top of the fuel channel and one to the bottom of the oxygen channel, to represent a cell in the center of a large stack. The cell model does not consider losses to the environment.



**Figure 1.** SOC configuration. Note that this diagram is not

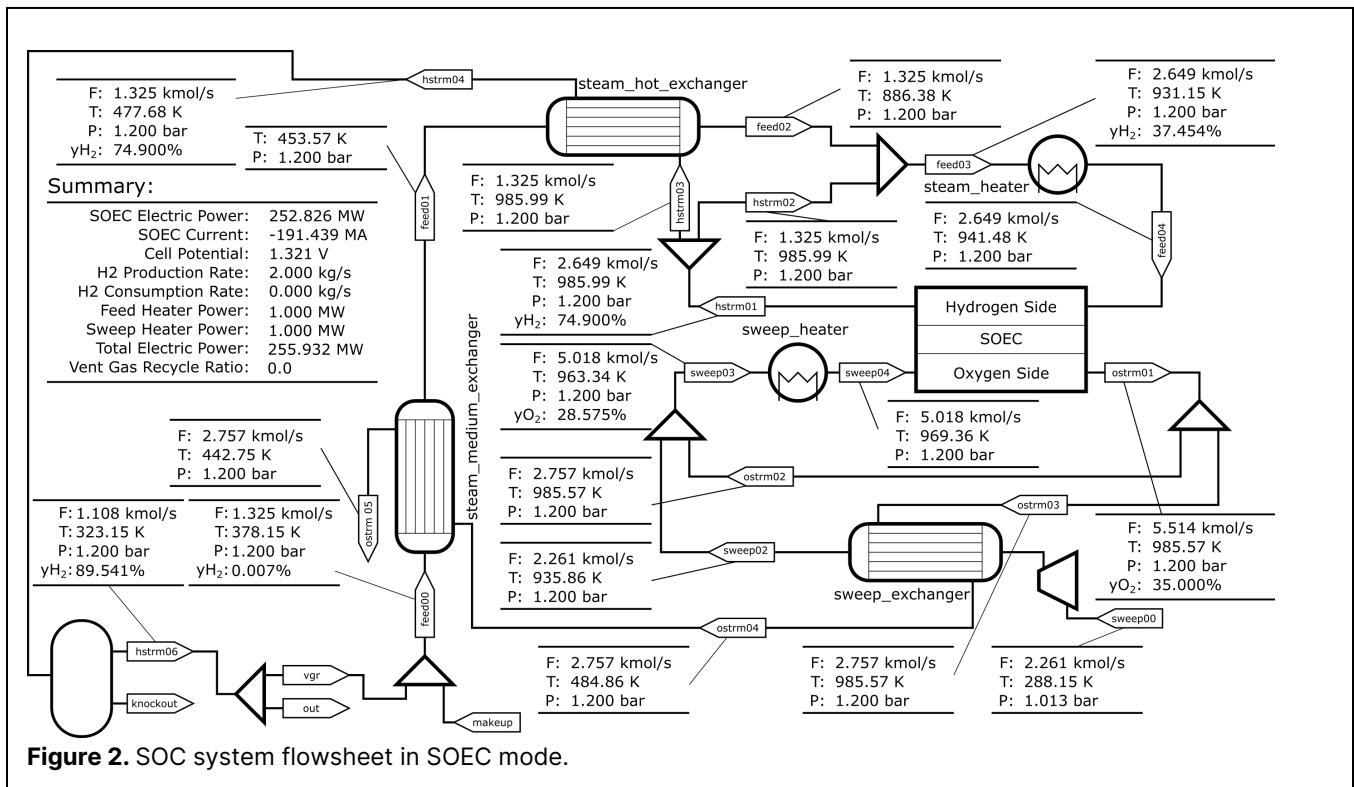


Figure 2. SOC system flowsheet in SOEC mode.

to scale—the cell is millimeters thick, and the electrolyte and oxygen electrode are hundreds of micrometers thick.

The SOC system flowsheet in Figure 2 consists of supporting equipment for the SOC module. Three cross-flow heat exchangers heat incoming air and water/hydrogen up to the cell’s operating temperature of approximately 975 K (in electrolysis mode). Two resistive trim heaters provide additional heat as well as control the temperature of the inlet streams to the stack. The oxygen stream is vented to the atmosphere, while the fuel stream goes to a condenser. Water is knocked out in a flash vessel. The hydrogen-rich gas exiting the flash is either mixed with the fresh fuel inlet to return to the SOC (in fuel-cell mode) or is taken out of the system for compression and further purification (in electrolysis mode). The heat exchangers and trim heaters use 1D models, whereas the other supporting equipment units use 0D models. Because of the short gas phase residence time in the system (about four seconds for the fuel side in electrolysis mode), gas phase holdups are turned off in all units, leaving thermal holdups as the primary source of dynamic behavior. For every discretized time step within the dynamic model, the SOC system flowsheet consists of 900 differential equations, 6955 algebraic equations, and 4311 inequalities.

In electrolysis mode, the hydrogen production rate is fixed at 2 kg/s, with the system aiming to minimize total power usage. The split fraction to the vent gas recycle stream is set to 0.0001 to avoid zero flows, and the cell

is sized to produce 2 kg/s with an average current density of  $-10,000 \text{ A/m}^2$ , allowing for a maximum current overshoot of 30%. Water conversion is constrained between 60% and 75% to optimize steam usage. In fuel cell mode, the cell operates at an average current density of  $4,000 \text{ A/m}^2$ , with an upper bound of  $5,200 \text{ A/m}^2$ . Hydrogen conversion has no specific bound due to recycling unconsumed hydrogen. Trim heaters provide 10 kW each to prevent issues with controller setpoints.

### CLASSICAL PROCESS CONTROL

The SOC control system must enforce a variety of process constraints. The fuel-side inlet stream must have at least 5% (mole-basis) of hydrogen to avoid oxidation of the electrode or interconnect. The fuel-side outlet must also have at least 25% hydrogen in fuel cell mode or water in electrolysis mode to avoid cell degradation from reactant starvation. The oxygen outlet stream can have at most 35% oxygen to avoid oxidation of process equipment. The cell voltage is limited between 0.7 and 1.4 V to avoid oxidation or reduction of the cell. The maximum temperature in the cell must be kept below 1023.15 K (750 °C) to avoid degradation, and the temperature difference between the ends of the cell must be kept below 75 K to avoid thermal stress on the cell. To accomplish these goals, the following manipulated variables are used: cell potential, makeup steam/hydrogen feed rate, steam/hydrogen ratio in makeup stream, sweep feed rate, fuel trim heater duty, sweep gas trim heater duty,

fuel stream recycle ratio, sweep stream recycle ratio, and vent gas recycle ratio. Classical control uses many P and PI control loops, detailed in Table 1. Derivative action was not used, because its benefits are severely degraded by measurement noise, which cannot be added to the PETSc TS integrator used in this study for dynamic simulations with classical PI control. The PI controller model in IDAES supports both variable bounds and anti-windup, both of which were used in these simulations.

**Table 1.** Variable pairings for classical control.

Controller Type	Manipulated Variable (MV)	Controlled Variable (CV)
PI	Cell potential	SOC fuel outlet H <sub>2</sub> mole fraction
P	Makeup feed rate	Hydrogen production rate
P	Sweep feed rate	SOC stack core temperature
PI (C11)	Steam heater duty	Steam heater outlet temperature
PI (C21)	Sweep heater duty	Sweep heater outlet temperature
P (C10)	Steam heater outlet temperature setpoint	SOC feed outlet temperature
P (C20)	Sweep heater outlet temperature setpoint	SOC sweep outlet temperature
None	Feed & sweep recycle ratios, makeup H <sub>2</sub> & H <sub>2</sub> O mole fractions, vent gas recycle ratio (used by NMPC, not PI)	

The cell potential has an immediate effect on the hydrogen composition of the fuel outlet stream; therefore, these two variables are paired. The makeup feed rate is then paired with the net hydrogen production/consumption rate (which can be calculated from the total current flowing through the cell, which is easily measurable). The trim heaters have an immediate impact on the stack inlet temperatures of their respective streams and are thus paired. However, it is desirable that the trim heaters also help adjust the SOC module's temperatures as needed. Therefore, a cascade arrangement is used, with another controller adjusting the inlet stream's setpoint based on the outlet stream's temperature. Typically, in cascade control, the inner controller is P and the outer controller is PI; however, an arrangement with the inner controller

PI and the outer controller P was chosen to avoid controller conflicts. Finally, because the trim heaters are not engaged in SOFC mode, the sweep blower flow rate is paired with stack core temperature (measured via a thermocouple embedded in the stack) to cool the cell as needed.

## NONLINEAR MODEL PREDICTIVE CONTROL

To compare the performance of classical control and advanced control strategies, an NMPC framework was developed for setpoint tracking using eight non-artificial MVs in Table 1: cell potential, makeup and sweep feed rates, feed and sweep recycle ratios, makeup H<sub>2</sub> and H<sub>2</sub>O mole fractions, and vent gas recirculation (VGR) ratio. Trim heater duties are not directly tracked to enable more freedom of adjustment.

Because makeup mole fractions are tracked instead of being fixed along a linear trajectory as in the classical control case, an equality constraint pinning their sum at 0.999 is introduced to the NMPC formulation. The remainder consisted of inert gases, with fixed mole fractions of 0.0008 N<sub>2</sub> and 0.0002 Ar. Feed heater duty is bounded between 0 MW and 2 MW and sweep heater duty between 0 MW and 4 MW for reasonable capital equipment sizing. Condenser vapor outlet temperature is fixed at 323.15 K under the ideal condenser assumption and thus is not an MV.

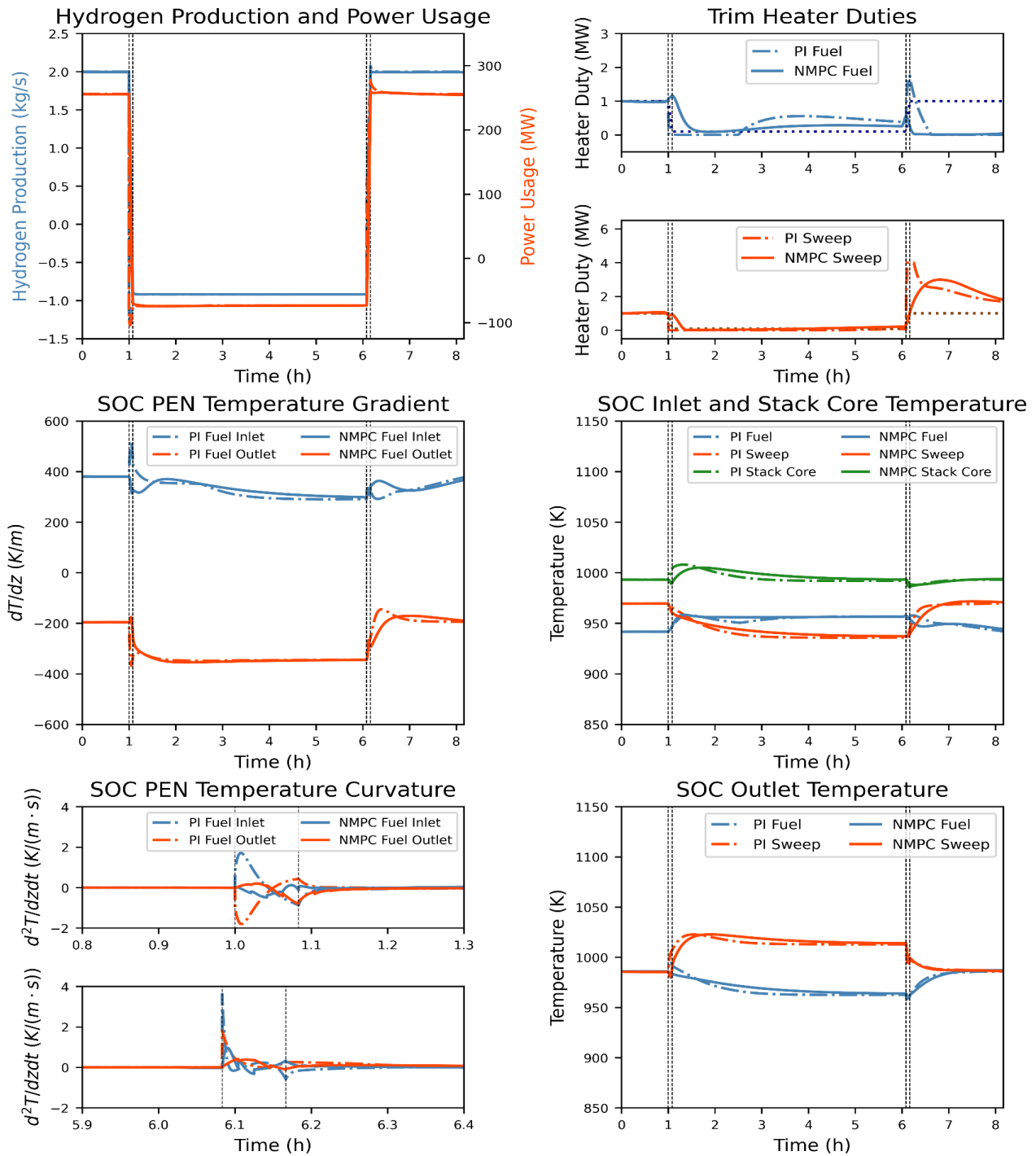
$$\begin{aligned}
 f_{obj} = & \sum_{i=0}^N \rho_{H_2} (y_i - y_i^R)^2 + \sum_{i=0}^N \sum_{j \in J} \rho_j (u_{ij} - u_{ij}^R)^2 \\
 & + \sum_{i=0}^N \sum_{k \in K} \rho'_k (x_{ik} - x_{ik}^R)^2 + \sum_{i=1}^N \rho' (v_i - v_{i-1})^2 \\
 & + \sum_{i=0}^N \sum_{z=1}^{z_L} \rho_M \left( \frac{\partial T_{iz}}{\partial z \partial t} \right)^2 \quad (1)
 \end{aligned}$$

As shown above, the objective function (eqn. 1) contains weighted sum of squared errors (SSE) of the trajectory-tracking of H<sub>2</sub> production rate as well as deviations of MVs and CVs from their reference values. A rate-of-change penalty on trim heater duties is added to attenuate oscillatory trajectories. In addition, to limit cell thermal degradation over time, the magnitude of PEN temperature curvature along the cell length (z-direction), is penalized. The first term in the objective function is the SSE of H<sub>2</sub> production rate,  $y_i$ , from its tracking target,  $y_i^R$ , at time  $t_i$ . The penalty on the H<sub>2</sub> rate tracking term,  $\rho_{H_2}$ , is selected to be 1. The second term involves penalties on SSEs of the tracked MVs,  $u_{ij}$ , from their references,  $u_{ij}^R$ ;  $J$  represents the set of tracked MVs. Similarly, the third term penalizes deviation of the CVs (represented by set  $K$ ) from their reference trajectories. The fourth term

is the rate of change penalty on trim heater duties, represented by  $v$ . Setpoint tracking and rate of change penalty terms are scaled to be  $O(1)$ , and 0.01 is selected for their penalties,  $\rho_i$ ,  $\rho_{k_i}$  and  $\rho'$ , to prioritize H<sub>2</sub> production rate tracking. The final term has a penalty of  $\rho_M$  on the sum of squares of PEN temperature curvatures in the z-direction.  $N$  is the number of steps in the prediction

horizon.

For this work, it is assumed that system dynamics is unambiguously captured in the controller model; in real-world application, potential plant-model mismatch can be addressed using moving horizon estimation (MHE) with an integrating disturbance model. The controller predicts



**Figure 3.** Comparison of classical control with NMPC in mode-switching operation.



system response for a given MV trajectory over the controller horizon with the system model and optimizes the MV trajectory for minimum objective function value. The first element of the solved trajectory is injected into the system, and the MV trajectory is re-optimized at the next sampling time over a shifted horizon. The actual decision variables in the NMPC problem are time derivatives of the MVs so that piecewise linear control profiles (first-order hold), instead of step inputs (zero-order hold), can be injected into the system. Placing a first-order hold on MVs, as opposed to a zero-order hold, smooths out the transition of control actions from the previous time step to the next and thus reduces spikiness in system response.

## SIMULATION RESULTS

For performance comparison of classical control and NMPC, the SOC system was linearly ramped from maximum hydrogen production (SOEC mode) to power production (SOFC mode) and back to maximum hydrogen production in simulation. A parametric sweep on  $\rho_M$  was conducted to investigate the trade-off between operating efficiency and cell degradation, which is dependent on PEN temperature curvatures. The ramps were carried out over 5 min, with the hydrogen-power ramp followed by 5 h settling at the new operating point and the power-hydrogen ramp followed by 2 h settling. Dynamic simulations using classical process control were conducted via the IDAES interface to the PETSc [12] suite of differential algebraic equation (DAE) solvers. Because this DAE system is stiff, a variable time step implicit Euler method was used. The time step was initialized at 0.1 s, after which it typically grew to 5 s to 10 s during the initial transient after ramping started or stopped, and then to 5 min to 10 min by the end of the integration interval. When anti-reset windup was turned on or off in the PI controllers, the time step decreased to 0.5 s to 1.0 s due to the steep transition, between error integrating and not. The fully discretized control problem for NMPC had approximately 47000 equations and variables. The studies were performed on an AMD Ryzen™ 9 processor @ 3.7 GHz with 16 GB memory. On average, the solution time was 2.4 CPU s for a prediction horizon of 375 s; the problem was solved well within the sampling time of 75 s.

Figure 3 compares the performance of the two control strategies, with  $\rho_M$  set to  $10^{-2}$  in the NMPC objective formulation;  $10^{-2}$  offers a good balance between settling speed of hydrogen production rate, the primary output variable, and addressing thermal degradation considerations. Both classical control and NMPC reach the SOFC mode production rate of  $-0.92$  kg/s from the SOEC mode production rate of  $2.00$  kg/s by the end of the five-minute ramp, with a considerable amount of overshoot observed in classical control. In the SOFC-SOEC ramp, both classical control and NMPC afford similar hydrogen production

rate tracking performance, with a small amount of overshoot in the former.

Curbing temperature gradient is critical to long-term operating performance of the cell. While PEN temperature gradients are not explicitly accounted for in either classical control or NMPC, both cases are well within the safety limit of  $1000$  K/m, and NMPC has advantages due to smoother trajectories and lower peak magnitudes at fuel inlet and outlet.

Moreover, the time derivative of SOC PEN temperature gradient describes the trend in the gradient's temporal variation. Here, NMPC affords lower peaks in the temperature curvatures at fuel inlet and outlet, as the final term in the NMPC objective (eqn. 1) imposes squeezing action on curvatures across all z-nodes. Cell temperatures in both control strategies generally take over 2 h to settle after the ramps finish. A slightly greater overshoot is observed in sweep channel inlet temperature for NMPC after the SOFC-SOEC ramp, and the same holds for classical control stack core temperature after the SOEC-SOFC ramp. While classical control temperatures mostly settle faster than the NMPC ones, this difference is less pronounced after the SOFC-SOEC ramp. Sweep heater duty in classical control saturates at the upper bound during the SOEC-SOFC switch, whereas NMPC produces smoother profiles well within design limits. Visible deviation of trim heater duties from setpoints reveals that the system has not reached steady state even by the time of the switch back from SOFC mode to SOEC mode. Despite the different heater duty profiles by classical control and NMPC, total power usage in both strategies is similar.

## CONCLUSIONS

In this work, a dynamic system flowsheet of an rSOC module and supporting auxiliary equipment was developed in the open-source, equation-based IDAES modeling framework. Control of this system for switching between maximum hydrogen production (SOEC mode) and power production (SOFC mode) was conducted with both classical control and NMPC. Dynamic simulation results show that although both control methods attain similar performance in many areas, a sophisticated classical control system involving non-intuitive cascade control arrangement was required to match the performance of NMPC in mode switching. Moreover, NMPC goes a step further by mitigating PEN temperature gradients and PEN temperature spatial-temporal derivatives along cell length more effectively than classical control does.

While classical control and NMPC attain similar total power usage in mode-switching operation, NMPC formulation in this work does not optimize for efficiency. Future work on economic MPC would allow the use of non-tracking objectives like efficiency optimization. Such

a scheme would be well suited to this rSOC system, since the system response time is slow compared to the rate at which electricity prices change. System performance while tracking more frequently alternating setpoint trajectories from fluctuating locational marginal prices (LMPs) is also critical in system economics evaluation. Another challenge is managing the trade-off between capital cost (cell degradation) and operating cost (setpoint tracking performance) over long-term operations. Finally, although the average CPU time in NMPC simulations already occupies only a small fraction of the sampling time, both advanced-step NMPC (asNMPC) and a distributed framework consisting of subsystem NMPCs can help drive down online computation delay; latency-free NMPC is a topic for further study.

## ACKNOWLEDGEMENTS

This work was conducted as part of the Institute for the Design of Advanced Energy Systems (IDAES) with support from the U.S. Department of Energy's Office of Fossil Energy and Carbon Management through the Simulation-based Engineering Research Program.

This project was funded by the Department of Energy, National Energy Technology Laboratory an agency of the United States Government, through a support contract. Neither the United States Government nor any agency thereof, nor any of its employees, nor the support contractor, nor any of their employees, makes any warranty, expressor implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## REFERENCES

1. Ferrari ML. Advanced control approach for hybrid systems based on solid oxide fuel cells. *Applied Energy* (145): 364-373 (2015).
2. Abbaker AM, Wang H, Tian Y. Enhanced Model-Free Discrete-Time Adaptive Terminal Sliding-Mode Control for SOFC Power Plant with Input Constraints. *AJSE* 47(3): 2851-2864 (2022).
3. Botta G, Romeo M, Fernandes A, Trabucchi S, Aravind P. Dynamic modeling of reversible solid oxide cell stack and control strategy development. *Energy Convers. Manag.* (185): 636-653 (2019).

4. Spivey BJ, Edgar TF. Dynamic modeling, simulation, and MIMO predictive control of a tubular solid oxide fuel cell. *J Process Control* 22(8): 1502-1520 (2012).
5. Schotman R. Dynamic modelling and nonlinear model predictive control of a reversible solid oxide fuel cell: for grid-tied power tracking. Delft University of Technology (2021).
6. Xing X, Lin J, Brandon N, Banerjee A, Song Y. Time Varying Model Predictive Control of a Reversible SOC Energy-Storage Plant Based on the Linear Parameter Varying Method. *IEEE Trans on Sustain Energy* 11(3): 1589-1600 (2020).
7. Raković SV, Levine WS Handbook of Model Predictive Control. Springer International Publishing (2019).
8. Rawlings JB, Mayne DQ, Diehl MM. Model Predictive Control: Theory, Computation, and Design. Nob Hill Publishing, LLC. (2020).
9. Lee A., Ghouse JH, Eslick JE, et al., The IDAES process modeling framework and model library— Flexibility for process simulation and optimization. *J Adv Manuf Process* 3(3): e10095(2021).
10. Allan, D. A. and Eslick, J. [https://github.com/IDAES/idaes-pse/blob/main/idaes/models\\_extra/power\\_generation/unit\\_models/soc\\_submodels/solid\\_oxide\\_module\\_simple.py](https://github.com/IDAES/idaes-pse/blob/main/idaes/models_extra/power_generation/unit_models/soc_submodels/solid_oxide_module_simple.py).
11. Allan, Douglas, et al. "NMPC for Setpoint Tracking Operation of a Solid Oxide Electrolysis Cell System." Retrieved from: <https://www.osti.gov/biblio/1964151> (2023).
12. Abhyankar S, Brown J, Constantinescu EM, Ghosh D, Smith BF, Zhang H. PETSc/TS: A Modern Scalable ODE/DAE Solver Library. *arXiv: Numer Anal* (2018).

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Integrated Design and Scheduling Optimization of Multi-product processes – case study of Nuclear-Based Hydrogen and Electricity Co-Production

Ruaridh Macdonald<sup>a</sup>, Dharik S. Mallapragada<sup>a\*</sup>

<sup>a</sup> MIT Energy Initiative, Massachusetts Institute of Technology, Cambridge, MA

\* Corresponding Author: [dharik@mit.edu](mailto:dharik@mit.edu)

## ABSTRACT

Increasing wind and solar electricity generation in power systems increases temporal variability in electricity prices which incentivizes the development of flexible processes for electricity generation and electricity-based fuels/chemicals production. Here, we develop a computational framework for the integrated design and optimization of multi-product processes interacting with the grid under time-varying electricity prices. Our analysis focuses on the case study of nuclear-based hydrogen (H<sub>2</sub>) and electricity generation, involving nuclear power plants (NPP) producing high temperature heat and electricity coupled with a high temperature steam electrolyzers (HTSE) for H<sub>2</sub> production. The ability to co-produce H<sub>2</sub> along with nuclear is widely seen as critical to improving the economics of nuclear energy technologies. To that end, our model focuses on evaluating the least-cost design and operations of the NPP-HTSE system while accounting for: a) power consumption variation with current density for the HTSE and the associated capital and operating cost trade-off, b) heat integration between NPP and HTSE and c) temporal variability in electricity prices and their impact on plant operations to meet a baseload hydrogen demand. Instead of formulating a monolithic optimization model, which would be computationally expensive, we propose a decomposition approach that reformulates the original problem into three sub-problems solved in an iterative manner to find near-optimal solutions. Through a numerical case study, we demonstrate the potential synergies of NPP and HTSE integration under alternative electricity price scenarios. This synergy is measured via the metric of relative breakeven H<sub>2</sub> selling price that accounts for the opportunity cost of reduced electricity sales from H<sub>2</sub> co-production.

**Keywords:** Hydrogen, Nuclear, Multiscale Modelling, Energy Systems, Electricity & Electrical Devices,

## INTRODUCTION

Despite its significant share of generation today and importance for climate change mitigation, nuclear power faces economic hurdles in many U.S. and other regions primarily due to depressed wholesale electricity prices stemming from increasing electricity generation from natural gas and variable renewable energy (VRE). For instance, between 2013 and 2021, 9.4 GW of existing nuclear power plants (NPP) have retired in the U.S. with an additional 7.2 GW scheduled to retire by 2025, mostly in regions with restructured electricity markets [1]. This economic outlook for existing U.S. NPPs also makes investments in new NPPs, based on next-gen small

modular reactor (SMR) concepts, challenging. However, despite their higher capital costs per kW relative to VRE generation sources, these NPP designs represent a type of low-carbon, dispatchable generation resource which has been shown to be critical to minimizing the cost of achieving deeply decarbonized power systems [2]. NPPs are not fully compensated for this benefit in current markets, so there is a need for alternative business models and revenue streams to support deployment of new NPPs to support economy-wide decarbonization goals.

In this context, the ability to deploy NPPs for simultaneously co-producing low-carbon hydrogen (H<sub>2</sub>) via water electrolysis that can be used for industrial applications is appealing for several reasons. First, industrial H<sub>2</sub>

demand, amounting to ~10 million tonnes per year in the U.S. in 2015 [3], tends to be centralized and constant in its consumption, both of which match well with least-cost NPP designs that tend to favor continuous operation and large-scale deployment. Second, H<sub>2</sub> demand is anticipated to grow, by up to 7X per one estimate [3], as part of efforts to decarbonize difficult to electrify end-uses. Third, high temperature heat available from next-generation nuclear reactors can be used to improve the electrical efficiency of H<sub>2</sub> production (and by extension, other industrial processes) by carrying it out at higher-than-ambient temperatures, such as high temperature steam electrolysis (HTSE) [4].

The design of integrated NPP-HTSE systems to co-produce H<sub>2</sub> and electricity needs to consider: a) the temporal dynamics in the economic value of electricity and its impact on co-product hydrogen price, b) the heterogeneity in design of new NPPs, particularly in terms of the maximum temperature of heat supply, c) the design of heat integration schemes between NPP and HTSE systems and d) the operating performance of the HTSE system, particularly the current density-dependent efficiency and heat management.

Previous NPP-HTSE system analyses have generally focused on light-water or pressurized water NPPs with lower temperatures of heat generation supply [5-7]. These analyses address detailed assessment of plant-level dynamics operations [5] as well as overall economic optimization using less-detailed plant-level models and soft-linking plant and grid-centric models [6-7].

Here, we develop an integrated design and scheduling optimization framework to address co-production of electricity and H<sub>2</sub> under time-varying electricity prices. For a given NPP design, electricity price series and exogenous H<sub>2</sub> demand, the model can evaluate the cost-effective sizing of HTSE, on-site H<sub>2</sub> storage and other auxiliary units (e.g. heat exchangers) as well as operation over a representative year, while adhering to a range of operational constraints. Crucially, we account for energetic and economic impacts of HTSE operation across a range of current densities & NPP-HTSE heat integrations.

Solving the proposed model in its original form takes several hours, even for representative periods as short as 12 hours. To include much longer representative periods, up to a year at hourly resolution, we accelerate solution of the proposed model by applying a decomposition strategy. This involves iteratively solving an upper-level investment problem sizing the HTSE using a line search algorithm, mid-level problem to size the heat exchanger network, and a lower-level operational model with fixed HTSE and heat exchangers which is formulated as a mixed integer linear program (MILP) and solved via Gurobi [19]. In addition, we utilize a Taylor expansion-based approximation of bilinear terms to speed up computation of the middle and lower-level operational model.

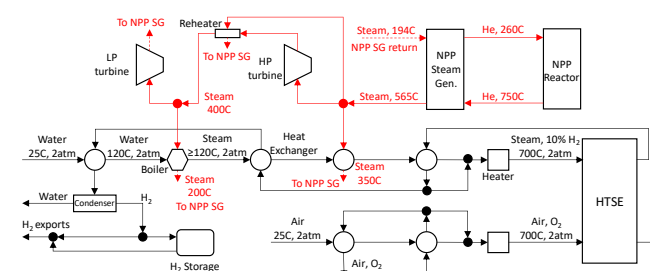
Finally, we propose a new metric of relative break-even price for H<sub>2</sub> production, which as opposed to the leveled cost of hydrogen (LCOH), accounts for the opportunity costs of forgoing electricity outputs in favor of H<sub>2</sub>. While the analytical framework has been developed for NPP-HTSE systems, it can be applied to other processes involving multiple co-products and dynamic interaction with the grid (e.g. electrified chemical production).

As a case study, we evaluate the integration of a high-temperature gas cooled reactor (HTGR) based NPP (750 °C) with an HTSE operating at 750 °C, where we quantify the value of heat integration between HTSE and NPP on the cost of hydrogen produced. We test the economics of the proposed system under current-day and future electricity prices and show how the relative break-even price as opposed to the LCOH is the appropriate metric for such multi-product systems.

## METHODOLOGY

### Process description

Figure 1 shows the schematic for the process where the HTGR is the primary heat source for the power generation and electrolysis processes. The HTGR produces hot gas leaving the reactor at 750 °C that is fed to steam generator producing steam at 565 °C and 16.5 MPa, that is used for power generation via a Rankine cycle [8]. A similar concept was chosen for a pilot project to provide heat and power for a U.S. petrochemical facility [9]. Part of the heat available from the HTGR can be used to improve the energy efficiency of HTSE H<sub>2</sub> production through feed heating. As seen in Figure 1, we allow for reactor to provide high temperature heat for steam superheating and low temperature heat for steam generation.



**Figure 1.** Overview of a high-temperature gas cooled reactor (HTGR) NPP integrated with HTSE operating at 750 °C. The black arrows relate to the HTSE flows and the red lines the nuclear power plant (NPP) flows. For simplicity, we have not shown the return of some steam flows back to the NPP steam generator (SG). All of the temperature and mass flow rates are free variables unless a value is given. The HTSE operating temperature is assumed constant and equal to the feed. Temperature of exhaust streams from HTSE are model variables but



must be within 100 °C of the inlet temperature to avoid thermal gradients [14].

## HTSE modeling

A key goal of our analysis was to understand how NPP and HTSEs might vary their electricity and H<sub>2</sub> outputs in response to electricity prices to maximize their profit. Therefore, it was important that we accurately model the dynamic operation of HTSE and how the efficiency varies as a function of the current density.

We developed a 0-D HTSE model based off the generic electrolyzer model produced by Orella [10] along with overpotential calculations from Buttler et al. [11]. The HTSE potential ( $V$ ) as function of current density ( $j$ ) at each time-step is calculated as sum of thermodynamic potential ( $V_{th}$ , Eq. 2), ohmic overpotential ( $V_{ohm}$ , Eq. 3), concentration overpotential ( $V_{conc}$ , Eq. 4-5), and activation overpotential for each electrode ( $V_{act,k}$ , Eq. 6). Each of these terms are related to SOEC system parameters, such as cell temperature ( $T$ ), exchange current density ( $j_{exchange,k}$ ), and partial pressures of components. Full details of the terms are given in [11].

$$V = V_{th} + V_{ohm} + V_{conc} + V_{act,c} + V_{act,a} \quad (1)$$

$$V_{th} = V_0 + \frac{RT}{2F} \ln \left( \frac{p_{O_2}^{0.5} p_{H_2}}{p_{steam}} \right) \quad (2)$$

$$V_{ohm} = -2.99 \times 10^{-5} j \delta e^{-\frac{10300}{T}} \quad (3)$$

$$c = -\frac{100\sqrt{2}jRT_{cath}}{2FPD_{steam}} \quad (4)$$

$$V_{conc} = \frac{RT}{2F} \ln \left( \frac{1 + \frac{c}{j_{H_2}}}{1 - \frac{c}{j_{steam}}} \right) \quad (5)$$

$$V_{act,k} = \frac{RT}{F} \sinh^{-1} \left( \frac{j}{2j_{exchange,k}} \right) \quad \forall k = a, c \quad (6)$$

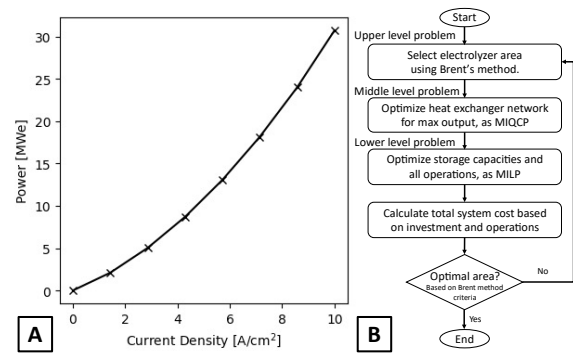
Given the HTSE stack area ( $A$ ) and time-dependent current density  $j_t$  (which is linearly proportional to the H<sub>2</sub> production rate for a given HTSE area by Faraday's Law), the power required to operate the HTSE at each time step  $t$  is given by Eq. 7, where the cell voltage ( $V$ ) is calculated as per Eq. 1-6 for the given current density.

$$P(j, A, t) = j_t A V_t(j_t) \quad (7)$$

As shown in Fig. 2A, the power consumption per unit HTSE area varies quadratically with the current density, meaning that the electrolyzer energy efficiency decreases with increasing current density owing to greater overpotentials. This creates an incentive to oversize the electrolyzer relative to demand, increasing the electrolyzer area ( $A$ ) and CAPEX to reduce the current density and hence power consumption required to achieve a given hydrogen production rate. Our modelling results exhibit this CAPEX - OPEX trade-off as we include a piece-wise approximation of the power vs. H<sub>2</sub> production curve for a fixed HTSE area.

The energy balance around the HTSE system is modeled with the constraint that feed stream and HTSE

operating temperature are fixed at 750 °C, and the ratio of hydrogen to steam at the cathode is held constant. At each time step, the heat balance of the HTSE is determined by the difference between its electric power demand and the enthalpy of the water splitting reaction, with high (low) current density operation generally leading to greater (less) overpotentials and heat generation. The HTSE is endothermic when operated below the thermoneutral potential ( $V_{th}$ ). The deficit or excess heat in the HTSE is balanced by either cooling or heating, respectively, the outlet streams. Outlet stream temperatures are model operational variables along with the HTSE stack current density. In this way, the HTSE dynamic operation is coupled with the operation of the heat exchanger network associated with feed preheating and product cooling shown in Fig. 1.



**Figure 2.** A: example piecewise power – current density relationship used in the model, for a 100m<sup>2</sup> HTSE producing up to 0.36 tonne H<sub>2</sub> per hour. We used 5 elements in our piecewise approximation. B: Overview of the three-part solution strategy used to optimize our model. MIQCP= Mixed integer quadratically constrained programme. MILP = Mixed integer linear programme.

## Model description and solution strategy

As shown in Eq. 8, the overall problem can be defined as a single mixed integer nonlinear program (MINLP). The objective function of this model is to minimize the total system cost while meeting a constant hourly hydrogen demand and several other operational and capacity constraints associated with units in Fig. 1.

$$\begin{aligned} \min_{x,y,z} & c_1 x + c_2^T y + c_3^T z \\ \text{s.t.} & f(x, y, z) = 0 \\ & h(x, y, z) \leq 0 \\ & g(x, y) \leq 0 \\ & x, y, z \geq 0 \end{aligned} \quad (8)$$

In Eq. 8,  $x$  is the stack area of the HTSE,  $y$  represents the vector of variables corresponding to the areas of the heat exchangers and boilers in the heat exchanger



network, and  $\mathbf{z}$  is the vector of variables corresponding to the capacity of the electricity and H<sub>2</sub> storage as well as the scheduling decisions for the entire system. Note that  $\mathbf{z}$  includes both continuous and binary variables, where the latter are associated with NPP start-up/shutdown [8].

Since we are interested in accounting for plant operational dynamics subject to hourly changes in electricity price over the year, solving the monolithic model of Eq. 8 becomes computationally challenging for representative periods longer than 12 hours. Therefore, we decomposed the original problem into three sub-problems that are iteratively solved to find the optimal solution. Fig. 2B shows the sequence of operations, where: a) the upper-level problem sizes the HTSE area, b) the middle-level problem sizes the heat exchanger network for a given HTSE area and maximum H<sub>2</sub> output for HTSE and c) lower-level operational problem evaluates cost-optimal plant operation over the entire year and sizes on-site H<sub>2</sub> storage capacity for a given HTSE area and heat exchanger network (via part a and b).

The upper-level investment problem is solved using Brent's method [12] to search for the HTSE area which minimizes the total system cost (capex + opex, Eq. 8). Brent's method is a gradient-free 1D minima-finding algorithm which combines the inverse quadratic interpolation, secant method, and bisection method. At each iteration, the ordering and value of function evaluations of the previous iterates, i.e. the results of solving the intermediate the lower problems, are compared to select the best methods to calculate the next iterate.

The middle-level problem optimizes the design of the heat exchanger network for a given HTSE area and several operating state of the system, corresponding to zero H<sub>2</sub> output, the mean output, and the maximum H<sub>2</sub> output. The only non-quadratic nonlinear constraints in the model pertain to the logarithmic mean temperature difference ( $\Delta T_{LMTD}$ ) in the heat exchanger sizing constraint. We approximate this by its first order Taylor expansion around initial guesses of the temperature differences ( $\Delta_{a,0}, \Delta_{b,0}$ ) for the various streams, as shown in Eq. 9. We set upper and lower bounds on each stream temperature based on the known temperatures (e.g. the HTSE operating temperature) and used the mid-point of these ranges as the initial temperature guesses.

$$\Delta_a = (T_{h,in} - T_{c,out}), \Delta_b = (T_{h,out} - T_{c,in}) \quad (9)$$

$$\Delta T_{LMTD} = \frac{\Delta_a - \Delta_b}{\ln(\Delta_a/\Delta_b)} \quad (10)$$

$$\Delta T_{LMTD} \approx \frac{\Delta_{a,0} - \Delta_{b,0}}{\ln(\Delta_{a,0}/\Delta_{b,0})} + \sum_{q \in [(h,in),(c,in)]} \frac{(T_q - T_{q,0})}{\ln(\Delta_{a,0}/\Delta_{b,0})^2} \left( \frac{\Delta_{b,0} - \Delta_{a,0}}{\Delta_{a,0}} \pm \ln \frac{\Delta_{a,0}}{\Delta_{b,0}} \right) + \sum_{q \in [(h,out),(c,out)]} \frac{(T_q - T_{q,0})}{\ln(\Delta_{a,0}/\Delta_{b,0})^2} \left( \frac{\Delta_{a,0} - \Delta_{b,0}}{\Delta_{b,0}} \pm \ln \frac{\Delta_{a,0}}{\Delta_{b,0}} \right) \quad (11)$$

This approximation is correct to within a few percent for the temperature ranges in our problem and is more accurate than using the arithmetic mean of the temperature differences when calculating heat fluxes. We

determined the heat exchanger areas by solving this middle-level problem for one hour of operation assuming the HTSE is operating at its mean hydrogen output for the given HTSE area. This problem is small, consisting of approximately 100 variables and 100 constraints with 27 quadratic constraints, and thus can be solved in under a second using Gurobi 10.0.

The lower-level problem is a mixed-integer quadratically constrained program (MIQCP) which optimizes the capacities of the battery and H<sub>2</sub> storage as well as the operating decisions of all the components for each of the 8760 hours of the year. The role of the H<sub>2</sub> storage is to allow for flexible operation of HTSE while meeting base-load H<sub>2</sub> demand. The battery storage could allow for enhancing flexibility of NPP by storing electricity at times of low electricity prices and discharging to operate HTSE or export to grid during high price periods.

To speed up solution of lower-level problem, we make two further approximations to convert it to a mixed-integer linear program (MILP): a) we approximate the HTSE power demand as a 1D piece-wise linear function of the current density using SOS2 constraints (Fig. 2A). b) the remaining quadratic constraints are related to the energy and mass balances in the heat exchanger network and splitters/mixers, respectively. For the bilinear terms, we employed a Taylor expansion-based linearization approach per Eq. 12, similar to the approach undertaken to approximate the LMTD in the middle-level problem.

$$xy \approx x_0 y_0 + (x - x_0) y_0 + (y - y_0) x_0 \quad (12)$$

We use the temperatures and mass flow rates from the intermediate problem solution to set  $x_0$  and  $y_0$ . The resulting MILP for a full year at hourly resolution has approximately 400,000 constraints and 400,000 variables and 1.5M non-zero terms after presolve. The solution time typically requires less than 30 seconds to solve the root relaxation followed by about three hours for branch and bound using 16 cores and Gurobi 10.0. The overall solution procedure of Fig. 2B typically requires 8-10 iterations to converge to the cost-optimal investment and operation so the total run time for the overall optimization is approximately 24 hours.

Our decomposition method can be improved in several ways. The upper problem could be solved using a gradient-based solver, making use of the duals of the lower problem. We could also solve more states of the system in the intermediate problem to both find Taylor expansion values for each time step, improving the accuracy of the linearization of the lower problem, and creating a warm start for the lower problem. This would reduce the time required for the branch and bound step.

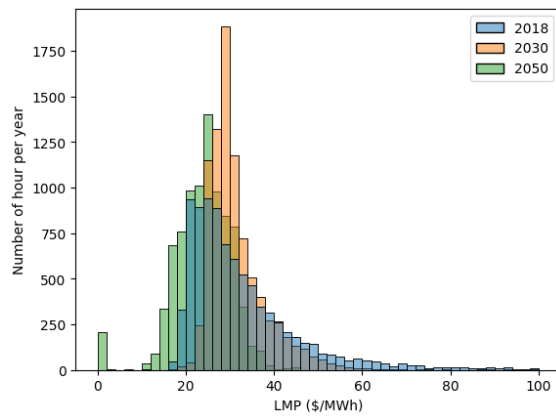
## Case study

We model the co-production of electricity and H<sub>2</sub> via the process in Figure 1 with a NPP with thermal capacity

200 MW<sub>t</sub> (80 MW<sub>e</sub>), located in Waterford, LA, USA. This location has an existing NPP and is close to many existing H<sub>2</sub> consumers (e.g. petrochemical plants). We assume the plant being optimized must supply a constant H<sub>2</sub> demand of 20 tonne/day, simulating H<sub>2</sub> demand from an industrial user. This means our optimization does not depend on the price of hydrogen. We do not consider cases where a facility can maximize revenue by freely choosing between selling electricity and hydrogen as this would require hourly timeseries of hydrogen prices.

### Electricity price scenarios

We evaluate the model outcomes for current grid conditions, represented by 2018 wholesale electricity prices at the Waterford site, as well as future grid scenarios for 2030 and 2050 by NREL for the same region [13]. Figure 3 shows the hourly electricity price distribution of the three price timeseries. The median 2030 price is greater than in 2018 but the distribution has shorter high and low tails. The mean and median 2050 prices are lower than both 2018 and 2030 and there are around 250 hours where electricity is priced at \$0/MWh.



**Figure 3.** Hourly electricity price distribution for the 3 scenarios evaluated in the study. 2018 prices are realized prices at the Waterford, LA node, while 2030 and 2050 scenarios are sourced from NREL ReEDS capacity expansion model outcomes [13]. Median prices for 2018, 2030 and 2050 are 29, 30, and 25 \$/MWh [13,24].

### Technology cost and performance assumptions

The major cost and performance assumptions used in the study are summarized in Table 1. We trial cases with and without a \$10/MWh transmission charges applied to imports. This charge reflects the difference in wholesale and industrial electricity prices [18].

### Relative breakeven price

We evaluated the cost benefits of co-producing electricity and H<sub>2</sub> by comparing the minimum selling price of H<sub>2</sub> produced by a NPP-HTSE co-production facility with the minimum selling price for a standalone HTSE

using grid electricity. The minimum H<sub>2</sub> selling price for a standalone HTSE is the price at which the project net present value is zero, i.e. annualized costs are equal to annualized revenues. This is the LCOH of the facility. However, computing the minimum H<sub>2</sub> selling price for a facility selling electricity and H<sub>2</sub> is more complicated since producing H<sub>2</sub> entails not selling the electricity used to produce the H<sub>2</sub>. This creates an opportunity cost which must also be recovered in the minimum selling price of the H<sub>2</sub>. We call this the relative breakeven price for H<sub>2</sub> associated with a co-production facility.

**Table 1.** Key cost and performance assumptions

Property	Value	Property	Value
<b>System</b>		<b>HTSE [10, 11,15, 16]</b>	
Hourly demand (tonne/hour)	0.83	Fixed costs (\$/m <sup>2</sup> /yr)	3,770
Discount rate	10%	Temperature (C)	750
<b>NPP [8, 17, 20, 21]</b>		Pressure (atm)	2
Reactor capacity (MWt)	200	Max. current density (A/cm <sup>2</sup> )	1
Turbine capacity (MWe)	80	H <sub>2</sub> mole fraction	0.1
CAPEX (\$/MWe/yr)	300,000	O <sub>2</sub> mole fraction	0.21
Fixed costs (\$/MWe/yr)	71,000	Cathode thickness (μm)	12.5
Variable Cost (\$/MWe)	15	Anode thickness (μm)	17.5
Minimum load	50%	Cell gap (μm)	12.5
Shutdown period (hours)	24	Diffusion coefficient (cm <sup>-2</sup> s <sup>-1</sup> )	5x10 <sup>-9</sup>
<b>H<sub>2</sub> storage [22]</b>		<b>Battery [23]</b>	
Fixed costs (\$/MWe/yr)	44,987	Fixed costs (\$/MWe/yr)	258,277
Compression energy (kWhe/kg)	0.4	Duration (hours)	4

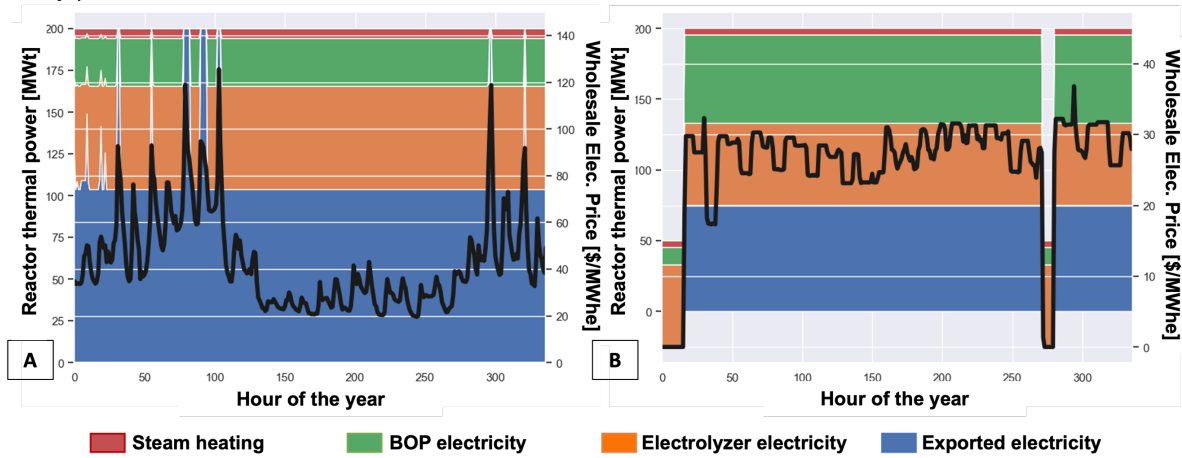
The relative breakeven price is calculated by finding the H<sub>2</sub> price which ensures a co-production facility earns at least as much profit as the NPP would make if operating independently and only sold electricity. This is done by solving the model in Eq. 8 twice: once as described for the full co-production facility and a second time where x = 0 (i.e. without a HTSE) and the exogenous hydrogen demand is set to zero. In the latter case, the NPP will operate independently selling electricity.

## RESULTS

### Base system design and operation results

Figure 4A shows the optimal dispatch of the base case system over two weeks using 2018 electricity prices. Given the high cost of HTSEs (Table 1), it is not

**Figure 4:** Optimal dispatch of NPP and HTSE over two weeks of operation for the 2018 and 2050 electricity price scenario and assumptions summarized in Table 1. Left panel shows reactor thermal balance (left axis) overlaid with electricity price profile (right axis, black line) for 2018 while right panel shows results for 2050. In all cases, system meets 20 tonne/day of baseload H<sub>2</sub> demand while adjusting grid electricity exports depending on electricity prices.



surprising to see that HTSE utilization is quite high despite fluctuating electricity prices. The HTSE output is only reduced during periods of very high electricity prices, when it is more profitable to reduce H<sub>2</sub> production and maximize power exports to the grid. This is enabled by 3 hours of H<sub>2</sub> storage and the HTSE being 8% oversized relative to the minimum HTSE area required to meet the H<sub>2</sub> demand constraint.

It might be expected that the facility would export less power during periods of high prices and use the cheaper electricity generated by the on-site NPP to reduce its average cost of electricity. However, deferring H<sub>2</sub> production till grid prices fall is less expensive overall – as long as storage is cheap enough – because less valuable electricity is used to produce H<sub>2</sub>, reducing the opportunity cost of hydrogen production.

The relative breakeven price of the co-produced H<sub>2</sub> accounts for the opportunity cost of lost electricity sales and should be compared to the LCOH of the independent HTSE. These figures are shown in Table 2. In this instance, an independent NPP will be loss-making so the opportunity cost of lost electricity sales is negative. This makes the relative breakeven price of NPP-HTSE H<sub>2</sub> less than its LCOH. An existing co-production facility should be willing to offer H<sub>2</sub> at the relative breakeven price, even though it will lose money on each kilo of H<sub>2</sub>, because it will lose less money than if it sold the same electricity.

Clearly, more revenue must be found to make the NPP-HTSE facility profitable overall especially for new facilities. One option is to sell H<sub>2</sub> priced at the LCOH of the NPP-HTSE facility. However, the losses being covered are from the electricity-side of the co-production facility. The relative breakeven price is the price required for the H<sub>2</sub>-side of the co-production facility to be profitable. In a

scenario with profitable independent NPPs, the relative breakeven price of NPP-HTSE H<sub>2</sub> is higher than its LCOH as the H<sub>2</sub> revenue must also recover the lost NPP profit.

The technical synergies from using NPP heat in the HTSE system reduces the cost of H<sub>2</sub> by approximately \$1/kg. We know this because the relative breakeven price of an NPP-HTSE which only exchanges electricity and where no import tariffs were in place should equal the LCOH of an independent HTSE in the same scenario. When there is no asymmetry in the price of buying and selling electricity, the cost of electricity from an onsite NPP is the same as that of purchasing grid power if the opportunity costs of lost NPP sales are also accounted for in the H<sub>2</sub> price. The results in Table 2 show that this is not the case. The relative breakeven price of the NPP-HTSE is \$1/kg cheaper than the no-tariff LCOH of the independent electrolyzer, indicating savings from using nuclear heat. This heating is used 5:1 to boil water versus superheat it. This could also be done using low-temperature NPPs, which operate at ~300 C.

H<sub>2</sub> produced by NPP-HTSEs is competitive with that from an independent HTSE, particularly if the grid has import tariffs. However, H<sub>2</sub> produced by both facilities is expensive, especially compared to the \$1/kg target. This is due to the relatively high capex of HTSE vs. state-of-art proton exchange membrane (PEM) electrolyzers. In addition, while HTSEs require less electricity input per kg of H<sub>2</sub> vs. PEM, they have lower current density limits than PEM electrolyzers, typically 1A/cm<sup>2</sup> vs. 2-3A/cm<sup>2</sup>. This means a large HTSE is required for the same H<sub>2</sub> output, compounding the cost difference.

### Impact of varying electricity price scenarios

Fig. 4B shows the optimal dispatch of the co-

production facility over two weeks under 2050 scenarios, for which the average electricity price is \$24/MWhe versus \$34/MWhe in 2018. The overall pattern of operation is similar to that in Fig. 4A. The HTSE is only 4% oversized, so there is little scope to vary its output much from the average H<sub>2</sub> demand. The NPP reduces its output during periods when the wholesale price falls to \$0.01/MWhe. This period only lasts seven hours so the NPP does not shutdown, as then it would have to wait a further 17 hours due to the 24-hour minimum shutdown constraint. The HTSE operates throughout.

The LCOH of the co-production facility is higher in 2050 than in 2018. This is because the price of electricity is lower and the NPP earns less revenue. However, the 2050 relative breakeven price of the NPP-HTSE is lower as this only considers the costs of producing H<sub>2</sub>, which is lower as electricity is cheaper. The reduction in prices also reduces the LCOH of the independent HTSE.

**Table 2.** The levelized cost of hydrogen (LCOH) and relative breakeven price of hydrogen for the independent HTSE and co-production facility in each of the three scenarios. The relative breakeven price of the independent electrolyzer is equal to its LCOH as it never incurs an opportunity cost when it produces H<sub>2</sub>.

	Electricity price timeseries year [13,24]		
	2018	2030	2050
Mean electricity price (\$/MWhe)	34.30	32.00	23.85
Median electricity price (\$/MWhe)	28.95	29.65	24.50
Co-production LCOH (\$/kg)	4.73	4.72	5.14
Co-production Relative breakeven price: (\$/kg)	2.54	2.29	1.95
Independent HTSE, no tariff LCOH (\$/kg)	3.62	3.43	2.90
Independent HTSE, \$10/MWh tariff LCOH (\$/kg)	4.52	3.79	3.38

H<sub>2</sub> produced by the NPP-HTSE facility in 2030 is \$0.3/kg cheaper than in 2018. While the mean price of electricity in the 2018 and 2030 scenarios is almost the same, the median price is higher, and the distribution of prices is shifted to the left and its right-hand tail is gone. This means there are more periods of low prices. By varying its output, both HTSE facilities can reduce their average price of electricity. In these runs, the effective price of electricity was as low as \$10/MWh. This trend continues in 2050, where the price distribution has a bimodal peak at \$0/MWh (see Fig. 3). However, the relatively high cost of H<sub>2</sub> storage limits the extent to which these periods of low prices can be taken advantage of.

## Sensitivity to the HTSE cost and current density limit

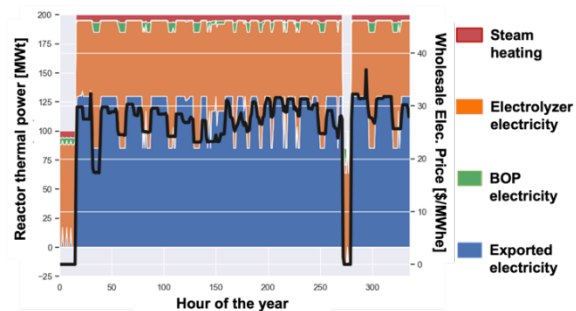
The optimized HTSE and NPP-HTSE facilities produce expensive H<sub>2</sub>. This is largely a function of the high cost of HTSEs and their low current density limit. We perform a sensitivity study of these values to see how they would affect the cost of H<sub>2</sub> and see which would be most impactful to improve.

Table 3 shows the results of the sensitivity study for the NPP-HTSE facility. As expected, increasing the current density or reducing the cost of the HTSE cause the biggest reduction in the H<sub>2</sub> LCOH and relative breakeven price. In both cases, it is economic to oversize the NPP-HTSE system significantly and mostly produce H<sub>2</sub> during periods of low prices. Fig. 5 shows an example of this for the 6A/cm<sup>2</sup> NPP-HTSE. The NPP exports and NPP-HTSE energy balance vary very closely with the price of electricity. As before, the NPP reduces its power output during zero-priced periods.

The other two changes had smaller impacts on the cost of H<sub>2</sub>. Reducing the cost of the NPP reduces the LCOH of H<sub>2</sub> as the NPP fixed costs are less, but has no impact on the relative breakeven price as the cost of electricity is unaffected.

**Table 3.** The LCOH and relative breakeven price of the four sensitivity study cases. Each was evaluated using the 2050 electricity price data.

	2050 Base-case	Low-cost HTSE [16]	900C HTSE	\$35 / MWh NPP [20]	6A/cm <sup>2</sup> HTSE
HTSE oversizing	4%	9%	1%	12%	70%
LCOH (\$/kg)	5.07	3.74	4.67	2.70	3.71
Rel. break. price: (\$/kg)	1.87	0.54	1.47	1.87	0.51



**Figure 5.** Dispatch and energy balance of the NPP over two weeks as part of the NPP-HTSE system for the 2050 system and a 6A/cm<sup>2</sup> HTSE.



## CONCLUSION

In this paper we have developed an integrated design and scheduling optimization model to study grid-interactive processes producing multiple products. To solve the model, a nonconvex MINLP, we proposed a three-level decomposition approach and range of tailored approximations to identify near-optimal solutions. We then demonstrated the approach in a short case study of electricity and hydrogen co-production using a high-temperature nuclear power plant and HTSE. While there are opportunities to improve the numerical stability and runtime of the method, we have shown that it can be used to capture HTSE part load operations and high temporal resolution efficiently.

In our case study, we demonstrated the cost advantage of co-producing electricity and hydrogen using an NPP and HTSE. We have shown that co-production facilities must consider the opportunity cost of not selling electricity when they price their hydrogen. We call this combination of the levelized cost of producing H<sub>2</sub> and the opportunity cost of not selling electricity the relative breakeven price. Through our sensitivity analysis, we highlighted how electricity prices, HTSE capital costs and current density limits, impact the minimum price of H<sub>2</sub>

## ACKNOWLEDGEMENTS

The authors acknowledge funding from Shell and feedback from Jacopo Buongiorno on the framing and approach through the course of the project.

## REFERENCES

1. CRS. U.S. Nuclear Plant Shutdowns, State Interventions, and Policy Concerns (2021)
2. Sepulveda, N. A., Jenkins, J. D., de Sisternes, F. J. & Lester, R. K. The Role of Firm Low-Carbon Electricity Resources in Deep Decarbonization of Power Generation. *Joule* 2, 2403–2420 (2018).
3. Ruth, M. et al. The Technical and Economic Potential of the H<sub>2</sub>@Scale Concept within the United States (2020). NREL
4. Buttler, A., Spliethoff, H. Current status of water electrolysis for energy storage, grid balancing and sector coupling via power-to-gas and power-to-liquids: A review. *Renewable and Sustainable Energy Reviews* 82, 2440–2454 (2018).
5. Kim, J. S., Boardman, R. D. & Bragg-Sitton, S. M. Dynamic performance analysis of a high-temperature steam electrolysis plant integrated within nuclear-renewable hybrid energy systems. *Applied Energy* 228, 2090–2110 (2018).
6. Frick, K., Wendt, D., Talbot, P., Rabiti, C. & Boardman, R. Technoeconomic assessment of hydrogen cogeneration via high temperature steam electrolysis with a light-water reactor. *Applied Energy* 306, 118044 (2022).
7. Frew, B., et al. Analysis of multi-output hybrid energy systems interacting with the grid: Application of improved price-taker and price-maker approaches to nuclear-hydrogen systems. *Applied Energy* 329, 120184 (2023).
8. Mulder, Eben. Overview of X-Energy's 200 MWth Xe-100 Reactor, NAS (2021)
9. Dow, Dow and X-energy advance efforts, 2023
10. Orella, M. J., et al. A General Technoeconomic Model for Evaluating Emerging Electrolytic Processes. *Energy Technology* 8, 1900994 (2020).
11. Buttler, A., et al. detailed techno-economic analysis of heat integration in high temperature electrolysis for efficient hydrogen production. *International Journal of Hydrogen Energy* 40, 38–50 (2015).
12. Brents' Method. [https://en.wikipedia.org/wiki/Brent%27s\\_method](https://en.wikipedia.org/wiki/Brent%27s_method)
13. National Renewable Energy Laboratory, <https://www.nrel.gov/analysis/cambium.html>
14. Petipas, Floriane, et al. "Thermal management of solid oxide electrolysis cell systems through air flow regulation." *Chem Eng Trans* 61 (2017)
15. Wendt, Daniel S., et al. High Temperature Steam Electrolysis Process Performance and Cost Estimates. No. INL/RPT-22-66117-Rev000. Idaho National Lab (2022)
16. James, Brian D., Prosser, Jacob H., Das, Sujit. HTE Stack Manufacturing Cost Analysis (2022)
17. Expert Finance Working Group. "Economic and Finance Working Group: SMR Roadmap." (2018).
18. EIA. Electric Power Monthly (2023)
19. Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual (2023)
20. Mulder, Eben, et al. Advanced Operation & Maintenance Techniques Implemented in the Xe-100 Plant Digital Twin to Reduce Fixed O&M (2022)
21. Davis, Ian, Braudt, Tom, Rackiewicz, David. Xe-100 Maintenance Staffing Optimization (2021)
22. NREL. 2023 Annual Technology Baseline (2023) <https://atb.nrel.gov/>.
23. Houchins, Cassidy, James, Brian D. Hydrogen Storage Cost Analysis (2022)
24. MISO. Real Time Market Report (2022)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Stochastic Programming Models for Long-Term Energy Transition Planning

Molly A. McDonald<sup>a,b</sup>, Christos T. Maravelias<sup>a,b,c,\*</sup>

<sup>a</sup> Princeton University, Department of Chemical and Biological Engineering, Princeton, NJ 08540, United States of America

<sup>b</sup> DOE Great Lakes Bioenergy Research Center, Princeton University

<sup>c</sup> Princeton University, Andlinger Center for Energy and the Environment, Princeton, NJ 08540, United States of America

\* Corresponding Author: [maravelias@princeton.edu](mailto:maravelias@princeton.edu).

---

## ABSTRACT

With growing concern over the effects of green-house gas emissions, there has been an increase in emission-reducing policies by governments around the world, with over 70 countries having set net-zero emission goals by 2050-2060. These are ambitious goals that will require large investments into the expansion of renewable and low-carbon technologies. The decisions about which technologies should be invested in can be difficult to make since they are based on information about the future, which is uncertain. When considering emerging technologies, a source of uncertainty to consider is how the costs will develop over time. Learning curves are used to model the decrease in cost as the total installed capacity of a technology increases. However, the extent to which the cost decreases is uncertain. To address the uncertainty present in multiple aspects of the energy sector, multistage stochastic programming is employed considering both exogenous and endogenous uncertainties. It is observed in scenarios when costs of emerging technologies decrease to competitive prices, decisions to invest in these technologies should be made earlier to allow for the decrease in costs to be taken advantage of in the future. Noticeably, a wider variety of energy and biofuel technologies are invested in when uncertainty is included. Interestingly, it is also seen that there are lower carbon emissions when uncertainty is considered.

---

**Keywords:** Stochastic Optimization, Design Under Uncertainty, Energy Systems

## INTRODUCTION

Due to concerns of global warming, many countries are setting goals to reach net-zero emissions sometime between the years 2050-2060 [1]. These goals will require large investments into low-carbon and renewable technologies, which will allow for emerging and underdeveloped technologies to advance. These technologies will need to be integrated into our current energy system efficiently. Energy system modeling can be used to inform decision makers about possible investment pathways that can help to reach net-zero emissions.

Energy system models are used to analyze how we can expand and operate a given set of technologies to meet specified goals. The model results can be used to inform policy makers of possible energy futures and show guiding decisions on the best investment choices to make today [2]. The results are not meant to be taken as the only way to reach the specified goals, but show

guiding decisions, such as what technologies are often invested in, what technologies are rarely invested in, or what technologies should be built in which regions. Energy system models tend to be over many years and include many sources of uncertainty since it is not known how parameter values are going to change over time.

When considering emerging or underdeveloped technologies, a source of uncertainty to consider is how the costs will develop over time. Learning curves are used to model the decrease in cost as the total installed capacity of a technology increases. The steepness of this curve, or how quickly the cost decrease, is based on a learning rate [3]. These learning rates can be estimated for any given technology, but ultimately are uncertain.

In order to address the uncertainty present in an energy system model, a sensitivity analysis is often completed to test how the outcomes will change as the input parameters are varied [2]. While this method does offer insight into how decisions might change as parameters

change, it does not offer a unique decision that can be made today to address the different parameter levels. The goal of this paper is to develop a multistage stochastic programming (MSSP) model in order to account for uncertainty and is applied to a problem for the United States using real data, allowing for the distinct parameter outcomes to be related today through first-stage decisions.

## METHODS

### Stochastic Programming

Stochastic programming is used to account for uncertainty in model parameters through the use of scenarios. In a two-stage stochastic program, the decision variables are partitioned into two sets: first-stage decisions and second-stage decisions. First-stage decisions are made before any uncertainty is observed and are the same for all scenarios, while second-stage decisions are made after uncertainty is observed and depend on the scenario. In a multistage stochastic program, there are more than two stages of decisions because uncertainty is observed sequentially throughout the time horizon. A multistage stochastic program is shown in equations 1-5 [4]:

$$\min z = c^1 x^1 + E_{s \in S^2} [\min q_s^2 y_s^2] + \dots + E_{s \in S^H} [\min q_s^H y_s^H] \quad (1)$$

$$\text{s.t. } W^1 x^1 = h^1 \quad (2)$$

$$T_s^2 x^1 + W^2 y_s^2 = h_s^2, \forall s \in S \quad (3)$$

...

$$T_s^H x^{H-1} + W^H y_s^H = h_s^H, \forall s \in S \quad (4)$$

$$x^1, y_s^t \geq 0, \quad t = 2, \dots, H; \quad (5)$$

There are H stages that the following vectors and matrices relate to.  $E_{s \in S^H}$  is the expectation of the costs across the scenarios in a given stage,  $c$  is the first-stage cost vector,  $q_s^H$  is the cost vector related to stage H,  $W^H$  is a recourse matrix,  $h_s^H$  is the right-hand side, and  $T_s^H$  is a technology matrix that is used to relate the current and previous stage decisions together. First-stage decisions are shown as  $x^1$  and  $y_s^H$  refer to the decisions at stage H and scenario  $s$ .

Uncertainty is classified into two types: exogenous and endogenous. If the uncertainty is exogenous, it is known when in time the uncertainty will be observed and scenarios will become distinguishable. In the case of endogenous uncertainty, it is not known when uncertainty will be observed and, thus, in a MSSP setting when scenarios will become distinguishable; rather, the timing depends on decisions [5]. For example, a cost reduction for a technology will only be observed if that technology is

chosen to be invested in; otherwise, it will remain unknown. Nonanticipativity constraints are used to group together scenarios that have not yet become distinguishable. For the case of exogenous uncertainty, a variable  $y$  must have the same value across scenario pairs until a certain time period when the uncertainty is observed, seen in equation 6.

$$y_{t,s} = y_{t,s'}, \forall (s, s', t) \in \Omega_t \quad (6)$$

Where  $\Omega_t$  is the set of scenario pairs that are indistinguishable at time period  $t$ .

For endogenous uncertainty, when the uncertainty will be observed depends on which decisions are made. Equation 7 is the nonanticipativity constraint for endogenous uncertainty. The variable  $y$  is now bounded by another variable  $z$ , which is the decision variable that relates to when the observation of uncertainty will occur. The timing of when this decision will be made is unknown.

$$-z_{t,s,s'} \leq y_{t,s} - y_{t,s'} \leq z_{t,s,s'}, \forall (s, s') \in \Psi \quad (7)$$

Where  $\Psi$  is the set of scenario pairs that are indistinguishable until uncertainty is observed.

### Value of the Stochastic Solution

The value of the stochastic solution (VSS) is used to measure the possible benefit of using stochastic programming. The calculation of the VSS is shown in equation 8.

$$VSS = EEV - OBJ\_MSSP \quad (8)$$

Where EEV is the expected result of using the solution of the deterministic model and OBJ\_MSSP is the optimal objective function value of the MSSP model. EEV is calculated by solving the deterministic model along each scenario path of the stochastic model and finding the expected value at the end of the time horizon using the same probabilities as the stochastic model [6]. The EEV is used in calculating the VSS rather than the optimal objective function value of the deterministic model because the deterministic model only considers one outcome of all parameters that is within the range of the scenarios that the stochastic model considers. By calculating the EEV, the deterministic model is exposed to the same uncertainty that the stochastic model is accounting for. This is a fairer comparison between accounting for uncertainty through stochastic programming and using a perfect foresight (deterministic) model when uncertainty is present.

## MODEL FORMULATION

The developed model is an energy system model similar to Rathi et al. with the addition of constraints on transportation and trade between regions and is applied to a problem for the entire United States [7]. The model

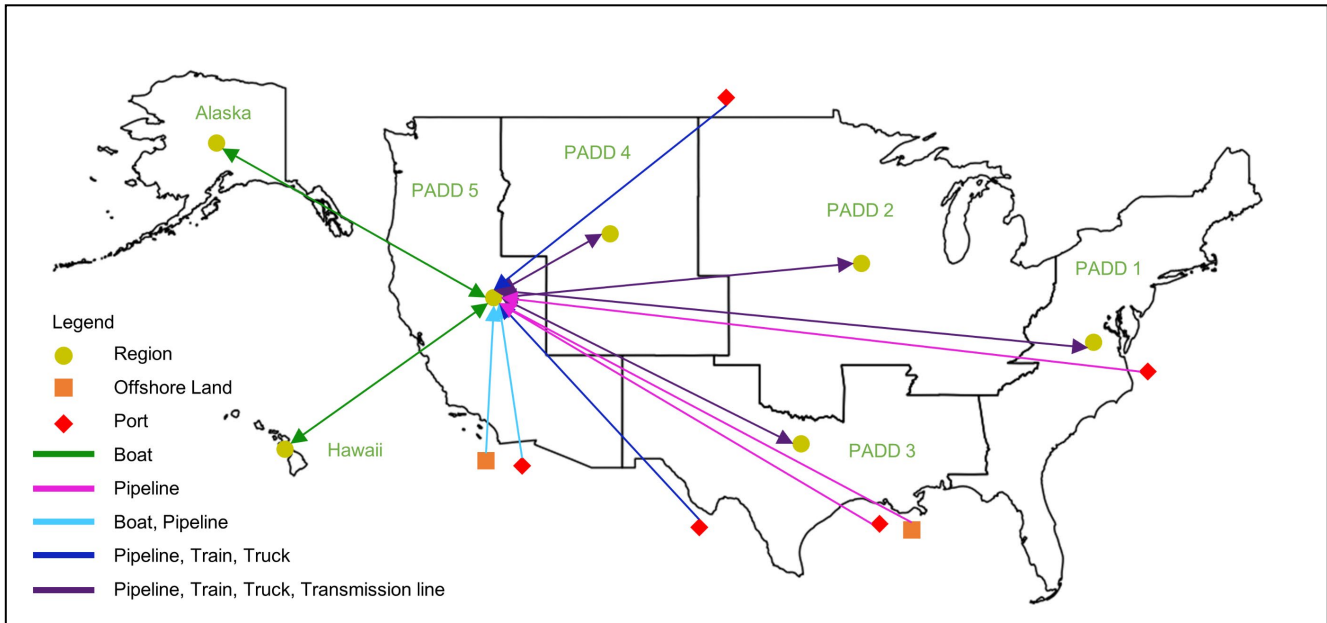
is a capacity expansion model that considers ten liquid fuel producing technologies, nine electricity generating technologies, and two intermediate technologies and must meet the expected demand in the United States over a 30-year time horizon that is divided into six 5-year time periods. For this problem, the United States is represented by 7 regions: Alaska, Hawaii, and 5 regions for the continental United States that follow the Petroleum Administration for Defense Districts (PADDs) as defined by the EIA [8]. Alaska and Hawaii are considered separate regions since they are disconnected from the continental United States. Trade between and within the different regions is also considered. The regions for the United States as well as an example of how one region is connected to the rest of the country is illustrated in Figure 1.

A deterministic and stochastic mixed integer linear model are developed that include constraints on capacity expansion, demand and supply, mass/energy flows through technologies, and conversion of components in

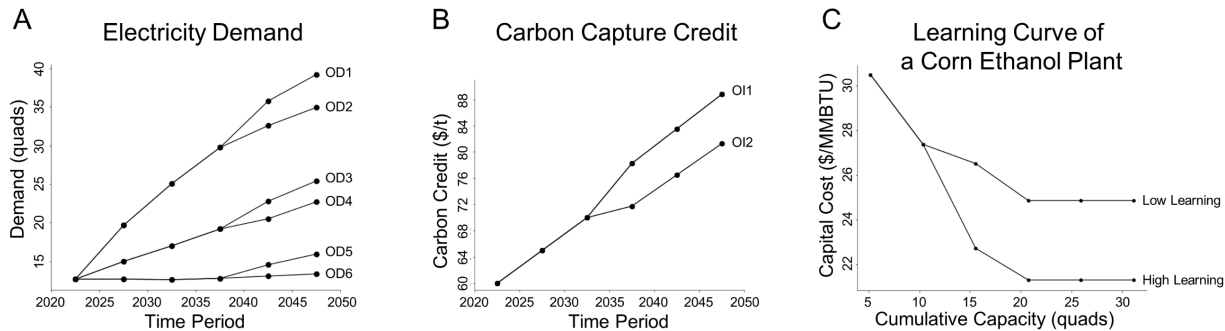
each technology. The models also include constraints on transportation between and within the regions, the material allowed to be transported by different transportation modes, and the direction material is allowed to flow between the regions. The objective function, shown in equation 9 for the deterministic model, is to minimize the cost of the system. The costs considered are the capital costs ( $C^{CAP}$ ), purchasing cost of raw material and final products ( $C^{PUR}$ ), operating costs ( $C^{OP}$ ), transportation costs ( $C^{TRANS}$ ), and a penalty for unmet demand ( $C^{PEN}$ ). The costs can be reduced through credits for carbon capture ( $C^{CAPTURE}$ ) and production of electricity from renewable sources ( $C^{REN}$ ).

$$\min COST = C^{CAP} + C^{PUR} + C^{OP} + C^{TRANS} + C^{PEN} - C^{CAPTURE} - C^{REN} \quad (9)$$

The stochastic model has the same costs in the objective function, but the objective is to minimize the



**Figure 1:** Map of the United States and example of how PADD5 is connected to the rest of the country. The color of the arcs indicate different transportation modes that are allowed between the different regions. Trade can occur in both direction of arcs with double sided arrows, while trade is restricted to one direction along arcs with single sided arrows.



**Figure 2:** Scenario trees for (A) the demand of electricity, (B) the credit for carbon capture, and (C) the learning curve for a corn ethanol plant.

expected cost across all the scenarios.

The system we study includes 24 scenarios representing the observation of two exogenous random parameters (demand and credits related to carbon emission reduction) and one endogenous random parameter (learning rate of biofuel technologies). The scenario trees for some of the random parameters can be seen in Figure 2. The outcome of the demand and credits related to carbon emission reduction are denoted as OD and OI in Figure 2A and 2B respectively. The total energy demand is assumed to be known, but what fraction that will be met by electricity or liquid fuels is unknown. The potential demand for electricity is presented in Figure 2A. There are six possible outcomes that relate to different levels of electrification with the highest electrification denoted as OD1 and the lowest level of electrification (or highest liquid fuel demand) denoted as OD6.

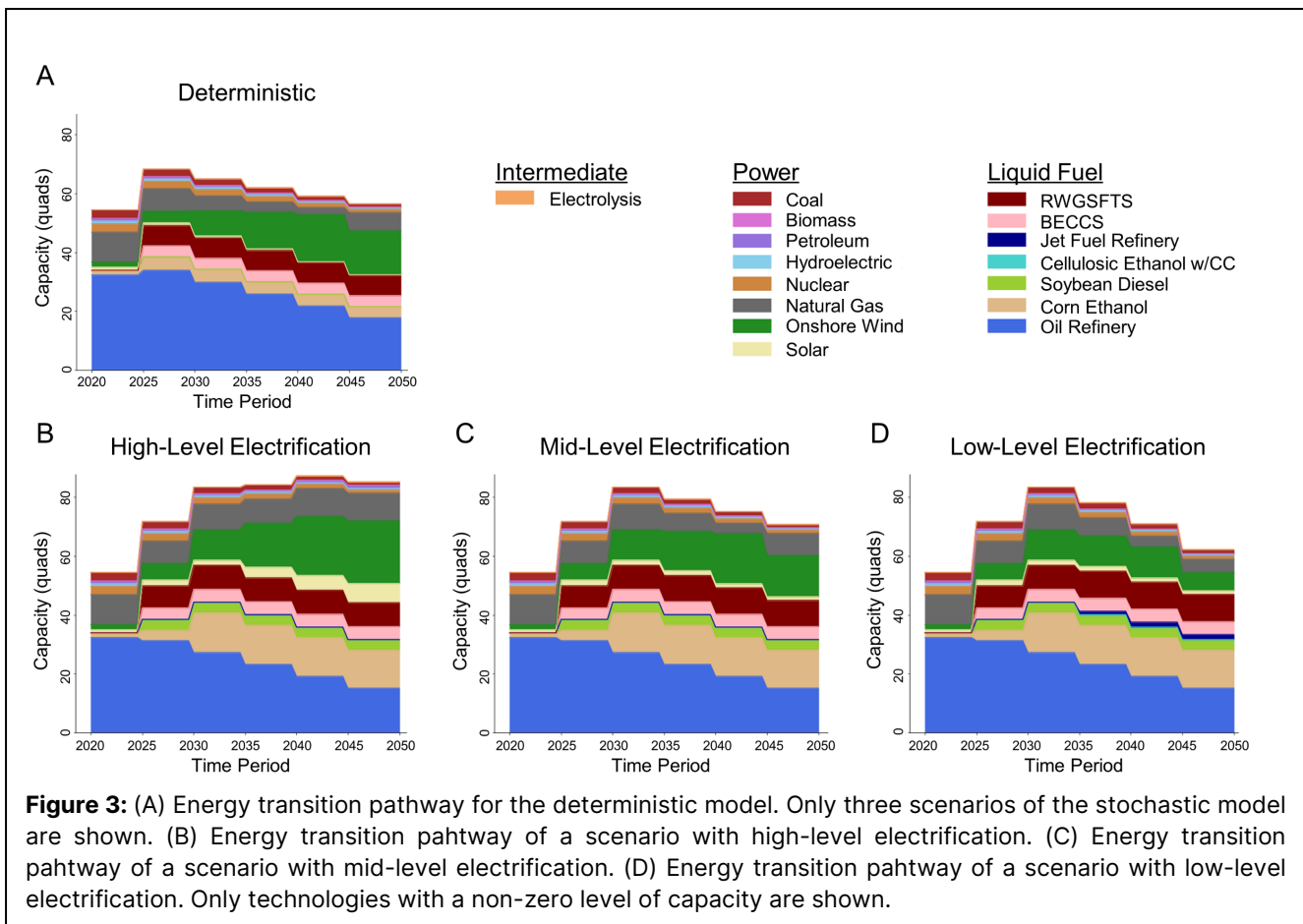
It is assumed that all biofuel technologies will learn together, meaning the capacity that will indicate a decrease in the capital cost along the x-axis of the learning curve, as shown in Figure 2C, will be based on the cumulative capacity of all biofuel technologies. For example, if an investment is made to expand corn ethanol plants, and the total installed capacity is increased from 5 quads to 10 quads, the cost will now decrease for all biofuel

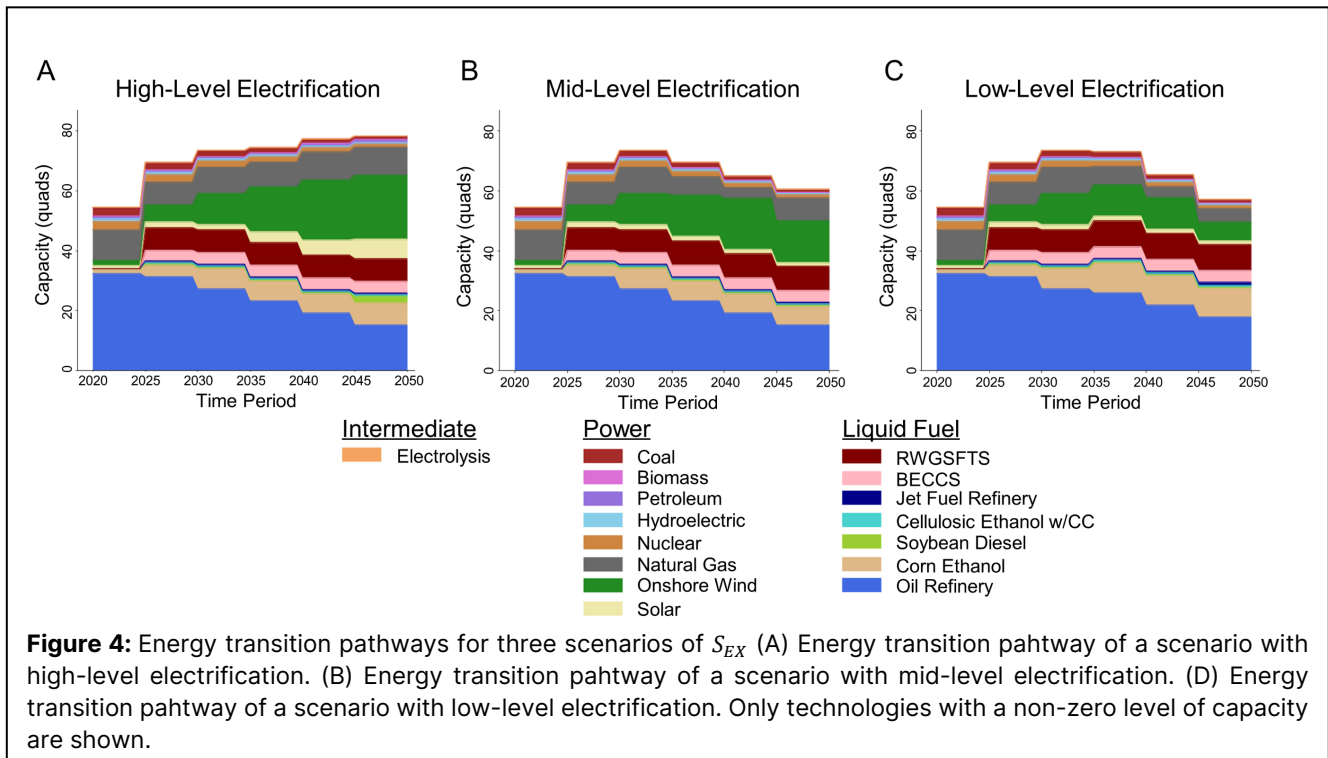
technologies. For instance, bioenergy with carbon capture and storage (BECCS) plants will also have a decrease in cost even though the total capacity of BECCS plants in the system has not changed. The decrease in cost will depend on the learning rate of the individual technology.

## RESULTS

### Energy Transition Pathway of the Deterministic and Stochastic Models

The initial investment decisions of the solution to the stochastic model are different than the initial investment decisions of the solution to the deterministic model. The energy transition pathway for the deterministic model and three scenarios of the stochastic model are presented in Figure 3A and Figure 3B-D respectively. The initial investment decisions of the stochastic model are the first stage decisions, meaning that they are the same across all scenarios. These decisions are seen in the second time period of Figure 3B-D. The capacity is not available until the second time period because we consider a 5-year lag between the time an investment decision is





made and when the capacity is available for use. When the first-stage decisions of the stochastic model are compared with the initial investment decisions of the deterministic model, it is observed that the stochastic model has more investment into onshore wind, solar, reverse water-gas shift combined with Fischer-Tropsch (RWGSFTS), soybean diesel, and cellulosic ethanol with carbon capture. This is because the stochastic model considers scenarios with both higher electricity and liquid fuel demand along with the possibility of lower capital costs in biofuel technologies.

The displayed scenarios have the same level of credits related to carbon emission reduction and learning rate but differ in the outcome of the demand. In the scenario with high and mid electrification there is an increase in electricity generating technologies by the end of the time horizon compared to both the pathway of the deterministic model and the low electrification scenario. The low electrification scenario has more liquid fuel production capacity and less electricity generating capacity by the end of the time horizon compared to the pathway of the deterministic model and the other electrification scenarios.

### Comparison of Accounting for Endogenous and Exogenous Uncertainty

The previous results were for a stochastic model that accounts for endogenous uncertainty. A stochastic model that considers only exogenous uncertainty was also considered. For the remainder of the paper, the stochastic model that accounts for endogenous uncertainty

will be referenced as  $S_{EN}$  and the stochastic model that accounts for exogenous uncertainty will be referenced as  $S_{EX}$ .

In  $S_{EX}$ , there is uncertainty in the fraction of the demand that will be met by liquid fuels or electricity and uncertainty in policy related to the reduction of carbon emissions. The uncertain parameters have the same scenario trees used in the  $S_{EN}$  shown in Figure 2. There is still a learning curve for each biofuel technologies that is based on the cumulative installed capacity of all biofuel technologies, but the learning rate is now assumed to be certain. The energy transition pathway for three scenarios of  $S_{EX}$  are shown in Figure 4. The scenarios shown in Figure 4A-C have the same level of demand and credits related to carbon emission reduction as Figure 3B-D. When only exogenous uncertainty is considered, there is less investment into the different biofuel technologies since there is not an opportunity for higher learning rates to be observed. Notably, there is less investment into corn ethanol and soybean diesel plants. This is because the initial capital cost of these technologies is less than other biofuel technologies; therefore, when endogenous uncertainty is considered through uncertain learning rates based on cumulative installed capacity, there are large investments into corn ethanol and soybean diesel plants in order to reach a cumulative installed capacity that is large enough for the other biofuel technologies to have competitive prices. Also, there is less total installed capacity compared to  $S_{EN}$ . The optimal objective function value of  $S_{EN}$  is less than 1% higher than the optimal objective function value of  $S_{EX}$ . This demonstrates that



while the objective value obtained considering endogenous uncertainty is only slightly higher than considering purely exogenous uncertainty, the investment decisions can be different, which was seen with higher investment into low-carbon technologies in  $S_{EN}$ .

The stochastic models are larger than their deterministic counterparts since there are additional variables and equations, defined for all scenarios, as can be seen in the formulation shown in equations 1-7. All models were implemented in GAMS 42.1.0 and solved using Gurobi 10.0.0 to a 1% optimality gap on 2.6 GHz Intel cluster machines. The deterministic model has 26,637 equations, 108,971 continuous variables, 1,773 binary variables and is solved in 2 seconds.  $S_{EX}$  has 345,277 equations, 1.3 million continuous variables, 21,276 binary variables and is solved in 1.5 hours. Lastly,  $S_{EN}$  has 739,021 equations, 2.6 million continuous variables, 42,552 binary variables and is solved in about 48 hours. Although the stochastic models are significantly more expensive, considering uncertainty can lead to significant cost savings compared to the perfect foresight approach.

### Benefit of Accounting for Uncertainty

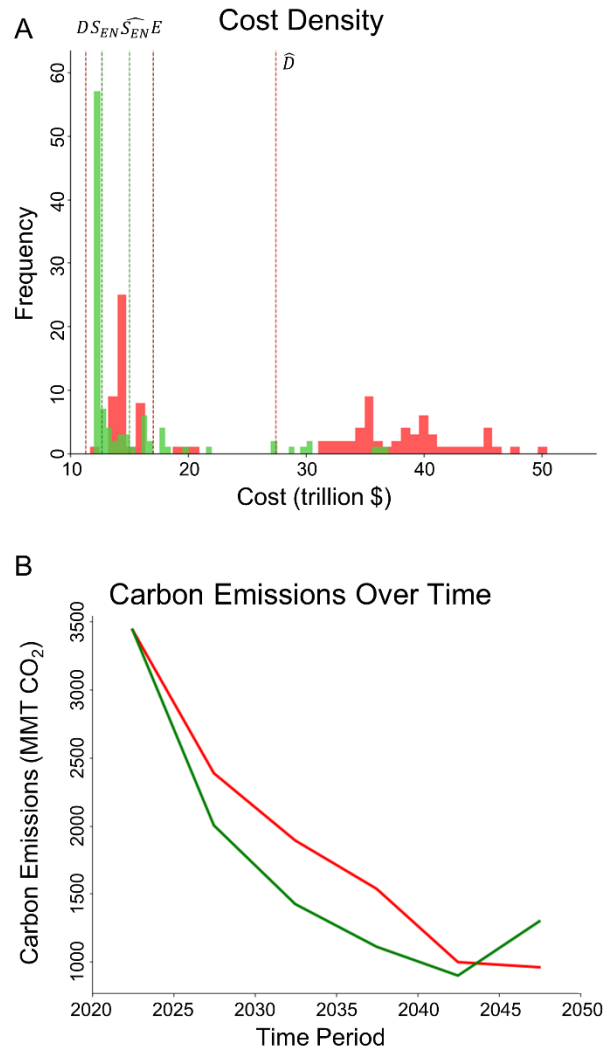
In order to test the benefit of using stochastic programming, the VSS was calculated. The optimal objective function value of the deterministic model is \$11.28 trillion and the optimal objective function value of  $S_{EN}$  is \$12.93 trillion. As explained earlier, the VSS is calculated with the EEV in order to have a fairer comparison between a model that is accounting for uncertainty and a model with perfect foresight. The EEV is \$17.07 trillion. This leads to a VSS of \$4.15 trillion, showing that there is a benefit to accounting for uncertainty through the use of stochastic programming.

It was also desired to test the sensitivity of the investment decisions of the deterministic and stochastic models to varying parameter values. To do this, the investment decisions of the deterministic model found when calculating the EEV and the investment decisions of the stochastic models were used as input parameters and the models were run 100 times with different parameter values. The parameter values that are varied are the random parameters considered in the stochastic models.

### Sensitivity of the results of the deterministic model and the stochastic model that considers endogenous uncertainty

Figure 5 shows the results when the parameter values were varied 100 times for the deterministic model and  $S_{EN}$ . The optimal objective function values (costs) for the deterministic model and  $S_{EN}$  as well as the EEV are shown as D, S, and E respectively along the top of Figure 5A. The average cost for  $S_{EN}$  and the average cost of the deterministic model are denoted as  $\widehat{S}_{EN}$  and  $\widehat{D}$  respectively.  $\widehat{S}_{EN}$  is lower than  $\widehat{D}$ , and the difference is larger

than the VSS calculated earlier. This shows that considering uncertainty with stochastic programming leads to a system that is more robust to changes in the input parameters.

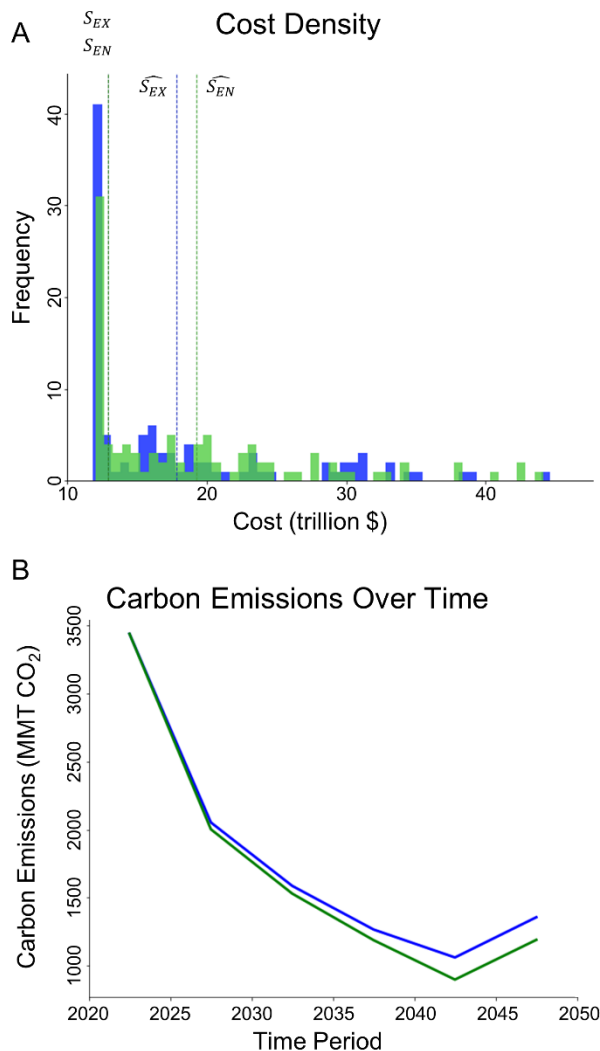


**Figure 5.** (A) Distribution of the cost of the deterministic model (red) and  $S_{EN}$  (green) for 100 varied parameter values. (B) Carbon emissions over time for the deterministic model and  $S_{EN}$ .

Figure 5B shows carbon emissions over time for the deterministic model and  $S_{EN}$ . The carbon emissions of  $S_{EN}$  are lower for most of the time horizon, due to higher investments into renewable and low-carbon technologies. There is an increase in emissions in the last time period because there is a finite horizon. Since there is no information past the last period on demand or credits to reduce carbon emissions, the model will increase emissions in order to meet the highest level of demands at the cheapest cost.

### Sensitivity of the results of the stochastic model that considers endogenous uncertainty the stochastic model that considers exogenous uncertainty

The sensitivity of  $S_{EX}$  and  $S_{EN}$  are shown in Figure 6. The parameter values were varied for 100 simulations with the same values used in the simulations to compare the investment decisions of the deterministic model and  $S_{EN}$ . The results of  $S_{EN}$  are the same in Figure 5 and Figure 6.



**Figure 6.** (A) Distribution of the cost of  $S_{EX}$  (blue) and  $S_{EN}$  (green) for 100 varied parameter values. (B) Carbon emissions over time for  $S_{EX}$  and  $S_{EN}$ .

The costs for  $S_{EN}$  and  $S_{EX}$  are shown along the top of Figure 6A and are similar, where the cost of  $S_{EN}$  is less than 1% higher than the cost of  $S_{EX}$ . The average cost of  $S_{EN}$  and  $S_{EX}$  are denoted as  $\widehat{S}_{EN}$  and  $\widehat{S}_{EX}$  respectively. The average costs are similar but  $\widehat{S}_{EN}$  is slightly higher. The carbon emissions over time (Figure 6B) are also similar,

but  $S_{EN}$  has slightly lower emissions than  $S_{EX}$ . Although the costs are not significantly different, it is important to note that the technology mix can be different, as seen in the differences in capacities of Figure 3B-D and Figure 4A-C.

## CONCLUSIONS

An energy system model that considers capacity expansion of liquid fuel producing technologies and energy generating technologies to meet the energy demand was developed and applied to a problem for the United States using real data. As most literature energy system models, it was assumed that there is a single decision maker for the entire system, which, obviously, is not the case. Nevertheless, the results of the proposed models allow us to better understand the energy sector by identifying what levels and types of investments are necessary to meet certain goals and constraints, such as net zero emissions. A multistage stochastic programming model was developed to consider both exogenous and endogenous uncertainty. The exogenous parameters considered were the demand in electricity and liquid fuels, and credits related to reducing carbon emissions. Endogenous uncertainty was introduced through uncertain learning rates of biofuel technologies. When uncertainty is considered through stochastic programming there are higher initial investments into renewable and low-carbon technologies, which will be required in order to meet net-zero emission goals. With the calculation of the VSS, it was seen that there is a cost benefit when we consider uncertainty rather than assuming perfect foresight. Also, it was shown that the results of the stochastic model led to solution with additional desired characteristics, such as lower carbon emissions and a technology mix that is more robust to varying input parameter values. Lastly, it was seen that introducing endogenous uncertainty does not increase the cost of the system significantly compared to considering only exogenous uncertainty, but there can be a difference in the investment decisions.

## ACKNOWLEDGEMENTS

This material is based upon work supported in part by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Biological and Environmental Research Program under Award Number DE-SC0018409 and by the Innovation Research Grant from the Andlinger Center for Energy and the Environment at Princeton University.

## REFERENCES

1. United Nations.

<https://www.un.org/en/climatechange/net-zero-coalition>

2. Larson E, Greig C, Jenkins J, Mayfield E, Pascale A, Zhang C, Drossman J, Williams R, Pacala S, Socolow R, Baik E, Birdsey R, Duke R, Jones R, Haley B, Leslie E, Paustian K, Swan A. Net-zero america: potential pathways, infrastructure, and impacts. Princeton University (2021)
3. Daugaard T, Mutti LA, Wright MM, Brown RC, Componation P. Learning rates and their impacts on the optimal capacities and production costs of biorefineries. *Biofuels Bioprod. Biorefining* 9:82–94 (2015)
4. Birge JR, Louveaux F. Introduction to Stochastic Programming. Springer Science & Business Media. (2011)
5. Goel V, Grossmann EI. A class of stochastic programs with decision dependent uncertainty. *Math. Program.*, 355–394 (2005)
6. Escudero LF, Garín A, Merino M, Pérez G. The value of the stochastic solution in multistage problems. *TOP*, 15:48–64 (2007)
7. Rathi, T., Zhang, Q. (2022). Capacity planning with uncertain endogenous technology learning. *Computers and Chemical Engineering*, 164:107868
8. U.S. Energy Information Administration. <https://www.eia.gov/todayinenergy/detail.php?id=4890>

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# An MINLP Formulation for Global Optimization of Heat Integration-Heat Pump Assisted Distillations

Akash Nogaja<sup>a</sup>, Mohit Tawarmalani<sup>a,b</sup>, and Rakesh Agrawal<sup>a\*</sup>

<sup>a</sup> Purdue University, Davidson School of Chemical Engineering, West Lafayette, IN, USA

<sup>b</sup> Purdue University, Daniels School of Business, West Lafayette, IN, USA

\* Corresponding Author: [agrawalr@purdue.edu](mailto:agrawalr@purdue.edu).

---

## ABSTRACT

Thermal separation processes, such as distillation, play a pivotal role in the chemical and petrochemical sectors, constituting a substantial portion of the industrial energy consumption. Consequently, owing to their huge application scales, these processes contribute significantly to greenhouse gas (GHG) emissions. Decarbonizing distillation units could mitigate carbon emissions substantially. Heat Pumps (HP), that recycle lower quality heat from the condenser to the reboiler by electric work present a unique opportunity to electrify distillation systems. In this research we try to answer the following question in the context of multi-component distillation – Do HPs actually reduce the effective fuel consumption or just merely shift the fuel demand from chemical industry to the power plant? If they do, what strategies consume minimum energy? To address these inquiries, we construct various simplified surrogate and shortcut models designed to effectively encapsulate the fundamental physics of the system. These models are integrated into a superstructure-based Mixed-Integer Nonlinear Programming (MINLP) formulation, which is amenable to global optimization algorithms aimed at minimizing the effective fuel consumption of the system. Moreover, through the examination of a toy 4-component alcohol separation example, we demonstrate how HPs can notably reduce carbon emissions, even when the consumed electricity is generated by burning fossil fuels.

---

**Keywords:** Distillation, Optimization, Energy Efficiency, Process Design

## INTRODUCTION

With more than 40,000 operational columns in the USA alone, distillation accounts for 90-95% of all fluid separations [9]. Distillation is a thermal method that uses differences in component volatilities to achieve separation. In the current economy, thermal energy is generally provided as heat by the combustion of fossil fuels. Given the immense scale at which distillation is deployed, its cumulative energy consumption constitutes a substantial 2.5% of the overall energy consumption within the United States [6]. Consequently, distillation is also responsible for a significant fraction of GHG emissions.

One of the solutions to minimize this significant carbon footprint of distillation systems is to electrify them. Two predominant advantages of electrification are as follows:

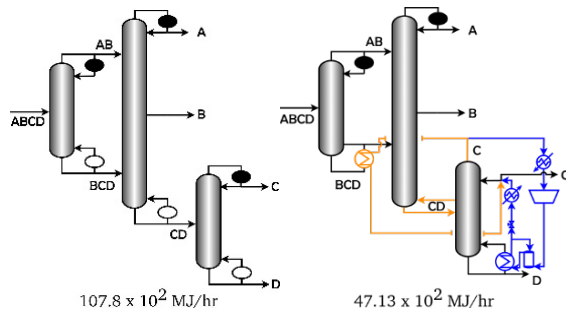
1. Even within the current economic framework,

which includes a portion of electricity generation sourced from non-renewables, electrification can potentially curtail fuel consumption.

2. Additionally, it steers the chemical industry toward the imminent shift of energy sources to renewable electricity [2].

In this study, we delve into the application of heat integration (HI) and heat pumps (HP) for electrifying distillation systems. Heat Integration involves the transfer of energy from a heat source with a higher temperature to a lower temperature level where heat is needed [13]. In contrast, heat pumps are devices that can upgrade the heat from a lower temperature source to a higher temperature sink using mechanical work [15]. They are particularly useful in binary and multi-component distillation systems because the condenser's heat at lower temperatures can be upgraded effectively to meet the reboiler

heat demands [3].



**Figure 1.** (a) One of the many possible configurations (b) HI-HPAD version of (a) (The part in yellow represents heat integration, the part in blue represents heat pump)

We demonstrate the synergistic use of the two technologies using an equimolar 4-component feed (Benzene, Toluene, p-Xylene and o-Xylene) separation. For a 100 kmol/hr feed, the configuration in Figure 1 (a) consumes about 130% more energy than the HI-HPAD version of the same configuration in Figure 1 (b).

Figure 1 (a) shows one of the many possible sequences for a 4-component separation. It's important to note that in the first column, the feed 'ABCD' undergoes a sloppy split i.e. one or more of its components are distributed in the distillate and the residue. Considering these splits is crucial during the design of a distillation sequence. By modifying the distribution of components, it becomes possible to influence the temperature of the condenser and reboiler. This, in turn, facilitates the establishment of heat integration links between condensers and reboilers. For example, by altering the split of component 'B', the temperature of the first column's reboiler is brought down such that vapor from the third column can heat the reboiler of the first column (see Figure 1(b)). Further, in this research, we focus on a specific type of heat pump called as – 'Mechanical Vapor Recompression' (MVR) due to its widespread applicability in the literature and industry. In this setup, the vapors from the overhead are compressed to a higher pressure and then condensed against the boiling liquid in the reboiler as in the third column in Figure 1 (b). Note that the reboiler of the third column is partly powered by steam (not shown in the diagram). The potentially high coefficient of performance (COP) and versatile applicability make the technology an appealing option for the electrification process.

The decisions pertaining to the number of configurations, their operating conditions, and HI-HP links for multicomponent separations increase combinatorially, prohibiting industrial practitioners from evaluating every possibility in commercial process simulators [3,17]. This can result in suboptimal configurations being employed and a higher energy penalty. The goal of this research is to develop a method that generates a rank list of configurations exploring all HI-HP possibilities that help

industrial practitioners in decision making.

Towards that goal, in the following section 'Model Development', we propose equations that capture the essential thermodynamics with high fidelity while still being tractable by global optimization solvers like Gurobi and BARON. Further, the developed equations are organized in a superstructure based MINLP formulation that embeds all possible distillation sequences (including sloppy splits), heat integration links and heat pump connections. Subsequently, we validate the proposed model by applying it to a binary alcohol separation system using the commercial process simulator ASPEN Plus V11. Additionally, through a demonstration involving a toy 4-component alcohol separation problem, we showcase the model's efficacy in identifying configurations that significantly reduce effective fuel consumption.

## MODEL DEVELOPMENT

Developing such a framework presents three challenges – determining the energy consumption of a split given the component distribution, estimating temperature of multi-component phase changing streams and the electricity consumption of compressor in the heat pump. We solve the three problems systematically in this section -

### Heat Duty Estimation in Multicomponent Distillation Sequencing

The distillation sequencing problem has received extensive attention in the literature [1,4,5,17,18]. A comparative analysis among different configurations is possible when each configuration operates at its global minimum energy. Hence, we use state of the art superstructure based global optimization model relying on the Underwood's method to compute minimum energy consumption of a given split [7].

State-of-the-art methods directly using Underwood's equations minimize vapor consumption and employ the Constant Molar Overflow (CMO) assumption. We relax this assumption using the CHT (Constant Heat Transport) model [12] which employs a simple variable transformation for the flow variables -

$$F_i^{LH} = F_i \lambda_i$$

$$z_i^{LH} = \frac{z_i \lambda_i}{\sum_{j=1}^n z_j \lambda_j} \quad (1)$$

$F_i^{LH}$  is the latent heat variable defining heat flow,  $F_i$  is the molar flow and  $\lambda_i$  is the molar enthalpy of vaporization of component 'i' in the feed. This straightforward transformation allows the computation of the needed heat duties for the heat pump calculations.



## Multi-component Phase Change Temperature Estimation

In multi-component distillation systems, we frequently encounter streams that undergo non-constant temperature phase changes. To guarantee a positive driving force in the heat exchanger, involving the condensation of compressed vapor and boiling liquid, we have developed an easy-to-use surrogate model [14] –

$$T = \frac{B_{\text{mix}}}{A_{\text{mix}} + \ln(\rho)} - C_{\text{mix}}; \rho = \frac{P^{\text{ref}}}{P} \sum_i \alpha_i X_i \quad (2)$$

The parameters  $A_{\text{mix}}$ ,  $B_{\text{mix}}$  and  $C_{\text{mix}}$  are specific to a mixture and are trained using data from experiments or detailed thermodynamic models. The variable  $\rho$  is called the 'pressure scaled pseudo relative volatility'. It captures the thermodynamic state (liquid fraction) of the stream by changing liquid mole fraction variables. In our recent article [14], the model underwent testing across various non-azeotropic multi-component systems, yielding a high coefficient of determination ( $R^2$ ) of approximately 0.99 for all the cases. An expedient and accurate estimation of the temperatures of fluids involved in heat pumping is essential.

## Heat Pump Electricity Consumption

Similar to the temperature model, the electricity consumption of the compressor is modeled by making a 'reversible heat pump' approximation –

$$W = F^{\text{LH}} \int_0^1 \left( \frac{T(P_2)}{T(P_1)} - 1 \right) dq^{\text{LH}} \quad (3)$$

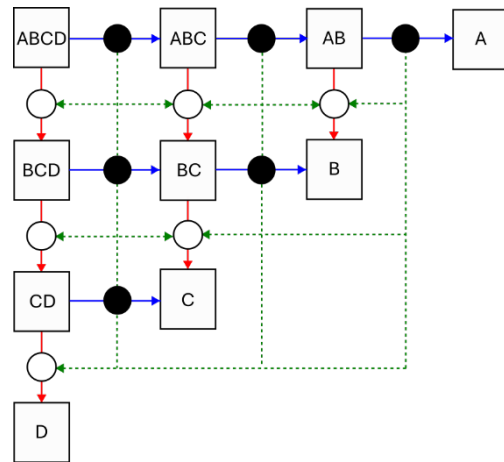
$$q^{\text{LH}} = \frac{L^{\text{LH}}}{F^{\text{LH}}}$$

$P_2$  is the pressure of the compressed overhead vapor in the heat pump while  $P_1$  is the pressure of the column. The integral factorizes the effect of changing temperature as the stream undergoes phase transition.

Leveraging temperature and work consumption models, in the next section, we construct a superstructure atop the MINLP model by Tumbalam Gooty et al. (2022) incorporating heat pump links between various condensing and boiling streams.

## MINLP SUPERSTRUCTURE

The components in the MINLP formulation are represented by alphabets, with their order corresponding to decreasing relative volatility. Each stream is represented by a sequence of alphabets, where each alphabet corresponds to a distinct component present in the stream (see Figure 2).



**Figure 2.** MINLP Superstructure for HI-HP integrated distillation column sequencing. Streams (Squares), Condensers (filled circles), Reboilers (empty circles) for the separation of a 4-component mixture. The dotted green paths represent the heat-work network.

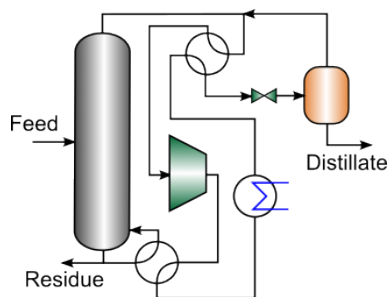
To establish a heat-work network within the existing sequencing superstructure, we connect all condensers to all reboilers and assign a binary variable representing the presence or absence of a connection between each pair. The state of these binary variables, activated or deactivated, is determined by the driving force for heat exchange.

This superstructure accommodates both heat pumps and steam utility to power the reboilers. Our objective is to minimize the effective fuel consumption for generating the required electricity and steam, assuming a boiler efficiency of 85% and a power plant efficiency of 50%.

## MODEL VALIDATION:

The effectiveness of the MINLP model for multi-component heat duty prediction has been thoroughly established through extensive testing on a variety of mixtures [12,16]. Additionally, the temperature model's accuracy has been verified by comparing it to detailed thermodynamic calculations [14]. In this study, we focus solely on validating the performance of the 'reversible heat pump model'.

We validate the model using two test cases – Methanol & Ethanol, Methanol & Propanol. These cases are specifically chosen to evaluate the model performance with increasing boiling point differences in the feed components. The flowsheet simulated in ASPEN Plus V11 is shown in Figure 3.



**Figure 3.** HPAD for binary distillation simulated in ASPEN Plus V11

**Table 1:** Effective fuel consumption for Steam driven and HPAD distillation. The saturated liquid feed mixture is set to be equimolar. The energy consumption values are expressed in the unit MJ/kmol feed.

Binary Mixture	ASPEN Steam Driven		Shortcut Model Steam Driven	
	Steam Driven	HPAD	Steam Driven	HPAD
Methanol - Ethanol	76.1	10.2	74.7	8.9
Methanol-Propanol	39.4	9.7	37.1	9.2

The feed to the distillation column is configured to be equimolar, and the target product purity is set at 99%. As the research primarily centers around a comparative analysis of various strategies, the number of stages' degrees of freedom (DOF) is excluded from the analysis. In ASPEN simulation, a uniform number of stages, specifically 200, is applied for all simulations. Additionally, the minimum approach temperature in the heat pump systems is standardized to be 5°C. Note that the shortcut model performs calculations for a pinched column, therefore, a lower energy demand than the corresponding ASPEN simulation is not unexpected.

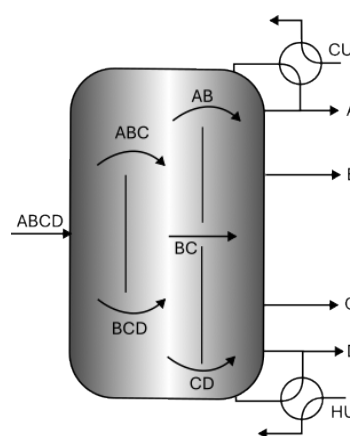
The data presented in Table 1 clearly indicates that the shortcut model demonstrates commendable performance for both mixtures. This suggests that the model is proficient in simulating the separation of components with moderate temperature differences. However, based on our experience, as the boiling points of the mixture components diverge significantly, the difference between the calculated energy consumption values from the proposed model and the ASPEN simulation tends to increase. Nevertheless, when comparing different configurations, the model accurately preserves the relative benefits in them.

In the next section, we demonstrate the efficacy of the model using a toy 4-component alcohol separation problem.

## CASE STUDY – ALCOHOL SEPARATION

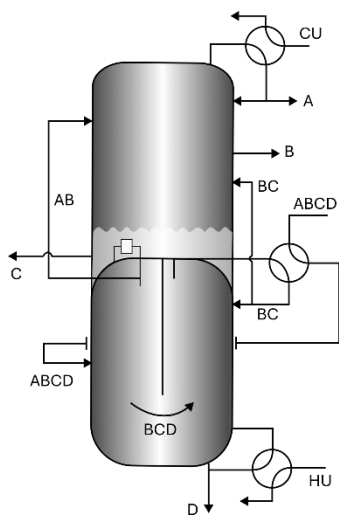
We proceed to evaluate the framework on a toy multi-component (ABCD) alcohol mixture separation problem. The saturated liquid feed comprises 35% methanol (A), 35% ethanol (B), 20% propanol (C), and 10% butanol (D). The problem is reformulated as a Mixed-Integer Quadratically Constrained Programming (MIQCP) problem and solved using Gurobi 9.1.

Figure 4 illustrates the configuration with the lowest effective fuel consumption, achieved without employing heat integration or a heat pump (HI/HP). This configuration is referred to as a fully thermally coupled system (FTC). The calculated effective fuel consumption using the shortcut model for this configuration is 56.7 MJ/kmol feed.



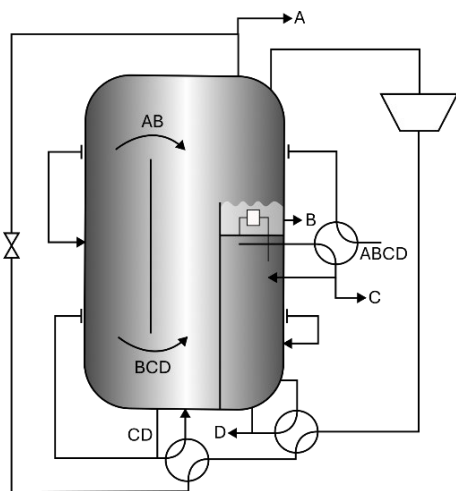
**Figure 4.** Lowest energy consuming configuration that doesn't involve heat integration or heat pumps. This configuration is the dividing wall column version of fully thermally coupled (FTC) system.

Subsequently, we introduce heat integration (HI) links in the model, allowing for varying column pressures while still utilizing steam as the energy source. Figure 5 depicts the optimal heat-integrated configuration, resulting in a reduced effective fuel consumption of approximately 17.5% to 46.7 MJ/kmol feed. The dividing wall column versions of multi-component configurations can be drawn using the method proposed by Ramapriya et al. [10,11]. Similar to the fully thermally coupled system (FTC), the proposed optimal configuration employs steam at the highest temperature, specifically in reboiler D. The divided wall column beneath operates at a higher pressure, facilitating heat transfer to the column above. It is noteworthy that the condensation of stream AB at a higher pressure provides heat to reboiler C at a lower pressure. Such integrations remain unexplored if column pressure variation is not considered and demonstrates the efficacy of our model.



**Figure 5.** Lowest energy consuming configuration involving heat integration but not heat pumps.

Furthermore, the problem is constrained to include only one heat pump compressor. Figure 6 illustrates one of the top configurations that minimizes effective fuel consumption to 21.4 MJ/kmol feed. The fuel consumption of this is approximately 63% less compared to FTC and 54% lower as compared to the heat integrated configuration. Two notable features of the presented configuration are: (a) The configuration does not rely on any steam utility, and (b) The entire separation unit operates using only a single compressor. Given that compressors contribute significantly to the capital expenditures (CAPEX) of the system, the use of a single compressor results in reduced capital costs.



**Figure 6.** One of the optimal configurations to separate the 4-comp alcohol mixture involving heat integration and heat pumps. The configuration involves a single compressor as the energy input.

## CONCLUSIONS

In the era of decarbonization, it is crucial to evaluate whether the proposed electrification of unit operations genuinely reduces fuel consumption or merely shifts the demand for fuel from the chemical industry to power plants. In this article, we develop a systematic framework to evaluate the aforementioned metric for multi-component distillation systems.

The framework relies on different shortcut and surrogate models which capture the thermodynamics of the system efficiently. Through a couple of binary distillation examples, we establish the accuracy of the model in comparing different alternatives and reduction in fuel consumption. Finally, using a toy multicomponent problem, we demonstrate the efficacy of the proposed framework to identify configurations consuming significantly lower fuel than traditional steam heated distillation units.

For most cases, HP assisted configurations can help mitigate carbon emissions in the current economy. A further reduction in energy consumption can be brought about by using intermediate reboilers/condensers, Multi-effect distillation etc which can be explored in the context of multi-component distillation systems.

## TABLE OF ACRONYMS

Acronym	Full Form
HI	Heat Integration
HP	Heat Pump
HPAD	Heat Pump Assisted Distillation
CAPEX	Capital Expenditure
FTC	Fully Thermally Coupled Column
MINLP	Mixed Integer Non-Linear Programming
CMO	Constant molar Overflow
CHT	Constant Heat Transport

## ACKNOWLEDGEMENTS

We thank the National Science Foundation under Cooperative Agreement No. EEC-1647722 for funding. We also thank Dr. Radhakrishna Tumbalam Gooty and Dr. Tony Mathew for their discussions on the work presented in this paper.

## REFERENCES

1. Rakesh Agrawal. Synthesis of multicomponent distillation column configurations. *AIChE Journal* 49, 2: 379–401 (2003)
2. Rakesh Agrawal and Jeffrey J Sirola. Decarbonization of Chemical Process Industries via Electrification. *The Bridge*, 32–40. (2023)
3. Rakesh Agrawal and Terrence F. Yee. Heat Pumps for Thermally Linked Distillation Columns: An

- Exercise for Argon Production from Air. *Industrial & Engineering Chemistry Research* 33, 11: 2717–2730 (1994)
4. M. J. Andrecovich and A. W. Westerberg. An MILP formulation for heat-integrated distillation sequence synthesis. *AIChE Journal* 31, 9: 1461–1474 (1985)
  5. José A. Caballero and Ignacio E. Grossmann. Design of distillation sequences: From conventional to fully thermally coupled distillation systems. *Computers and Chemical Engineering* 28, 11 (2004)
  6. Richard B. Chapas and Jeffery A. Colwell. Industrial Technologies Program Research Plan for Energy-Intensive Process Industries. (2007)
  7. Radhakrishna Tumbalam Gooty, Rakesh Agrawal, and Mohit Tawarmalani. Advances in MINLP to Identify Energy-efficient Distillation Configurations. (2020)
  8. Radhakrishna Tumbalam Gooty, Rakesh Agrawal, and Mohit Tawarmalani. Advances in MINLP to Identify Energy-Efficient Distillation Configurations. *Operations Research* (2022)
  9. Jimmy L. Humphrey and George E. Keller II. Separation Process Technology: Performance, Selection, Scaleup. *McGraw Hill* (1997)
  10. Gautham Madenoor Ramapriya, Mohit Tawarmalani, and Rakesh Agrawal. A systematic method to synthesize all dividing wall columns for  $n$ -component separation—Part I. *AIChE Journal* 64, 2: 649–659 (2018)
  11. Gautham Madenoor Ramapriya, Mohit Tawarmalani, and Rakesh Agrawal. A systematic method to synthesize all dividing wall columns for  $n$ -component separation: Part II. *AIChE Journal* 64, 2: 660–672 (2018)
  12. Tony Joseph Mathew, Mohit Tawarmalani, and Rakesh Agrawal. Relaxing the constant molar overflow assumption in distillation optimization. *AIChE Journal* 69, 9 (2023)
  13. Tony Joseph Mathew, Radhakrishna Tumbalam Gooty, Mohit Tawarmalani, and Rakesh Agrawal. 110th Anniversary: Thermal Coupling via Heat Transfer: A Potential Route to Simple Distillation Configurations with Lower Heat Duty. *Industrial and Engineering Chemistry Research* 58, 47: 21671–21678 (2019)
  14. Akash Sanjay Nogaja, Tony Joseph Mathew, Mohit Tawarmalani, and Rakesh Agrawal. Identifying Heat-Integrated Energy-Efficient Multicomponent Distillation Configurations. *Industrial and Engineering Chemistry Research* 61, 37 (2022)
  15. H.R. Null. Heat Pumps in Distillation. *Chem. Eng. Prog.* 58: 58–64 (1976)
  16. Gautham Madenoor Ramapriya, Ajiththaa Selvarajah, Luis Eduardo Jimenez Cucaita, Joshua Huff, Mohit Tawarmalani, and Rakesh Agrawal. Short-Cut Methods versus Rigorous Methods for Performance-Evaluation of Distillation Configurations. *Industrial and Engineering Chemistry Research* 57, 22: 7726–7731 (2018)
  17. Vishesh H. Shah and Rakesh Agrawal. A matrix method for multicomponent distillation sequences. *AIChE Journal* 56, 7: 1759–1775 (2010)
  18. Radhakrishna Tumbalam Gooty, Rakesh Agrawal, and Mohit Tawarmalani. An MINLP formulation for the optimization of multicomponent distillation configurations. *Computers and Chemical Engineering* 125: 13–30 (2019)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Role of Hydrogen as Fuel in Decarbonizing US Clinker Manufacturing for Cement Production: Costs and CO<sub>2</sub> Emissions Reduction Potentials

**Ikenna J. Okeke<sup>a\*</sup>, Sachin U. Nimbalkar<sup>a</sup>, Kiran Thirumaran<sup>a</sup>, and Joe Cresko<sup>b</sup>**

<sup>a</sup> *Manufacturing Energy Efficiency Research and Analysis Group, Oak Ridge National Laboratory, Oak Ridge, TN, USA*

<sup>b</sup> *Industrial Efficiency and Decarbonization Office, US Department of Energy, SW Washington, DC, USA*

\* Corresponding Author: [okekeij@ornl.gov](mailto:okekeij@ornl.gov).

---

## ABSTRACT

As a low-carbon fuel, feedstock, and energy source, hydrogen is expected to play a vital role in the decarbonization of high-temperature process heat during the pyroprocessing steps of clinker production in cement manufacturing. However, to accurately assess its potential for reducing CO<sub>2</sub> emissions and the associated costs in clinker production applications, a techno-economic analysis and a study of facility-level CO<sub>2</sub> emissions are necessary. Assuming that up to 20% hydrogen can be blended in clinker fuel mix without significant changes in equipment configuration, this study evaluates the potential reduction in CO<sub>2</sub> emissions (scopes 1 and 2) and cost implications when replacing current carbon-intensive fuels with hydrogen. Using the direct energy substitution method, we developed an Excel-based model of clinker production, considering different hydrogen-blend scenarios. Hydrogen from steam methane reformer (gray) and renewable-based electrolysis (green) are considered as sources of hydrogen fuel for blend scenarios of 5%–20%. Metrics such as the cost of cement production, facility-level CO<sub>2</sub> emissions, and cost of CO<sub>2</sub> avoided were computed. Results show that for hydrogen blends (gray or green) between 5% and 20%, the cost of cement increases by 0.6% to 16%, with only a 0.4% to 6% reduction in CO<sub>2</sub> emissions. When the cost of CO<sub>2</sub> avoided was computed, the extra cost required to reduce CO<sub>2</sub> emissions is \$229 to \$358/metric ton CO<sub>2</sub>. In summary, although green hydrogen shows promise as a low-carbon fuel, its adoption for decarbonizing clinker production is currently impeded by costs.

---

**Keywords:** Cement, Clinker, Hydrogen, Decarbonization, CO<sub>2</sub>.

## INTRODUCTION

In the United States, cement (and lime) production in 2022 is responsible for 31 million metric tons (MMT) CO<sub>2</sub> emissions of the industry sector's total CO<sub>2</sub> emissions [1]. These GHG emissions attributed to the cement sector are directly associated with clinker manufacturing, which constitutes 70%–90% of conventional cement blends. With clinker production volume of 79 MMT in 2022 [2], if the status quo is maintained, reduction of these emissions is vital if we intend to achieve the 1.5°C target.

Because clinker manufacturing involves the breakdown of limestone to lime and the concurrent production of byproduct CO<sub>2</sub> in a high-temperature pyroprocessing

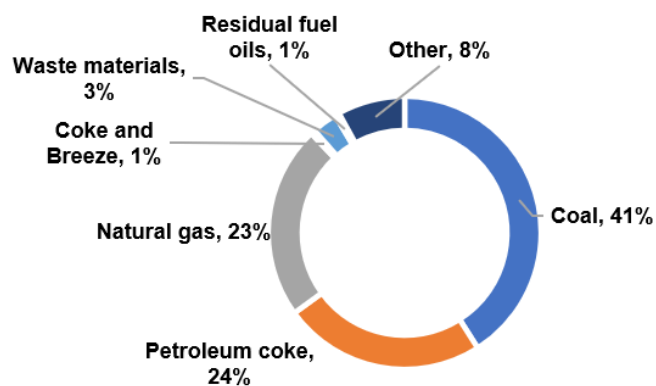
step, the generation of process- and combustion-related CO<sub>2</sub> emissions is currently inevitable. Owing to the chemistry of the reaction, process-related CO<sub>2</sub> emissions account for up to 60% of total clinker production CO<sub>2</sub> emissions with the remaining associated with energy use. Given the twofold sources of CO<sub>2</sub> emissions in cement production, efforts are channeled toward decarbonizing one or both.

To decarbonize the cement industry, implementation of the crosscutting decarbonization pillars—namely energy efficiency; low-carbon fuels, feedstocks, and energy sources (LCFFES); industrial electrification; and carbon capture, utilization, and storage (CCUS)—are necessary conditions to attain a low-carbon cement industry. Hence, over the years, different research efforts have



been focused on these carbon-reducing strategies, technologies, and practices.

For example, Worrell et al. [3] conducted an energy efficiency analysis using 30 energy-efficient technologies and measures in the US cement sector. Results of the analysis showed energy and CO<sub>2</sub> savings of 11% and 5%, respectively. Nevertheless, current typical energy use in the cement sector still revolves between 3.5 and 4.1 GJ/MT [4–6] and with emissions intensity as high as 900 kgCO<sub>2</sub>/MT cement [7]. Advanced waste heat recovery and process intensification are still needed for the contribution of energy efficiency pillars to be significant. For LCFES applications, the use of supplementary cementitious materials such as blast furnace slag, volcanic ash, pozzolans, fly ash, and calcined clay can reduce the clinker-to-cement ratio to about 65%–75% with the innovative limestone calcined clay allowing higher clinker substitution up to 50% [8]—thereby reducing energy use and emissions typically associated with clinker production. However, a drawback to using slag and fly ash is the transition from blast furnace and coal power plants to more sustainable alternatives [9]. Industrial electrification research has focused on the electricity-driven calciner or calciner/kiln system [4–5]—offsetting on-site scope 1 CO<sub>2</sub> emissions emanating from fuel combustion and enabling a CO<sub>2</sub>-rich flue gas (pyroprocessing step), which minimizes the energy requirement for carbon capture (CC). An innovative electrified cement production in progress is the Leilac project designed to indirectly heat raw meal in a shell and tube system, efficiently separating the process CO<sub>2</sub> for direct capture or use [10]. Finally, CCUS is primed as the sole decarbonization pillar that can significantly reduce CO<sub>2</sub> intensity of the cement sector, depending on the technology implemented. However, the most feasible cement CC technologies are still at a low technology readiness level (TRL), and the high TRL technologies are very energy intensive [11], which the current cement facility is not equipped to support.



**Figure 1.** Average cement fuel mix [12].

Among all decarbonization pillars, fuel switching has received little or no attention. Traditionally, US cement

facilities use a range of fuel mix, as depicted in Figure 1. Coal and petroleum coke with high-carbon content account for 65% of this fuel mix, dominating the energy-related emissions. In practice, cement facilities usually combust any cheap fuel in the market or even receive a tipping fee to dispose of industrial wastes in the kiln [13]. To decarbonize the cement sector, broad commitment toward clean or low-carbon fuels is needed.

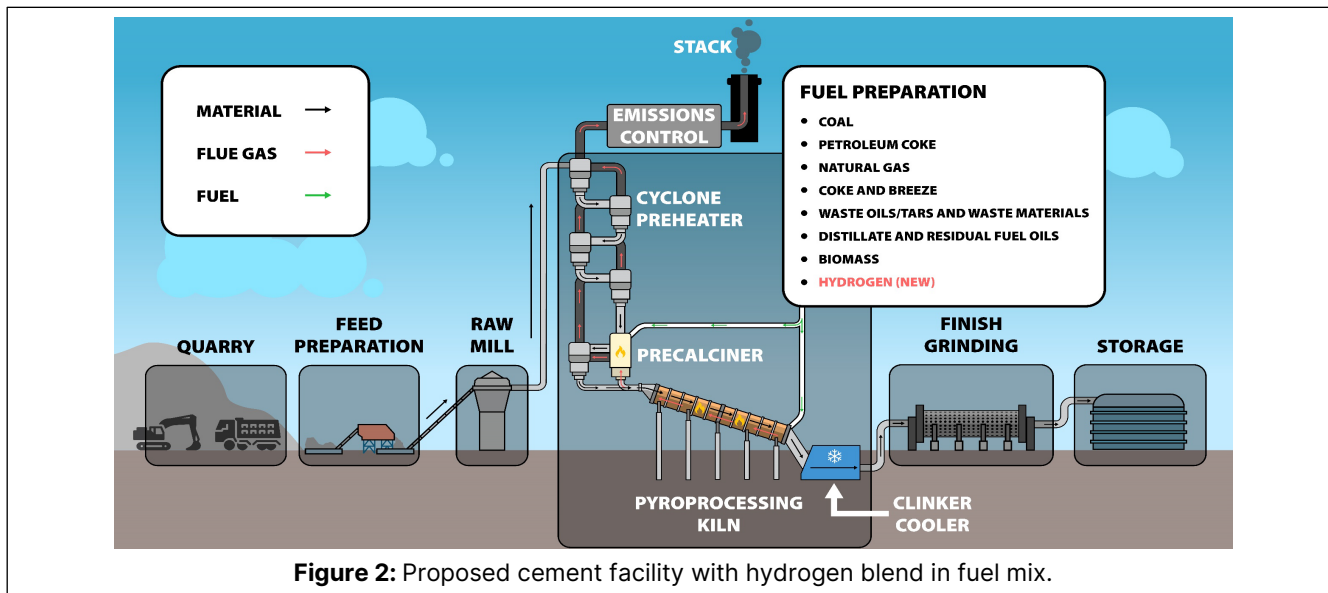
Within the hydrogen economy, hydrogen as a carbon-free energy carrier is being pushed as a fuel that can decarbonize the cement industry sector. For instance, CEMEX has announced the use of hydrogen as a share in cement fuel mix for all cement facilities in Europe to attain zero CO<sub>2</sub> combustion emissions and improve energy efficiency [14]. Nevertheless, whereas hydrogen seems to be a viable means to decarbonize cement production, the technical challenges, costs, and potential emissions reduction that could impede the full adoption of hydrogen as a fuel need to be evaluated.

Hence, in this study, we assessed the cost and potential CO<sub>2</sub> emissions reduction of blending hydrogen as a fuel in cement production. This analysis explored the use of hydrogen from both conventional steam methane reformer (SMR) hydrogen (gray) and renewable-powered electrolysis hydrogen (green). In this context, scope 2 emissions of hydrogen production were considered. In addition, discussions on the near-term technical challenges currently impeding hydrogen use in the industry sector are highlighted. Performance metrics, such as facility-level CO<sub>2</sub> emissions, minimum cement selling price (MCSP), and cost of CO<sub>2</sub> avoided (CCA), were computed. Overall, insights on the current state of hydrogen adoption as a fuel in the cement sector are elucidated.

## METHODOLOGY

### Process modeling and assumptions

This work employs a brown-field cement facility with a five-stage cyclone preheater with precalciner system for dry portland cement manufacturing. Clinker production is sized to be about 1.2 MMT (clinker to cement ratio = 81%) corresponding to an average US cement facility. In this proposed facility, hydrogen is assumed to substitute 5%–20% of the primary energy demand. At the time of this analysis, it is assumed that a hydrogen blend of up to 20% would not necessitate any significant modification of the equipment [24] used in the cement facility. The authors developed an Excel-based model for the energy demand of hydrogen-fired cement production using a direct energy substation method. Hydrogen supply to the cement facility is assumed to be met via SMR hydrogen (gray) and electrolysis hydrogen (green)—with associated scope 2 emissions included. For comparison, the conventional cement (based on traditional fuel mix) was also modeled. Intrinsicly, it is assumed that the cement



**Figure 2:** Proposed cement facility with hydrogen blend in fuel mix.

produced in each scenario meets the portland cement standard. Key parameters and assumptions used in the model are shown in Table 1.

Figure 2 shows the proposed cement production facility. As a decarbonization strategy, different hydrogen blend ratios are designed to substitute carbon-intensive fuels such as coal and petroleum coke while keeping natural gas and biomass fuels constant. Overall, the performance of the cement facility with different hydrogen sources at different blend ratios was assessed in terms of MCSP, facility-level CO<sub>2</sub> emissions (scopes 1 and 2), and CCA.

**Table 1:** Parameters used in the model.

Parameter	Value	Ref
<i>Energy demand (GJ/t)</i>		
Feed preparation	0.13	[6]
Fuel preparation	0.13	[6]
Pyroprocessing (with cooling)	3.61	[6]
Finish grinding	0.21	[6]
Primary thermal energy	3.51	[6]
Electricity (site)	0.19	[6]
Raw meal to clinker ratio	1.6	[15]

### Process description

Portland cement manufacturing involves the general steps of mining or extraction of feedstock, feedstock preparation via crushing and grinding operations, fuels preparation, the pyroprocessing operation (preheater/precaliner), kiln with cooling system, and final grinding with additives, as shown in Figure 2.

Raw materials for cement production such as limestone, clay, shale, marl, and iron ore are mined or extracted at quarries typically located near the cement facility. Additional raw material such as gypsum needed in the finishing grind operation is also sourced. Cement

facilities currently use waste materials such as fly ash and blast furnace slag, etc. as raw materials [16].

To ensure the homogeneity of the mixture and that the appropriate chemical and physical properties are achieved, these raw materials are crushed into smaller particles using hammer or jaw crushers [6]. The crushed particles are further processed by grinding to form raw meal, which is sent to the preheater [6]. This raw material processing takes place in the feed preparation unit which is powered primarily by electricity.

Owing to the diverse nature of fuel mixes (solids, liquids, and gases) used in cement production, fuel preparation is a standard process. Conventional fuels like coal, petroleum coke, and waste tire are crushed, creating fine particles that improve combustion. These fuels are then combusted in the calciner and kiln to drive the calcination and sintering reactions.

Raw meal from the grinder is sent to the pyroprocessing step, where the moisture content is reduced in the cyclone preheater system before the meal is sent to the precaliner. In the precaliner, the initial calcination reaction takes place at about 900°C, driving off CO<sub>2</sub> from the raw meal and improving the overall thermal efficiency of the process. Clinker formation occurs in the rotary kiln, which operates at 1,450°C to produce hot solid clinker products.

The final cement production step involves air cooling and subsequent grinding of the clinker into a fine powder, which is mixed with gypsum and other additives to form cement.

### Facility-level CO<sub>2</sub> emissions assessment

In this analysis, the facility-level CO<sub>2</sub> emissions are based on scopes 1 and 2 emissions emanating from on-site (process and combustion) and upstream (electricity and hydrogen) emissions. Other on-site CO<sub>2</sub> emissions such as mobile equipment use are not considered. The

facility-level CO<sub>2</sub> emissions are computed as:

$$CO_{2t} = CO_{2p} + CO_{2f} + CO_{2e} \quad (1)$$

$$CO_{2f} = \sum_i CO_{2i} \quad (2)$$

where  $CO_{2t}$  is the total facility-level CO<sub>2</sub> emissions,  $CO_{2p}$  is the process-related CO<sub>2</sub> emissions,  $CO_{2f}$  is the fuel combustion CO<sub>2</sub> emissions,  $CO_{2i}$  accounts for the combustion-related CO<sub>2</sub> emissions of  $i^{th}$  fuel, and  $CO_{2e}$  is the electricity CO<sub>2</sub> intensity.

For the fuel-related emissions, the waste biomass used as fuel is assumed to be sustainable; hence, biomass combustion CO<sub>2</sub> equals zero. Whereas the fuel emissions intensity is obtained from EPA's emission factors for GHG inventories [17], we used the following emissions factors: for electricity, 0.4 kg/kWh; for conventional SMR hydrogen, 10 kgCO<sub>2</sub>e/kg H<sub>2</sub>; and for electrolysis green hydrogen, 0.97 kgCO<sub>2</sub>e/kg H<sub>2</sub> [18].

### Cost assessment

Cost performance of the proposed cement manufacturing is conducted to assess the economic viability of this decarbonization strategy when compared with the conventional cement pathway. The direct equipment cost (DEC) of the cement facility is estimated based on [4] using the sixth-tenth rule to adjust for plant capacity.

The equipment cost is presented in \$US2020 using the Chemical Engineering Plant Cost Index [19]. Based on the computed DEC, the total capital investment is calculated using correlations from [20,21]. The operating costs (variable and fixed) were estimated as a function of plant-size consumables (feedstock, fuel, electricity), labor, maintenance, operating overhead, and property insurance costs [4,21]. The MCSP is computed by conducting a discounted cash flow rate of return over the assumed lifetime of the plant using the economic parameters presented in Table 2.

Parameter	Value	Parameter	Value
Plant life (years)	30	Tax rate (%)	40
Plant avail. (%)	75	Rate of return (%)	12
Plant loan (years)	10	Depreciation	MACRS
Loan interest (%)	9.5	Working capital	15%
Debit/equity (%)	50	Operating hours	8,000
<b>Fuel costs (\$/GJ)</b>			
Coal	2.01	Distillate fuel oils	9.70
Petroleum coke	2.69	Biomass	3.34
Natural gas	2.93	SMR H <sub>2</sub>	7.09
Waste oils/tars	0.01	Electrolysis H <sub>2</sub>	31.91
Coke and breeze	2.02		

### Cost of CO<sub>2</sub> avoided

We computed the CCA to ascertain the additional cost incurred to offset a unit amount of CO<sub>2</sub> in cement manufacturing. The CCA formulation is given in Eq. (2) based on [22,23]:

$$CCA = \frac{MCSP_s - MCSP_c}{CO_{2,c} - CO_{2,s}} \quad (2)$$

where  $MCSP_c$  and  $MCSP_s$  are the minimum cement selling price of the conventional and scenario hydrogen fuel fired cement in \$/MT cement, respectively; and  $CO_{2,c}$  and  $CO_{2,s}$  are the CO<sub>2</sub> emissions of both the conventional cement and scenario hydrogen fuel blended cement in kgCO<sub>2</sub>eq/MT cement, respectively.

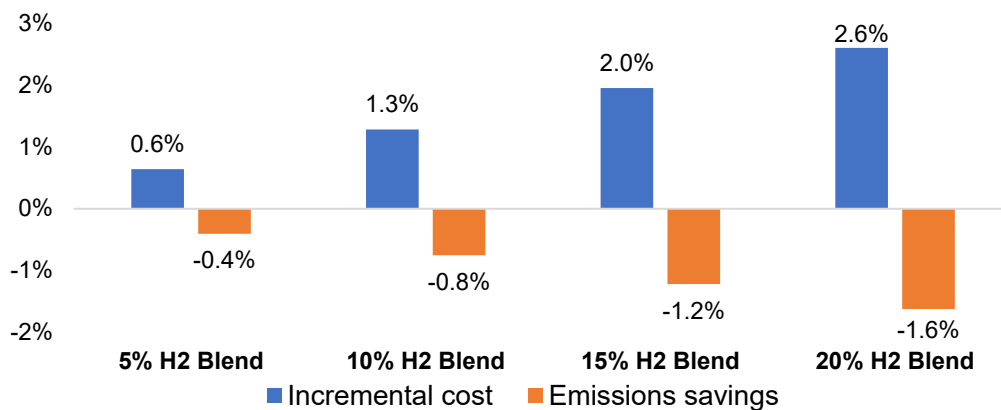
## RESULTS

### Systems performance: Gray hydrogen blend

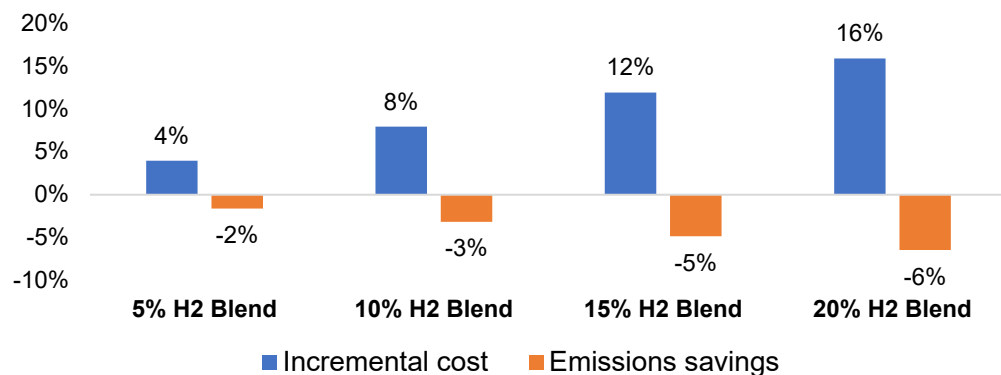
Table 3 shows the MCSP, facility-level CO<sub>2</sub> emissions, and CCA for different blends of gray hydrogen ratios when compared with the conventional clinker. The MCSP for conventional cement is \$106.9/MT, whereas a blend of hydrogen between 5% and 20% changes the

**Table 2:** Economic parameters and assumptions.

<b>Table 3:</b> Costs and performance analysis of SMR hydrogen blend ratios with conventional fuels						
Metric	Unit	Conventional	5% H <sub>2</sub>	10% H <sub>2</sub>	15% H <sub>2</sub>	20% H <sub>2</sub>
<b>Cost</b>						
Fuel	M\$/year	7.02	7.71	8.39	9.09	9.76
Electricity	M\$/year	4.03	4.03	4.03	4.03	4.03
Capital	M\$	369.3	369.3	369.3	369.3	369.3
Fixed O&M	M\$/year	36.42	36.42	36.42	36.42	36.42
Variable O&M	M\$/year	7.22	7.22	7.22	7.22	7.22
MCSP	\$/MT	106.9	107.6	108.2	108.9	109.7
<b>Emissions</b>						
Combustion CO <sub>2</sub>	kgCO <sub>2</sub> /MT	230	227	225	221	218
Process CO <sub>2</sub>	kgCO <sub>2</sub> /MT	518	518	518	518	518
Total CO <sub>2</sub>	kgCO <sub>2</sub> /MT	749	746	743	739	736
Cost of CO <sub>2</sub> avoided	\$/MT CO <sub>2</sub>	-	228.9	246.9	228.9	228.9



**Figure 3:** Percentage change in cost and emissions of cement for SMR hydrogen blends.



**Figure 4:** Percentage change in cost and emissions of cement for green hydrogen blends.

MCSP to \$107.6–\$109.6/MT. This represents an increase of up to \$2.8/MT cement. As can be observed (Table 1), the change in cement cost affects only the fuel cost owing to hydrogen use. Other cost components such as electricity (power), capital, fixed operating & maintenance (O&M), and variable O&M were assumed to remain the same. This assumption is feasible because hydrogen blends up to 20% are currently stipulated as the maximum limit before major infrastructure change is required [24].

In terms of emissions reduction, whereas the process emissions remain unchanged when hydrogen is blended in the fuel mix (Table 1), energy-related CO<sub>2</sub> emissions change depending on the amount of hydrogen blended. Hence, the total scopes 1 and 2 CO<sub>2</sub> emissions for the 5%–20% H<sub>2</sub> blends ranged between 736.4 and 745.5 kgCO<sub>2</sub>/MT cement. If only the energy-related CO<sub>2</sub> emissions are considered, a 5% maximum CO<sub>2</sub> emissions reduction is achieved. This extent of emissions reduction is due to the scope 2 emissions associated with hydrogen production. Hypothetically, only when the scope 2 emissions of SMR hydrogen are zero can a significant reduction (as high as 23%) in energy-related CO<sub>2</sub> emission be achieved.

The CCA—a function of the additional cost incurred

to avoid CO<sub>2</sub> emission—ranged between \$228.9 and \$246.9/MT CO<sub>2</sub>. An interesting observation in the computed CCA is the increase in CCA from \$228.9/MT (for 5% H<sub>2</sub>) to \$246.9/MT (for 10% H<sub>2</sub>), followed by a decrease back to \$228.9/MT (for 15% and 20% H<sub>2</sub> blends, respectively). This phenomenon is linked to the extent of emissions offset when compared with the extra cost incurred to achieve such an extent of reduction.

Figure 3 depicts the comparative emissions savings of an H<sub>2</sub> blend with the corresponding change in MCSP necessary to achieve the CO<sub>2</sub> reduction. Blending 5% H<sub>2</sub> in the clinker fuel mix had only a 0.4% reduction in total facility-level CO<sub>2</sub> emissions while increasing the cost by 0.6%. When 10% H<sub>2</sub> is blended, only 0.8% reduction in CO<sub>2</sub> emissions is achieved—increasing the cost of cement by 1.3%. At 20% H<sub>2</sub> blend, emissions reduction is still below 2%, whereas the clinker cost increases by 2.6%. One might wonder why the emissions reduction is not significant for a zero-carbon fuel like hydrogen. We considered indirect CO<sub>2</sub> emissions (scope 2) associated with hydrogen production, which, on a mass basis, are higher than the fossil fuels for which they substitute.

### Systems performance: Green hydrogen blend

When green hydrogen is considered as the source

**Table 4:** Costs and performance analysis of green hydrogen blend ratios with conventional fuels

Metric	Unit	Conventional	5% H <sub>2</sub>	10% H <sub>2</sub>	15% H <sub>2</sub>	20% H <sub>2</sub>
<i>Cost</i>						
Fuel	M\$/year	7.02	13.8	19.0	24.2	29.4
Electricity	M\$/year	4.03	4.03	4.03	4.03	4.03
Capital	M\$	369.3	369.3	369.3	369.3	369.3
Fixed O&M	M\$/year	36.42	36.42	36.42	36.42	36.42
Variable O&M	M\$/year	7.22	7.22	7.22	7.22	7.22
MCSP	\$/MT	106.9	111.1	115.4	119.6	123.9
<i>Emissions</i>						
Combustion CO <sub>2</sub>	kgCO <sub>2</sub> /MT	230	218	206	194	182
Process CO <sub>2</sub>	kgCO <sub>2</sub> /MT	518	518	518	518	518
Total CO <sub>2</sub>	kgCO <sub>2</sub> /MT	749	736	725	712	700
Cost of CO <sub>2</sub> avoided	\$/MT CO <sub>2</sub>	–	351.7	358.3	351.7	351.7

of hydrogen in the fuel mix, cost and emissions performance differs, as shown in Table 4. In terms of cost, H<sub>2</sub> blends between 5% and 20% yield an MCSP of \$111.1–\$123.9/MT with corresponding CO<sub>2</sub> emissions of 700–736 kgCO<sub>2</sub>/MT cement. Comparing the emission reduction achieved using green hydrogen to that of gray hydrogen, additional 5% and 16% improvements in facility-level and energy-related CO<sub>2</sub> emissions are achieved, respectively. For a 5% reduction in facility-level emissions, up to a 13% increase in MCSP is observed for green hydrogen-fired cement compared with the gray hydrogen counterpart. Clearly, this premium for decarbonized cement might be difficult for cement manufacturers to compete in the market, given the extent of reduction.

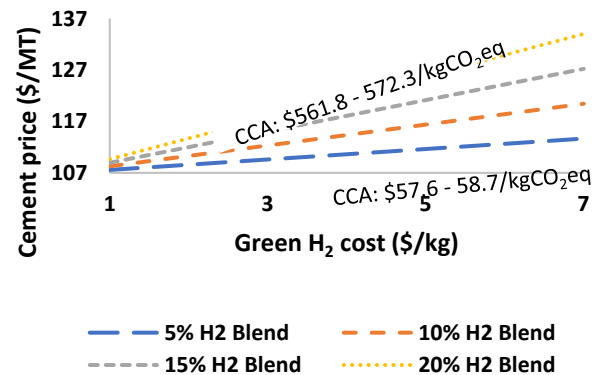
It is not surprising that the CCA for the use of green hydrogen ranges between \$351.7 and \$358.3/MT CO<sub>2</sub>. Again, this computed CCA is more expensive than the cost of carbon capture reported for the cement industry (\$50–\$60/MT) [25]. This implies that it is economically and even environmentally competitive to adopt and use CC in cement rather than use hydrogen as a fuel for decarbonization purposes.

Finally, Figure 4 directly compares the emissions reductions achieved and the associated costs for different H<sub>2</sub> blend ratios relative to conventional cement. For a 2% reduction in CO<sub>2</sub> emissions, a 4% increase in MCSP was observed for a 5% H<sub>2</sub> blend. A maximum of 6% reduction in facility-level CO<sub>2</sub> can be achieved when 20% H<sub>2</sub> is blended in a clinker fuel mix—at the expense of a 16% increase in MCSP. Given the minimal reduction in emissions at a high cost, at present it does not make economic sense to use hydrogen as a decarbonization lever in the cement industry.

### Sensitivity analysis

Because green hydrogen blend in cement fuel mix offers more CO<sub>2</sub> emissions reduction when compared with gray hydrogen, the effect of green hydrogen's cost on MSCP is assessed via sensitivity analysis. Figure 5

shows the effect of green hydrogen's cost on MCSP for different hydrogen blend ratios.



**Figure 5:** Sensitivity analysis of the effect of green hydrogen cost on cement price.

At a green hydrogen price of \$1/kg owing to the federal Infrastructure Reduction Act 45 V tax credit that incentivizes green hydrogen products, the MCSP is between \$107.6 and \$109.7/MT for 5%–20% hydrogen blend ratios. Even at \$1/kg for green hydrogen, the selling price of cement is still \$0.70/MT higher than that of conventional cement. The CCA of the cement at this hydrogen price is between \$57.6 and \$58.7/MT CO<sub>2</sub> avoided, which is competitive with that achieved via CCS. To a greater extent, because CCS implementation at a specific cement facility can be impeded by a lack of CO<sub>2</sub> transportation and sequestration sites, use of green hydrogen proves to be more suitable. Nevertheless, on an emissions intensity basis, a CCS-enabled cement facility at 85% capture (112.3 kgCO<sub>2</sub>/MT) outperforms the green hydrogen-fired (700 kgCO<sub>2</sub>/MT) cement facility.

At a \$7/kg green hydrogen price, the cost of cement increases between \$113.7 and \$134/MT with a corresponding CCA of \$562–\$572/kgCO<sub>2</sub>eq avoided. Overall, the high cost of hydrogen when compared with the extent of emissions reduction limits its use as a



decarbonization strategy in cement production.

## DISCUSSIONS

Decarbonizing cement production is vital to attaining US economy-wide GHG reduction goals. With an interest in decarbonizing industrial process heat, cement manufacturers are beginning to propose the adoption of hydrogen as fuel. Argos and CEMEX are pushing toward a hydrogen share in their cement fuel mix for facilities in Honduras [28] and Europe [14], respectively. However, this analysis shows that pursuing this route as a decarbonization strategy is highly limited by cost when compared with the amount of reduction attained. At best, when green hydrogen (with zero emissions) constitutes 100% of fuel used in cement production (not shown in this analysis), only a 29% reduction in scope 1 emissions is achieved. Nevertheless, when scope 2 emissions of hydrogen and electricity are considered, the reduction amount drops to 26%. Therefore, it is evident that hydrogen should be channeled to other industrial applications.

Despite the emissions- and cost-related challenges, supply chain and other technical barriers still impede the widespread adoption of hydrogen as fuel (blend mix or 100% H<sub>2</sub>). At 100% use of hydrogen in the cement sector, the demand for hydrogen per annum will be about 2 MMT (based on 2022 clinker production of 78 MMT). Because the current annual US hydrogen production is about 10 MMT, a supply-demand issue might arise. In addition, significant reconfiguration of existing cement facilities with high CAPEX will be required [26].

Technical challenges associated with hydrogen include storage owing to low volumetric energy density, increasing tendency of fire or explosion because of its wide flammability range in air, burning challenges linked to its burning velocity that will require complex control systems, high NO<sub>x</sub> formation because it has a high adiabatic flame temperature (burner modification required, with associated costs), and low thermal radiative heat transfer owing to the absence of soot exhibited by hydrogen flames [27] that affect clinker quality. Other technical challenges include hydrogen leakage and equipment/pipeline embrittlement, etc. Hydrogen leaks appear to be very detrimental because hydrogen tends to react with hydroxyl radicals, reducing the amount of OH needed to break down methane in the atmosphere [29].

Overall, more research and development are needed in the efforts to adopt hydrogen as fuel in the cement industry and other sectors in which hydrogen has the potential to contribute to decarbonization goals. In addition, government funding and policies will play a key role in reducing the cost of hydrogen (green) for industries to begin this much-needed shift.

## CONCLUSION

Clean hydrogen is a low-carbon fuel, feedstock, and energy source that can play a key role in the decarbonization goals of the US economy. Nevertheless, the use of hydrogen as a fuel has inherent supply chain technical challenges that must be overcome before it can safely and easily be adopted as a fuel mix. Even though these challenges are addressed, the cost implications could prove to be prohibitive for industry—particularly, the cement industry, with low margins—to adopt hydrogen as fuel. At 20% green H<sub>2</sub> blend, a CO<sub>2</sub> emissions reduction of 6% can be achieved at a significant 16% increase in cement cost. Therefore, it might not be economically viable for the cement industry to adopt hydrogen as a fuel in clinker manufacturing.

As observed, the CCA did not change for the hydrogen blend ratios except for 10%. The obvious explanation for this impact is that blending hydrogen as fuel (within this ratio) in clinker production does not really affect carbon emissions reduction, given the significant contribution to the overall emissions by the pyroprocessing step. Because there is an additional 8% increase in CCA for the 10% hydrogen blend, any cement facility considering blending hydrogen in the fuel mix is better off with any of the other blend scenarios because the cost increment is commensurate with the emissions offset.

## ACKNOWLEDGEMENTS

This research is funded by DOE Industrial Efficiency and Decarbonization Office Strategic Analysis.

## NOTICE OF COPYRIGHT

This manuscript has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

## REFERENCES

1. U.S. EIA Annual Energy Outlook 2023: Energy-Related Carbon Dioxide Emissions by End Use, [https://www.eia.gov/outlooks/aeo/tables\\_ref.php](https://www.eia.gov/outlooks/aeo/tables_ref.php). Accessed Nov. 17, 2023.
2. USGS Minerals commodities summaries 2023. <https://pubs.usgs.gov/periodicals/mcs2023/mcs2023-cement.pdf>. Accessed Nov. 17, 2023.

3. Ernst Worrell, Christina Galitsky, Lynn Price. Energy efficiency improvement opportunities for the cement industry (2008).
4. Bodil Wilhelmsson, Claes Kollberg, Johan Larsson, Jan Eriksson, and Magnus Eriksson. CemZero – A feasibility study evaluating ways to reach sustainable cement production via the use of electricity. Vattenfall Cementa (2018).
5. Jacob, R.M. and Tokheim, L.A. Electrified calciner concept for CO<sub>2</sub> capture in pyro-processing of a dry process cement plant. *Energy* 268 (2023): 126673.
6. U.S. DOE. Bandwidth study on energy use and potential energy savings opportunities in U.S. cement manufacturing. *Office of Energy Efficiency and Renewable Energy* (2017).
7. Portland Cement Association. Carbon footprint.
8. Scrivener, K., Martirena, F., Bishnoi, S., Maity, S., Calcined clay limestone cements (LC3). *Cement and concrete research* 114 (2018): 49-56.
9. Nature. Cement and steel – nine steps to net zero (2022) . <https://www.nature.com/articles/d41586-022-00758-4> Accessed Nov 21, 2023.
10. Leilac Technology. Roadmap to 2050: A cost-effective path to carbon neutral industrial production 2021. <https://www.leilac.com/wp-content/uploads/2022/09/LEILAC-Roadmap.pdf>.
11. Hills, T., Leeson, D., Florin, N., and Fennell, P., Carbon capture in the cement industry: technologies, progress, and retrofitting. *Environmental Science & Technology* 50.1 (2016): 368-377.
12. U.S. DOE. Manufacturing Energy and Carbon footprint, *Industrial Efficiency & Decarbonization Office* (2018 MECS). <https://www.energy.gov/eere/iedo/manufacturing-energy-and-carbon-footprints-2018-mecs>.
13. Personal communication.
14. FuelCellsWorks. CEMEX successfully deploys hydrogen-based ground-breaking technology (2021) . <https://fuelcellworks.com/news/cemex-successfully-deploys-hydrogen-based-ground-breaking-technology/>.
15. Marceau, M., Nisbet, M.A., and Van Geem, M.G. Life cycle inventory of portland cement manufacture. *Portland Cement Association*: No. PCA R&D Serial No. 2095b (2006).
16. U.S. EPA. Portland cement manufacturing. [https://www3.epa.gov/ttn/chief/old/ap42/ch11/s06/final/c11s06\\_1995.pdf](https://www3.epa.gov/ttn/chief/old/ap42/ch11/s06/final/c11s06_1995.pdf).
17. U.S. EPA. Emission factors for greenhouse gas inventories (2023). [https://www.epa.gov/system/files/documents/2023-03/ghg\\_emission\\_factors\\_hub.pdf](https://www.epa.gov/system/files/documents/2023-03/ghg_emission_factors_hub.pdf).
18. Spath, P.L., Mann, M.K., Life cycle assessment of renewable hydrogen production via wind electrolysis. *National Renewable Energy Laboratory* (2004).
19. Chemical Engineering Plant Cost Index, <https://www.chemengonline.com/pci-home>.
20. Peters, M.S., Timmerhaus, K.D. Plant design and economics for chemical engineers.
21. Okeke, I.J, Saville, B.A., MacLean H.L. Low carbon hydrogen production in Canada via natural gas pyrolysis, *International Journal of Hydrogen Energy* 48, No. 34 (2023): 12581-12599.
22. Adams, T.A., Hoseinzade L., Madabhushi P.B., Okeke, I.J., Comparison of CO<sub>2</sub> capture approaches for fossil-based power generation: review and meta-study, *Processes* 5, no. 3 (2017): 44.
23. Okeke, I.J, Adams II, T.A., Combining petroleum coke and natural gas for efficient liquid fuels production. *Energy* 163 (2018): 426-442.
24. Quintino, F.M., Nascimento., N, Fernandes., E.C. Aspects of hydrogen and biomethane introduction in natural gas infrastructure and equipment. *Hydrogen* 2, No. 3 (2021): 301-318.
25. Global CCS Institute. Technology readiness and costs of ccs (2021). <https://www.globalccsinstitute.com/wp-content/uploads/2021/03/Technology-Readiness-and-Costs-for-CCS-2021-1.pdf>.
26. U.S. DOE. Pathways to commercial liftoff: low-carbon cement (2023). <https://liftoff.energy.gov/wp-content/uploads/2023/09/20230918-Pathways-to-Commercial-Liftoff-Cement.pdf>.
27. Yin, Y., Medwell, P.R., Gee, A.J., Foo, K.K., Dally, B.B. Fundamental insights into the effect of blending hydrogen flames with sooting biofuels. *Fuel*, (2023). **331**: p. 125618.
28. Global Cement. Update on hydrogen injection in cement plants (2022). <https://www.globalcement.com/news/item/14637-update-on-hydrogen-injection-in-cement-plants>.
29. Katherine Bourzac, Carbon-free fuels could have a dark side. *Science*. 2023. <https://www.science.org/content/article/carbon-free-fuels-could-have-dark-side>.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Comparative Techno-economic Assessment of Hydrogen Production, Storage and Refueling Pathways

Minseong Park<sup>a</sup>, Hegwon Chung<sup>a</sup>, and Jiyong Kim<sup>a\*</sup>

<sup>a</sup> Sungkyunkwan University, School of Chemical Engineering, Suwon, Gyeonggi-do, South Korea

\* Corresponding Author: [jiyongkim@skku.edu](mailto:jiyongkim@skku.edu).

---

## ABSTRACT

Hydrogen, as a clean and versatile energy carrier, holds immense promise for addressing the world's growing energy and environmental challenges. However, hydrogen-based energy systems face challenges related to efficient storage methods, energy-intensive production, refueling processes, and overall cost-effectiveness. To solve this problem, a superstructure was developed that integrates overall technologies related to hydrogen energy transportation. This study synthesizes process pathways for hydrogen energy transportation method including energy carrier production, storage, and refueling, based on the developed superstructure. The techno-economic analysis was conducted to evaluate the performance of each transportation pathway and compare it with conventional fossil fuel transportation system. Process performance criteria, including unit production cost (UPC), energy efficiency (EEF), and net CO<sub>2</sub> equivalent emissions (NCE), serve as indicators for process performance. By comparing technological pathways, we can propose the most economically and environmentally optimal energy refueling route. Additionally, sensitivity analyses were performed on various external factors, identifying influential variables in the decision-making process for hydrogen production, storage, and refueling strategies, while also elucidating technological limitations.

---

**Keywords:** Hydrogen, Process synthesis, Environment, Techno-economic analysis, Energy refueling

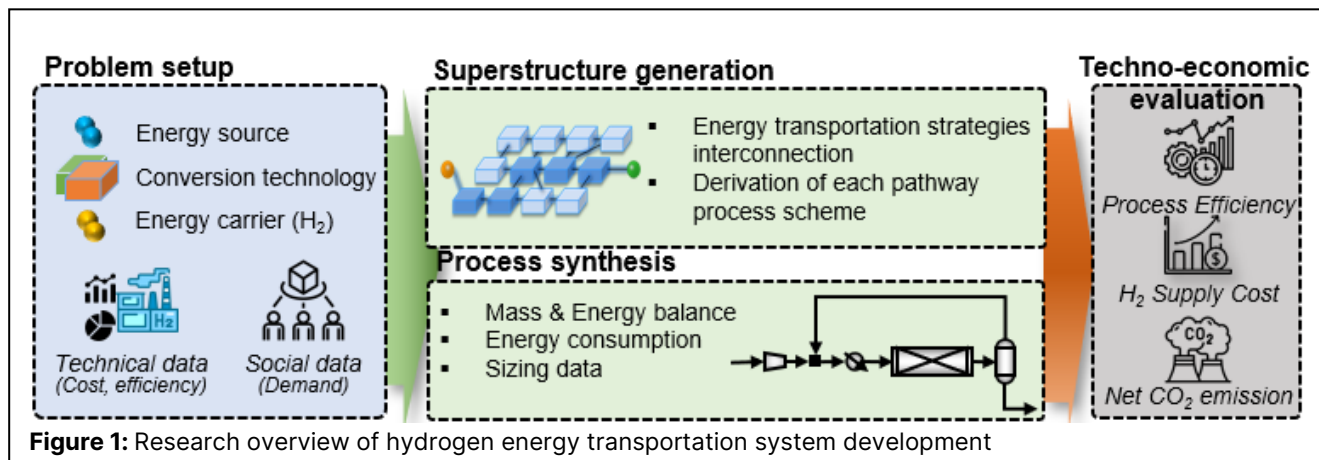
## 1. INTRODUCTION

One of the most important global environmental issues is the depletion of conventional energy sources and the climate change caused by greenhouse gas (GHG) emissions. Hydrogen is flexible energy carrier that can be converted into usable forms of energy and is an environmentally clean energy source. The concept of hydrogen economy, characterized by the production, distribution, and utilization of hydrogen as a primary energy carrier, has gained significant traction. The transformative vision envisions a departure from conventional fossil fuel dependency toward a circular and interconnected energy system where hydrogen acts as a linchpin for integration, storage, and efficient utilization of renewable energy sources.

To establish a hydrogen energy system, various preliminary studies are being conducted on developing hydrogen production, storage, and transportation systems. Li Lin et al. conducted the process development

and cost analysis of an ammonia-based hydrogen refueling station process [1]. A.D. Korberg et al. developed and evaluated a transport process for renewable energy resources [2]. Jack Shepherd et al. developed a green ammonia value chain and proposed the pathway with the highest economic feasibility [3]. Ayodeji Okunlola et al. proposed an optimal hydrogen transportation method through analysis of inter-country hydrogen transportation and supply [4].

In preliminary research, the development of technologies for the production, storage, and transportation of hydrogen is progressing, and research on establishing hydrogen energy system is also being conducted. However, the previous study has not determined and evaluated compared conventional fossil fuel transportation system. We compared and analyzed the energy system, including hydrogen production, storage, and transportation, with the conventional fossil fuel-based energy system to determine the possibility of hydrogen replacing conventional energy resources. This study focused on



various hydrogen storage and transportation method such as high-pressure compression, liquefaction, ammonia, liquid organic hydrogen carriers (LOHC). Each developed hydrogen storage and transportation method is synthesized with a process for production hydrogen from various energy sources.

This study developed the superstructure of the entire energy transfer based on hydrogen and fossil fuels. An energy transport pathway was developed by modeling various unit processes within the superstructure. Then, a techno-economic analysis is performed for each pathway to derive energy efficiency (EEF), levelized cost of energy (LOCE), and net CO<sub>2</sub> equivalent emissions (NCE), which are quantitative indicators of process performance. Through the results of techno-economic analysis, a process pathway with optimal economic or environmental feasibility for hydrogen energy transportation was derived. In addition, we identified technical limitations for establishing a hydrogen energy system and proposed technical and political strategies.

## 2. RESEARCH OVERVIEW AND METHODOLOGY

### Research Overview

The goals of this work are: i) to propose optimal economic or environmental hydrogen energy transportation pathway, ii) to determine technical limitation of hydrogen energy transportation system, iii) to propose technical and political strategies for hydrogen energy system. Figure 1 provides procedures for energy transportation system development. First, we set problem by investigating energy sources, conversion technologies, and energy carriers and making a database of the technical and social data. To identify the overall structure of the hydrogen energy transportation system, a superstructure connecting transportation technologies was developed based on technical data. Then, a unit process model on the energy transportation system superstructure was developed and the mass and energy balance for each energy

transport pathway was derived through process synthesis. Through a techno-economic analysis of each developed process model, process efficiency, cost, and CO<sub>2</sub> emissions are calculated.

Primary energy source considered in this study include conventional fossil fuels such as coal, crude oil, and natural gas, as well as renewable energy sources such as wind energy, solar energy, and biomass. These energy sources can be transported as is or converted into other forms of energy carriers. Table 1 shows information on the energy carriers considered in this study. The energy carriers considered in this study of fossil fuels transported as primary energy sources, hydrogen for transporting renewable energy, and hydrogen carriers for efficient transport of hydrogen. Among them, hydrogen is the most important material and serves as an intermediate for transportation with high energy density. However, hydrogen has a low density, making it difficult to transport. Therefore, convertible materials were considered to facilitate transportation of hydrogen from energy resources: ammonia, methanol, methyl cyclohexane (MCH), perhydro dibenzyl toluene (H<sub>18</sub>-DBT).

**Table 1:** Energy carrier for hydrogen transportation system.

Transportable primary energy source	
Coal	
Crude oil	
Natural gas	
Biomass	
Converted energy carrier	
Converted energy carrier	Description
Hydrogen	H <sub>2</sub>
Methanol	CH <sub>3</sub> OH
Ammonia	NH <sub>3</sub>
Methyl cyclohexane	C <sub>7</sub> H <sub>14</sub>
H <sub>18</sub> -Dibenzyl toluene	C <sub>21</sub> H <sub>38</sub>

### Process Analysis Method

In this study, the techno-economic and environmental feasibility was evaluated using the process

performance indicators such as EEF, LCOE and NCE.

To determine technical aspects of hydrogen energy transportation pathways, EEF serves as an assessment indicator. EEF is expressed as the ratio of transported energy to the total of primary energy and utility consumption for conversion, as shown in Eq (1).

$$EEF = \frac{\text{Transported Energy}}{\text{Primary energy source} + \text{Utilities}} \times 100 (\%) \quad (1)$$

To compare economic performance of the hydrogen energy transportation pathways, the LCOE was used as an indicator. The LCOE is calculated using annualized capital investment (ACI) and total operating cost (TOC) and transported energy as shown in Eq (2).

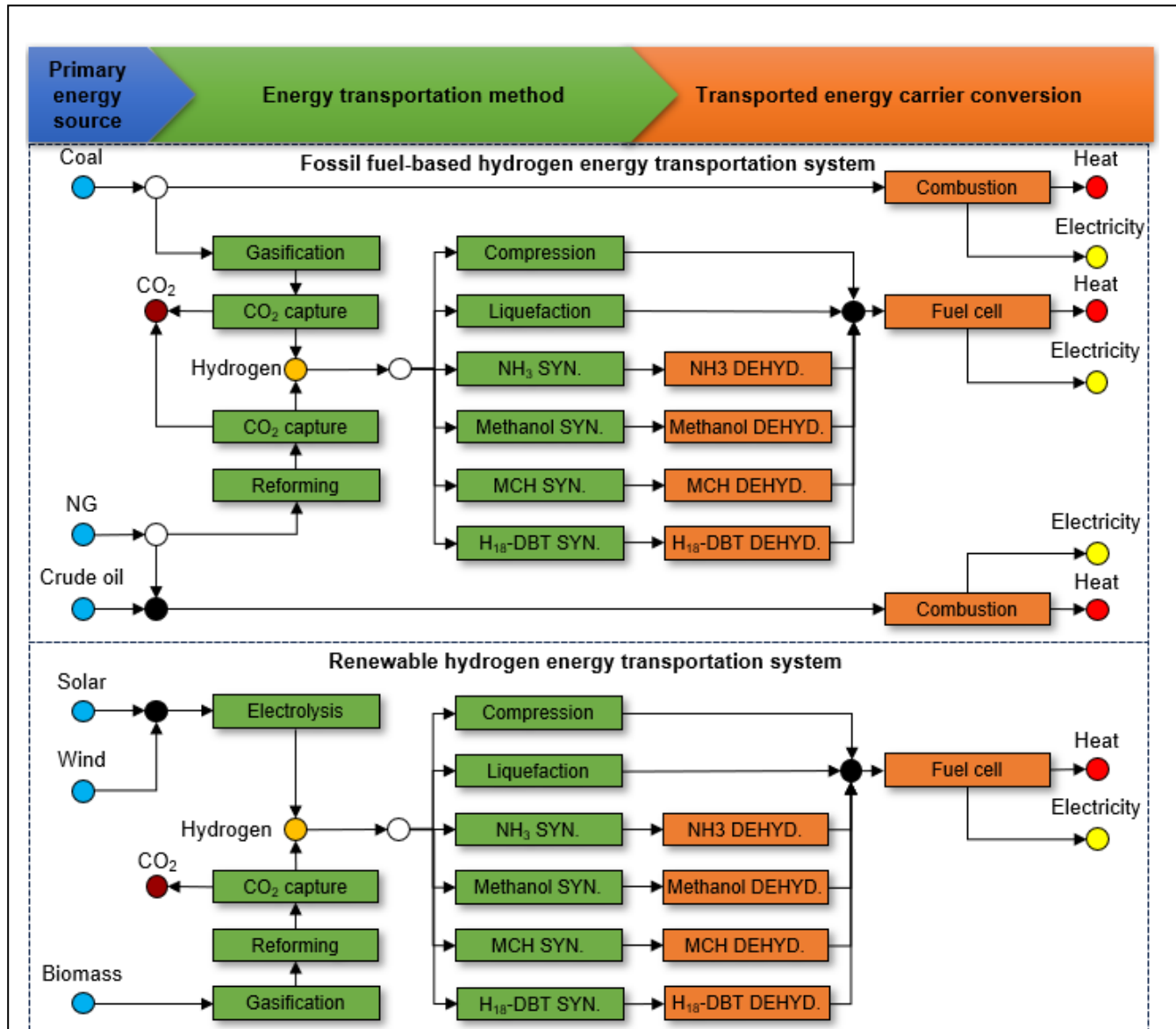
$$LCOE(\frac{\$}{kWh}) = \frac{ACI + TOC}{\text{Transported energy}} \quad (2)$$

ACI is an amortized cost of the capital expenditure (CAPEX), which is specified using a straightforward depreciation method, as shown in Eq. (3). The TOC consists of the primary energy source and utility cost which is specified Eq. (4).

$$ACI = CAPEX \times \frac{\text{interest rate} (1 + \text{interest rate})^{\text{lifetime}}}{(1 + \text{interest rate})^{\text{lifetime}} - 1} \quad (3)$$

$$TOC = \text{Raw material cost} + \text{Utility cost} \quad (4)$$

To assess the environmental performance of the



**Figure 2:** Superstructure of fossil fuel-based and renewable energy transportation system, NG: natural gas, MCH: methyl cyclohexane, H<sub>18</sub>-DBT: perhydro dibenzyl toluene, SYN.: synthesis process, DEHYD.: dehydrogenation process [5-10].



hydrogen energy transportation pathways, the NCE was calculated by including two different indicators to CO<sub>2</sub> emissions. First, the direct emission was GHGs (CO<sub>2</sub>, CH<sub>4</sub>) that are emitted directly through the exhaust and tail gas during process operation. The second is indirect emissions from the usage of utilities, during the process. Thus, the NCE as an environmental indicator is expressed as the total CO<sub>2</sub> emission per transported energy, as shown in Eq. (5).

$$NCE = \frac{\text{Direct emission} + \text{Indirect emission (kg CO}_2\text{-eq)}}{\text{Transported energy}} \quad (5)$$

### 3. RESULT AND DISCUSSION

#### Superstructure Development

As mentioned above, this study developed superstructure for energy transportation method. We developed two hydrogen energy transportation systems, illustrated in Fig 2: fossil fuel-based energy transportation system, renewable energy transportation system. In the case of fossil fuel-based hydrogen transportation systems, both direct transportation primary energy source and transportation via hydrogen are viable options. The pathway that converts fossil fuels into hydrogen and transport will benefit in terms of transport mass due to the high energy density of hydrogen and is expected to have better environmental performance by using the CO<sub>2</sub> capture process. Hydrogen-based energy carriers are essential to transport renewable energy. There are two reasons for this: the form of energy generated from wind or solar energy is electricity, and the low energy density of biomass.

In this study, a technical information investigation was conducted based on the developed superstructure to build a database. It is used for process simulation and techno-economic analysis through investigation of the operating conditions, raw material prices, and energy consumption of each unit process.

#### Process Development

In this study, we developed hydrogen production, storage and refueling process models were developed using Aspen Plus V12.0 according to hydrogen transportation superstructure. The information of each unit process is summarized in Table 2. It was assumed that each pathway feeds the same amount of energy based on the primary energy source.

The hydrogen transport pathway was modeled through synthesis between each unit process, mass and energy balance were derived. In addition, a techno-economic analysis is performed using process equipment sizing and costing data within Aspen Economic Analyzer based on mass and energy balance. In the separation and purification technologies, the pressure swing adsorption (PSA) amine-based CO<sub>2</sub> capture process, distillation

column and air separation unit (ASU) are mainly adopted for efficient each hydrogen transportation pathway.

**Table 2:** Unit process information for hydrogen energy transportation pathways [5-11].

Unit process	Operating conditions (°C, bar)
Coal gasification	700-1200, 30
Natural gas reforming	700-900, 1
Water gas shift	170-350, 1
CO <sub>2</sub> capture process	40, 1
H <sub>2</sub> liquefaction	-253, 20
Liquid H <sub>2</sub> vaporization	40, 30
NH <sub>3</sub> synthesis	300-350, 100-350
NH <sub>3</sub> dehydrogenation	500, 1
Methanol synthesis	220-250, 50
Methanol dehydrogenation	420, 1
MCH synthesis	210-280, 20
MCH dehydrogenation	320, 1
H <sub>18</sub> -DBT synthesis	150-270, 50
H <sub>18</sub> -DBT dehydrogenation	310, 1
Biomass gasification	450-800, 1

#### Comparative Analysis

Based on the technical, economic, and environmental parameters derived from the previous process simulation part, a techno-economic analysis was performed to compare and evaluate each hydrogen transportation pathway. As expected, the conventional system for directly transporting fossil fuels has the best efficiency and economic performance for obtaining electricity and thermal energy. The pathway with the economic feasibility compared to these direct fossil fuel transportation methods is natural gas-based hydrogen production and transportation through MCH (0.15 \$/GJ). Hydrogen production using natural gas is currently a representative method of producing hydrogen at the lowest cost, and the MCH hydrogen storage process also shows the best indicator for cost-effective hydrogen transportation with low energy consumption. Additionally, the pathway from an environmental perspective is to produce hydrogen from solar and wind electricity and then transport the hydrogen using H<sub>18</sub>-DBT. Although this has low economic feasibility due to the price of raw materials, it has the best performance among hydrogen transportation methods in terms of energy consumption, showing the lowest value in terms of NCE due to less use of additional utilities.

### 4. CONCLUSIONS

In this study, we proposed of most economic or environmental pathways in hydrogen energy transportation systems using it for process performance criteria: LCOE and NCE. This was achieved through process synthesis and evaluation based on hydrogen energy transportation superstructure, and the effect on the application of the hydrogen energy system was shown. The major findings from this study as follows:

- To derive the optimal hydrogen energy

transportation pathway according to evaluation standard, process synthesis for energy transport is performed to derive mass and energy balance.

- We conducted techno-economic analysis and derived cost-effective hydrogen transportation method as natural gas-based hydrogen transport to MCH.
- The environmentally best transportation pathway was found to be transporting renewable electricity-based hydrogen as H<sub>18</sub>-DBT.

In summary, we identified the hydrogen transportation pathway using organic compound is promising energy transportation pathways. As future work, we will expand the scope of energy transportation technologically, geographically, and policy-wise. In addition, we will derive major cost drivers of energy transportation through sensitivity analysis after the development of the energy transportation system and propose energy transportation system strategies through various scenario-based analysis.

## ACKNOWLEDGEMENTS

This work was supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry & Energy (MOTIE) of the Republic of Korea (No.20224C10300040).

## REFERENCES

1. Li Lin. et al. Techno-economic analysis and comprehensive optimization of an on-site hydrogen refuelling station system using ammonia: hybrid hydrogen purification with both high H<sub>2</sub> purity and high recovery., *Sustainable Energy & Fuels* 3006-3017 (2020).
2. A.D. Korberg. et Techno-economic assessment of advanced fuels and propulsion systems in future fossil-free ships., *Renewable and Sustainable Energy Reviews* 142 (2021).
3. Jack Shepherd. et al. Open-source project feasibility tools for supporting development of the green ammonia value chain., *Energy Conversion and Management* 274(2022).
4. Ayodeji Okunlola. et al. Techno-economic assessment of low-carbon hydrogen export from Western Canada to Eastern Canada, the USA, the Asia-Pacific, and Europe., *International Journal of Hydrogen Energy* 47 (2022).
5. Ju-Sung Lee. et al. Large-scale overseas transportation of hydrogen: Comparative techno-economic and environmental investigations., *Renewable and Sustainable Energy Reviews* 165 (2022).

6. Yuki Ishimoto. et al. Large-scale production and transport of hydrogen from Norway to Europe and Japan: Value chain analysis and comparison of liquid hydrogen and ammonia as energy carriers., *International Journal of Hydrogen Energy* 45 (2020).
7. Carlos Arnaiz. et al. Techno-economic assessment of blue and green ammonia as energy carriers in a low-carbon future., *Energy Conversion and Management* 255 (2022).
8. Omar Y.H. Elsermagawy. et al. Thermo-economic analysis of reverse water gas shift process with different temperatures for green methanol production as a hydrogen carrier., *Journal of CO<sub>2</sub> Utilization* 41 (2020).
9. M. Niermann. et al. Liquid organic hydrogen carriers (LOHCs) – techno-economic analysis of LOHCs in a defined process chain., *Energy & Environmental Science* 12 (2019).
10. Nobert Heublein. et al. Hydrogen storage using liquid organic carriers: Equilibrium simulation and dehydrogenation reactor design., *International Journal of Hydrogen Energy* 45 (2020).
11. Pedro J. Megia. et al. Hydrogen Production Technologies: From Fossil Fuels toward Renewable Sources. A Mini Review., *Energy & Fuels* 35 (2021).

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Impact of surrogate modeling in the formulation of pooling optimization problems for the CO<sub>2</sub> point sources

HA Pedrozo<sup>a</sup>, MA Zamarripa<sup>b</sup>, JP Osorio Suárez<sup>c</sup>, A Uribe-Rodríguez<sup>c</sup>, MS Diaz<sup>d</sup>,  
LT Biegler<sup>a\*</sup>

<sup>a</sup>Department of Chemical Engineering, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, United States

<sup>b</sup>KeyLogic Systems, Inc., 3168 Collins Ferry Road, Morgantown, WV 26505, United States

<sup>c</sup>Centre for Innovation and Technology Colombian Petroleum Institute, ECOPEPETROL, Piedecuesta 681011, Colombia

<sup>d</sup>Department of Chemical Engineering, Universidad Nacional del Sur, Bahía Blanca 8000, Argentina

\* Corresponding Author: <mailto:lb01@andrew.cmu.edu>

## ABSTRACT

Post-combustion carbon capture technologies have the potential to contribute significantly to achieving the environmental goals of reducing CO<sub>2</sub> emissions in the short term. However, these technologies are energy and cost-intensive, and the variability of flue gas represents important challenges. The optimal design and optimization of such systems are critical to reaching the net zero and net negative goals, in this context, the use of computer-aided process design can be very effective in overcoming these issues. In this study, we explore the implementation of carbon capture technologies within an industrial complex, by considering the pooling of CO<sub>2</sub> streams. We present an optimization formulation to design carbon capture plants with the goal of enhancing efficiency and minimizing the capture costs. Capital and operating costs are represented via surrogate models (SMs) that are trained using rigorous process models in Aspen Plus, each data point is obtained by solving an optimization problem in Aspen Plus equation-oriented approach. Since selecting the functional form of the surrogate model is crucial for the solution performance; we study different SM approaches (i.e., ALAMO, kriging, radial basis function, polynomials, and artificial neural networks) and analyze their impact on solver performance. Numerical results show the computational advantage of using ALAMO while highlighting the increased complexity of using ANN and kriging to formulate optimization problems. Regarding the pooling of CO<sub>2</sub> streams, the optimal designs for the network are not trivial, thus showing the importance of addressing the problem systematically.

**Keywords:** Carbon Capture, Pyomo, Surrogate Model, Process Design, Optimization

## INTRODUCTION

Climate change and global warming are relevant issues that we need to face as a society. In particular, the mitigation of CO<sub>2</sub> emissions is crucial to reduce its high atmospheric concentration, which exceeds 50% of pre-industrial era levels and has recently reached a historical maximum concentration of 424 ppm [1].

Carbon capture strategies play a crucial role in most decarbonization pathways, which seek a net-zero emission future to limit the global average temperature rise to below 1.5 °C [2]. These approaches allow retrofitting of power generation and industrial plants in the short and

medium term. Therefore, carbon capture plants can be installed in industrial complexes to reduce CO<sub>2</sub> emissions of the existing facilities, with minor structural modifications in the production process. Currently, CO<sub>2</sub> absorption using amine solvents is the most developed and reliable technology and it has been tested at an industrial scale for post-combustion processes [3]. However, industrial complexes can include several CO<sub>2</sub> streams to be captured (with different CO<sub>2</sub> flowrates and concentrations), and the implementation of post-combustion technologies for each point source in an industrial pole is challenging due to investment costs and operational complexity. Moreover, even one production process can

include several CO<sub>2</sub> point sources, where their mass flowrates and CO<sub>2</sub> concentrations can differ significantly.

An alternative approach involves the pooling of CO<sub>2</sub> streams from various point sources to optimize the distribution and mixture of CO<sub>2</sub> streams within a specific industrial complex. In this way, the total capture cost for the industries can be reduced, while the application of carbon capture is promoted. When engineering this approach, the problem of mixing the streams and designing carbon capture plants must be addressed systematically, using current knowledge and advances in optimization techniques [4,5].

In order to assist decision-making in the design of the carbon capture plants for a given set of point sources, pooling optimization problems could be formulated. The main decision variables are associated with capture plant capacities, CO<sub>2</sub> concentrations of their flue gas streams, and their CO<sub>2</sub> recoveries. Regarding the calculation of economic and performance metrics of the plants based on these variables, rigorous models are generally required to obtain accurate estimations [6], but including such rigorous models in the problem formulation represents numerical challenges.

Surrogate models (SMs) can help to overcome the numerical issues. In this case, SMs can be used to estimate capital and operating costs for carbon capture plants from the main decision variables, replacing the need for the rigorous CCS model. SMs are simplified approximations capable of estimating output data from a set of input variables, demanding negligible CPU resources. In addition, the mathematical expressions of relevant SMs are compatible with an equation-oriented optimization environment, which enables the use of advanced optimization solvers.

However, the accuracy of the RMs could affect the quality of the optimal solution, and their algebraic forms influence the efficiency of the solution algorithm.

Each surrogate model class offers unique advantages and may be chosen based on the specific requirements of the problem, such as the available data, computational resources, target accuracy, computational efficiency, and the complexity of the studied system. Polynomials are commonly used as reduced models that provide a straightforward and interpretable approach. Radial Basis Functions (RBFs), characterized by their reliance on distance from a central point, can represent complex nonlinear relationships in data, particularly in systems with localized behavior patterns [7]. Kriging is a geostatistical technique used for spatial interpolation and prediction [8,9]. It stands out as a robust prediction method, estimating system responses while considering spatial correlations in scattered or noisy data. Automated Learning and Advanced Modeling for Optimization (ALAMO) facilitates the creation of surrogate models, relying on techniques like linear regression and decision trees

[10]. Finally, Artificial Neural Networks (ANNs) show efficacy in capturing complex data relationships, making them useful for function approximation and pattern recognition [11].

This study considers the impact of using different SMs for the formulation of the CO<sub>2</sub> stream pooling problem. In particular, we use the IDAES computational platform [5,12] to test: i) ALAMO-based SMs, ii) Kriging SMs (PySMO), iii) Radial basis function SMs (PySMO), iv) Polynomials (PySMO), and artificial neural networks (ANN) SMs (Keras).

We compare the surrogate model performances for prediction and for the formulation of optimization problems. This analysis provides insights into efficient ways of including SMs to formulate optimization problems and to elucidate the advantages and limitations of using specific SMs.

## METHODOLOGY

### Formulation of the pooling problem

In this section, we address the mathematical formulation of CO<sub>2</sub> stream pooling problems for carbon capture plant design in industrial complexes with various CO<sub>2</sub> point sources. This problem can be tackled by formulating a nonlinear programming (NLP) approach, as shown in problem (1). We are given a set of point sources  $s \in S$ , with their respective CO<sub>2</sub> mass flowrates ( $F_s^{in}$ ) and molar fractions ( $x_s^{in,CO_2}$ ), and we can design a set of carbon capture plants  $p \in P$ , to achieve a global CO<sub>2</sub> recovery of 90 %. In this problem, we need to design a network to blend CO<sub>2</sub> streams for further treatment. So, important decision variables for the network are the mass flowrates ( $F_{s,p}$ ) from a CO<sub>2</sub> point source  $s$  to the carbon capture plant  $p$ . The capture plants are designed by considering three key decision variables, which are the plant capacity ( $C_p$ ), the CO<sub>2</sub> molar fraction of their input stream ( $x_p^{CO_2}$ ), and the number of transfer units ( $NTU_p$ ).

The objective is to minimize the cost of capture (1.1), and we calculate this metric from the total annualized capital cost (which is obtained from the product between the total capital cost,  $CAPEX^t$ , and the annualization factor,  $\phi$ ), the total operating cost ( $OPEX^t$ ), and the total captured carbon dioxide ( $CO_2^t$ ). In Eq. (1.2), we set constraints for the maximum amount of CO<sub>2</sub> ( $F_s^{in}$ ) from source  $s$ . In Eq. (1.3), we include mass balances for the streams fed to plant  $p$ . In Eq. (1.4), we calculate the CO<sub>2</sub> molar fraction ( $x_p^{CO_2}$ ) of the input stream to plant  $p$ , from the features of the blended streams. Eqs. (1.5) and (1.6) represent the surrogate model function correlating operating ( $OPEX_p$ ) and capital cost ( $CAPEX_p$ ) for each plant with the control variables, which are the plant capacity, CO<sub>2</sub> molar flow, and number of transfer units. This last variable is directly related to the CO<sub>2</sub> recovery ( $rec_p$ ), and it is shown in Eq.



(1.7). Using this transformation, we reduce the complexity associated with collinearity in parametric surrogate models [13]. In Eq. (1.8) and (1.9), we calculate the total operating ( $OPEX^t$ ) and capital costs ( $CAPEX^t$ ) by the summation of individual plant costs, and the cost related to the gas distribution. In this context, the functions  $f^{Od}$  and  $f^{Cd}$  determine the operating and capital costs associated with the utilization of blowers to send the flue gases from point sources to the capture plants.  $f^{Od}$  and  $f^{Cd}$  are based on unit models for pressure changes, which are described in the literature [14]. In Eq. (1.10), we calculate the total captured carbon dioxide ( $CO_2^t$ ) considering all plants, and we add a constraint to ensure a minimum global  $CO_2$  recovery of 90 % in Eq. (1.11). Finally, Eqs. (1.12)-(1.14) are the bounds on the main decision variables for designing the plants.

$$\min \frac{\phi CAPEX^t + OPEX^t}{CO_2^t} \quad (1.1)$$

$$s. t. F_s^{in} \geq \sum_{p \in P} F_{s,p} \quad s \in S \quad (1.2)$$

$$\sum_{s \in S} F_{s,p} = C_p \quad p \in P \quad (1.3)$$

$$x_p^{CO_2} \sum_s \frac{F_{s,p}}{x_s^{in,CO_2}} = C_p \quad p \in P \quad (1.4)$$

$$OPEX_p = f^O(C_p, x_p^{CO_2}, NTU_p) \quad p \in P \quad (1.5)$$

$$CAPEX_p = f^C(C_p, x_p^{CO_2}, NTU_p) \quad p \in P \quad (1.6)$$

$$rec_p = (1 - e^{-NTU_p}) \quad p \in P \quad (1.7)$$

$$\sum_{p \in P} OPEX_p + \sum_{s,p} f^{Od}(F_{s,p}) = OPEX^t \quad (1.8)$$

$$\sum_{p \in P} CAPEX_p + \sum_{s,p} f^{Cd}(F_{s,p}) = CAPEX^t \quad (1.9)$$

$$CO_2^t = \sum_{p \in P} C_p rec_p \quad (1.10)$$

$$CO_2^t \geq 0.9 \sum_i F_s^{in} \quad (1.11)$$

$$C_p \leq 1 \quad p \in P \quad (1.12)$$

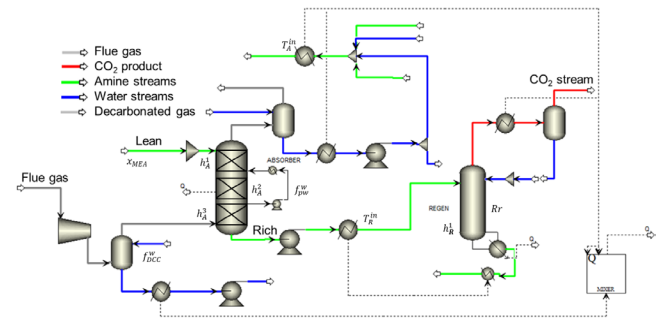
$$0.03 \leq x_p^{CO_2} \leq 0.4 \quad p \in P \quad (1.13)$$

$$0.8 \leq rec_p \leq 0.99 \quad p \in P \quad (1.14)$$

## Carbon capture process

In this study, we consider the conventional absorption-based process with intercooling using mono-ethanolamine (MEA) solvent, which is a well-established technology that serves as a baseline for comparison purposes. We developed a rigorous model for the MEA-based carbon capture process using templates from the Carbon Capture Simulation Initiative (CCSI), as discussed in [6]. The model is implemented in Aspen Plus, employing the electrolyte NRTL thermodynamic framework and incorporating specific interaction parameters derived from experimental data. The process flowsheet, see Figure 1, features a blower and direct contact cooler for flue

gas preconditioning, an absorber with a pump-around system, and a stripping column for solvent regeneration. The “Rich” amine stream from the absorber undergoes heating in a heat exchanger before entering the stripping column, where  $CO_2$  separation from the solvent occurs. The resulting  $CO_2$  product stream and regenerated solvent are obtained from the top and bottom of the stripping column, respectively. The Aspen Plus model is solved using the equation-oriented (EO) mode, in order to leverage the EO optimization features that are available in the software. In this way, the convergence properties of the model improve significantly [6]. Therefore, data generated to train the surrogate models with this method correspond to optimization runs.



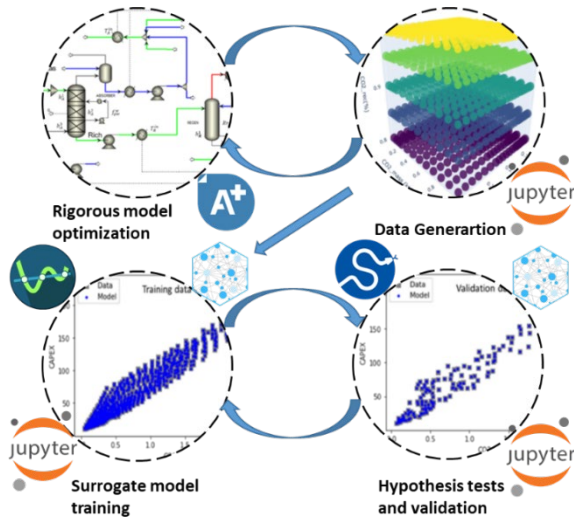
**Figure 1:** Aspen Plus® flowsheet for the carbon capture process

## Surrogate modeling approaches

We build surrogate models generated from optimization runs using the Aspen Plus® rigorous model described earlier. Constructing surrogate models requires a substantial amount of high-quality data points from rigorous models, and running each sample point manually is an inefficient method due to the time-consuming nature of data generation. Therefore, we developed a framework for automated optimization and simulation using the rigorous model [13], and integrating various computational platforms, as shown in Fig. 2. In this work, we employ three input variables (inlet  $CO_2$  mass flowrate in the flue gas, inlet  $CO_2$  mole fraction, and number of transfer units) to generate SMs for the calculation of logarithmic transformations of CAPEX and OPEX. In this case, using log-transformed output variables improves the residual patterns for the parametric SMs [15].

The IDAES computational platform [12] is employed for surrogate model generation, exploiting its capabilities for efficient data splitting (training and validation), different surrogate modeling approaches, and straightforward post-processing analysis (including the generation of 2D scatter plots and calculating fitting metrics).





**Figure 2:** Software workflow between Aspen, Python environments (SciPy, IDAES), and ALAMO in Jupyter Notebook

In order to assess the effect of the functional form of the surrogate model on the solution performance, we examine various surrogate models (SMs) ideas. This work includes surrogate models based on ALAMO (Automatic Learning of Algebraic Models for Optimization), PySMO (Python Surrogate Modeling Toolbox), and Tensorflow Keras.

ALAMO is a powerful tool for surrogate modeling [10]. ALAMO specializes in employing advanced algorithms to select appropriate basis functions and model parameters simultaneously. In this way, it identifies essential features and relationships within datasets. Additionally, ALAMO takes advantage of advanced optimization techniques to perform automated sampling. The AlamoPy application enables the use of this tool through Python code, facilitating post-processing tasks, like generating 2D scatter plots and calculating fitting metrics.

PySMO is a versatile and powerful tool to aid in the creation of surrogate models, offering a range of functionalities for efficient data analysis. The toolbox facilitates the integration of these surrogate models into optimization frameworks, enhancing the overall efficiency of computational processes. It supports the generation of surrogate models using various techniques, including kriging, Radial basis functions (RBF), and polynomials. In this work, we generate kriging and radial basis functions considering the regularization option to handle noise in data and to improve the smoothness of the curve. For RBF, we consider cubic basis transformation. For polynomials, we set a maximum polynomial order of four, including bilinear terms.

Tensorflow Keras is a user-friendly tool designed to simplify the implementation of artificial neural networks (ANNs). As an integral component of the Python ecosystem, Keras provides an intuitive interface for

constructing, training, and deploying neural networks. Its flexibility and modularity make it suitable for various applications, including surrogate modeling. We leverage Keras to develop ANN-based surrogate models, benefiting from its capabilities for efficient data analysis and optimization. We consider one hidden layer and hyperbolic tangent activation functions. The integration of the Keras ANN into the optimization formulation is done by using the toolkit OMLT (Optimization and Machine Learning Toolkit) [16].

## CASE STUDY

In this work, we consider the application of carbon capture to an iron and steel complex, which includes several point sources. The manufacturing process includes three main sectors: iron production, steel production, and a mini mill; producing 5 million tons of steel annually [17].

**Table 1:** CO<sub>2</sub> streams for iron and steel production [17,18]

	Notation	Flowrate of CO <sub>2</sub> (Mt/y)	CO <sub>2</sub> content (mol%)
Iron production			
Power plant stack	S0	3.69	23
Coke ovens gas	S1	1.73	27
Blast furnace	S2	1.94	21
Sinter plant stack	S3	1.67	8
Steel production			
BOF stack	S4	0.28	15
Hot strip mill stack	S5	0.14	7
Plate mill stack	S6	0.60	7
Lime kiln stack	S7	0.05	7
Mini mill			
Electric arc furnace off gas	S8	0.11	40

The blast furnace, fueled by coke from the coke oven, facilitates the reduction of iron ore to molten iron. CO<sub>2</sub> emissions originate from various sources, including coke oven gas generated during coal heating, ore preparation, and the basic oxygen steelmaking unit responsible for reducing carbon in pig iron. The plant is highly integrated, utilizing most blast furnaces and coke oven gases as low-grade fuel for the sinter plant, blast furnace stoves, and lime kiln. Table 1 shows detailed information on CO<sub>2</sub> emission sources within the process.

## OPTIMIZATION RESULTS

In this section, we discuss the results associated with surrogate model fitting and their efficiency within optimization formulation. The problems were implemented in Pyomo [19] and run with IPOPT [20], using the multistart strategy.

**Table 2:** CAPEX surrogate model fitting metrics.

CAPEX	R <sup>2</sup>	Mean squared error	Mean rel. error (%)	Max. rel. error (%)
ALAMO Training data	0.9987	1.02·10 <sup>-3</sup>	0.1411	0.6416
ALAMO Validation data	0.9986	1.13·10 <sup>-3</sup>	0.1442	0.6510
Polynomials Training data	0.9964	2.93·10 <sup>-3</sup>	0.2360	1.2025
Polynomials Validation data	0.9953	3.72·10 <sup>-3</sup>	0.2707	1.2503
RBF Training data	1.0000	1.13·10 <sup>-6</sup>	0.0030	0.0478
RBF Validation data	0.9999	1.10·10 <sup>-4</sup>	0.0357	0.3664
Kriging Training data	1.0000	3.02·10 <sup>-5</sup>	0.0221	0.1750
Kriging Validation data	0.9999	6.52·10 <sup>-5</sup>	0.0274	0.3303
ANN Training data	0.9997	2.56·10 <sup>-4</sup>	0.0684	0.4983
ANN Validation data	0.9996	2.88·10 <sup>-4</sup>	0.0726	0.4499

The computational experiments were performed on an Intel(R) Core(TM) i5-8250U CPU @1.80 GHz and 12 GB RAM. It should be mentioned that the problems were solved using the option "limited-memory" to approximate the Hessian of the Lagrangian using a quasi-Newton method; we observe that this enhances the convergence of the multistart strategy. Approximating the Hessian by this quasi-Newton approach ensures positive definite matrices, thus avoiding potential issues related to the use

of highly nonlinear functions and random initial guesses in the domain.

**Table 3:** OPEX surrogate model fitting metrics

OPEX	R <sup>2</sup>	Mean squared error	Mean rel. error (%)	Max. rel. error (%)
ALAMO Training data	0.9998	1.35·10 <sup>-4</sup>	0.0553	0.2986
ALAMO Validation data	0.9998	1.61·10 <sup>-4</sup>	0.0606	0.3008
Polynomials Training data	0.9979	1.77·10 <sup>-3</sup>	0.2313	0.7236
Polynomials Validation data	0.9975	2.04·10 <sup>-3</sup>	0.2448	0.8448
RBF Training data	1.0000	6.54·10 <sup>-7</sup>	0.0029	0.0479
RBF Validation data	0.9999	9.77·10 <sup>-5</sup>	0.0388	0.2848
Kriging Training data	1.0000	2.89·10 <sup>-5</sup>	0.0239	0.2351
Kriging Validation data	0.9999	4.92·10 <sup>-5</sup>	0.0310	0.2575
ANN Training data	0.9999	9.77·10 <sup>-5</sup>	0.0477	0.3056
ANN Validation data	0.9999	8.89·10 <sup>-5</sup>	0.0463	0.2822

Tables 2 and 3 show key statistical metrics for the CAPEX and OPEX surrogate models, respectively. In particular, the R<sup>2</sup> coefficients surpass 0.995, and the mean squared errors and mean relative errors exhibit consistency between training and validation data, for all surrogate modeling approaches. Only RBF-based SMs show important differences between the training and the validation data, however, the errors for the validation sets are still quite low, as compared to the other approaches. These results show that the selected approaches can fit the target set of data points accurately, to predict the log transformations of CAPEX and OPEX from the main input variables (CO<sub>2</sub> mass flowrate in the flue gas, CO<sub>2</sub> mole fraction, and NTU). It is highlighted that the ALAMO and

Polynomials approaches could present important errors (>50 %) when data is not transformed; in contrast, RBF, kriging, and ANN exhibit superior fitting capacity across diverse datasets. As observed in Tables 2 and 3, polynomial surrogate models yield the highest maximum relative errors (approximately 1.2% and 0.8% for CAPEX and OPEX, respectively). ALAMO and ANN models present relatively similar accuracy, with ANN exhibiting slightly superior indicators compared to ALAMO. In particular, the RBF approach has the best accuracy metrics, closely followed by Kriging surrogate models, indicating the fit capacity of these approaches.

Regarding optimization results using the different surrogate modeling approaches, Table 4 shows the main results. These results correspond to solving the optimization problem by considering 50 different initial guesses to increase the possibility of finding the global optimum.

The problem size of the NLP formulation is the same for the different cases, except when using ANN. We observe that the number of variables increases by 336 % for this case due to the additional variables in the hidden layer. On the other hand, no additional variables are required for the other surrogate model approaches.

CPU times show that ALAMO-based SM formulation is the most efficient strategy, requiring only 47 s to solve the problem with the multistart strategy. Using polynomials is the second most efficient approach, but it increases CPU time by 78 % with respect to the case of using ALAMO. On the other hand, the Kriging-based SM is the one that demands the highest CPU time, indicating the effect of using this highly nonlinear approach. Although the NLP formulation using ANNs has the most equations, it demands 30% lower CPU time, as compared to the use of Kriging.

The objective function values are also shown in Table 4. The lowest objective function is found when using RBF-based SM, and the highest value corresponds to the case of using ALAMO. However, the differences in the objective functions are lower than 3.7 %, suggesting no significant differences in the carbon capture cost obtained when using different surrogate modeling approaches.

In terms of the specific configurations resulting from the optimization process across different approaches, diverse pooling strategies result, based on the surrogate models. It is important to note that the pooling problem exhibits multiple optimal solutions, and in some instances, differences between them may not be substantial. Nevertheless, there are key ideas that are consistent in the solutions, and one of them is the use of a percentage of the point sources to achieve the 90 % target recovery, as detailed in Table 4. Point sources S0-S2, S4, and S8 are fully utilized in all cases due to their high CO<sub>2</sub> concentration. As S3 also presents a relatively high concentration of CO<sub>2</sub>, it is nearly fully employed across all

cases except in the scenario involving polynomial-based SMs, where it is completely used. On the other hand, given the relatively low CO<sub>2</sub> concentration in S5, S6, and S7, they are generally excluded from the CO<sub>2</sub> stream pooling, except when formulating the problem using polynomials, where 50% of S6 is utilized. We note that the constraint of achieving a minimum CO<sub>2</sub> recovery of 90% is active in all cases.

**Table 4:** Main results from the NLP optimization using different surrogate approaches

	Alamo	Polynomials	RBF	Kriging	ANN
N° de vars.	586	586	586	586	2533
N° de Eqs.	486	486	486	486	2433
CPU time	47	84	1176	2275	1594
Obj. function	40.6	40.0	39.1	39.1	40.3
Use S0	100%	100%	100%	100%	100%
Use S1	100%	100%	100%	100%	100%
Use S2	100%	100%	100%	100%	100%
Use S3	98%	100%	97%	97%	97%
Use S4	100%	100%	100%	100%	100%
Use S5	0%	0%	0%	0%	0%
Use S6	0%	50%	0%	0%	0%
Use S7	0%	0%	0%	0%	0%
Use S8	100%	100%	100%	100%	100%

### ALAMO solution

Given the efficiency of ALAMO surrogate models, we have selected this approach to present the optimization results. The optimal pooling of CO<sub>2</sub> streams for carbon capture purposes is shown in Fig. 3. We observe that some point sources are higher than the maximum plant capacity (see Eq. (1.12)); therefore, they could be sent to more than one plant. In Fig. 3, we also show the percentage distribution of each point source. In particular, point sources S2 and S3 are fed to 5 and 4 plants, respectively. Although each point source could be treated in two plants, numerical results suggest that these arrangements allow an improved capture cost. These could be associated with avoiding the design of plants operating with low CO<sub>2</sub> concentrations.

We also show the input variables associated with the design of each plant in Fig. 3. It is observed that two kinds of designs are selected to achieve the recovery target. On the one hand, p<sub>0</sub>, p<sub>3</sub>, p<sub>5</sub>, and p<sub>9</sub> operate with CO<sub>2</sub> concentrations of 0.095-0.102, CO<sub>2</sub> mass flowrates of 0.6Mt/y, and recoveries around 94 %. On the other hand, the remaining plants are designed to function with high CO<sub>2</sub> concentrations (0.21-0.28), while the resulting CO<sub>2</sub>

mass flowrate and CO<sub>2</sub> recovery are at their respective upper bounds (1 Mt/y and 99 %).

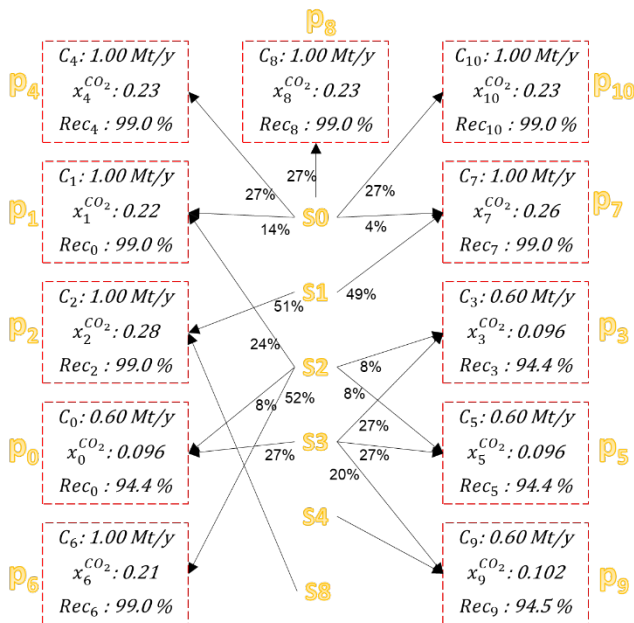


Figure 3: Optimal solution of the CO<sub>2</sub> streams pooling problem using ALAMO-based surrogate models. **p** and **S** correspond to the carbon capture plants and point sources (see references in Table 4), respectively.

## CONCLUSIONS

In this work, we evaluate different surrogate models (SMs) for formulating CO<sub>2</sub> stream pooling problems using the IDAES computational platform. We explore ALAMO-based SMs, Kriging SMs, Radial Basis Function SMs, Polynomials, and Artificial Neural Networks. ALAMO-based SMs prove the most efficient, with significantly lower CPU time. Despite differences in process configurations among the SMs, the impact on capture cost is minimal (below 3.7%). The optimal results associated with pooling configuration change with the surrogate modeling approach, but key ideas, like using high CO<sub>2</sub> concentration sources and maintaining a 90% recovery target, remain consistent across all cases.

## ACKNOWLEDGEMENTS

This work was funded by the Center for Advanced Process Decision-Making (CAPD) at Carnegie Mellon University.

## REFERENCES

- National Oceanic and Atmospheric Administration. Broken record: Atmospheric carbon dioxide levels jump again. (2023)

- Shu DY, Deutz S, Winter BA, Baumgärtner N, Leenders L, Bardow A. The role of carbon capture and storage to achieve net-zero energy systems: Trade-offs between economics and the environment. *Renew Sustain Energy Rev.* 178:113246 (2023)
- Kazemifar F. A review of technologies for carbon capture, sequestration, and utilization: Cost, capacity, and technology readiness. *Greenh Gases Sci Technol.* 12(1):200–230 (2022)
- Misener R, Floudas CA. Advances for the pooling problem: Modeling, global optimization, and computational studies. *Appl Comput Math.* 8(1):3–22 (2009)
- Biegler LT, Miller DC, Okoli CO. Chapter 2 - Don't search—Solve! Process optimization modeling with IDAES. In: *Simulation and Optimization in Process Engineering.* Ed: Bortz M, Asprion N. Elsevier. 33–55 (2022)
- Pedrozo A, Valderrama-Ríos CM, Zamarripa M, Morgan J, Osorio-Suárez JP, Uribe-Rodríguez A, and Biegler LT. Equation-Oriented Optimization Applied to the Optimal Design of Carbon Capture Plants Using Rigorous Models. *Ind Eng Chem Res.* 62(19):7539–7553 (2023)
- McDonald DB, Grantham WJ, Tabor WL, Murphy MJ. Global and local optimization using radial basis function response surface models. *Appl Math Model.* 31(10):2095–2110 (2007)
- Caballero JA, Grossmann IE. An algorithm for the use of surrogate models in modular flowsheet optimization. *AIChE J.* 54(10):2633–2650 (2008)
- Krige DG. A statistical approach to some mine valuation and allied problems on the Witwatersrand: By DG Krige. University of the Witwatersrand. (1951)
- Cozad A, Sahinidis N V, Miller DC. Learning surrogate models for simulation-based optimization. *AIChE J.* 60(6):2211–2227 (2014)
- Henao CA, Maravelias CT. Surrogate-based superstructure optimization framework. *AIChE J.* 57(5):1216–32 (2011)
- Lee A, Ghouse JH, Eslick JC, Laird CD, Sirola JD, Zamarripa MA, Gunter D, Shinn JH, Dowling AW, Bhattacharyya D, and Biegler LT. The IDAES process modeling framework and model library—Flexibility for process simulation and optimization. *J Adv Manuf Process.* 3(3):e10095 (2021)
- Pedrozo H, Valderrama-Ríos C, Zamarripa M, Morgan J, Osorio-Suárez J, Diaz S, Biegler L, and Uribe-Rodríguez A. Surrogate modeling for CO<sub>2</sub> capture by chemical absorption based on data from rigorous model optimization. *Authorea* (2023)
- Pedrozo HA, Rodriguez Reartes SB, Vecchiotti AR, Diaz MS, Grossmann IE. Optimal Design Of

- Ethylene And Propylene Coproduction Plants With Generalized Disjunctive Programming And State Equipment Network Models. *Comput Chem Eng.* 107295 (2021)
15. Seber GAF, Lee AJ. Linear Regression Analysis. Vol. 330. *John Wiley & Sons.* (2003)
  16. Ceccon F, Jalving J, Haddad J, Thebelt A, Tsay C, Laird CD, and Misener R. OMLT: Optimization & machine learning toolkit. *J Mach Learn Res.* 23(1):15829–15836 (2022)
  17. Wiley DE, Ho MT, Bustamante A. Assessment of opportunities for CO<sub>2</sub> capture at iron and steel mills: An Australian perspective. *Energy Procedia.* 4:2654–61 (2011)
  18. Scholes CA. Membrane gas-solvent contactor process for efficient carbon capture from multiple emission sources within an industrial complex. *Chem Eng Sci.* 278:118915 (2023)
  19. Bynum ML, Hackebeil GA, Hart WE, Laird CD, Nicholson BL, Siirola JD, Watson JP, Woodruff DL. Pyomo-optimization modeling in Python. Berlin/Heidelberg, Germany: Springer; (2021)
  20. Wächter A, Biegler LT. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math Program.* 106(1):25–57 (2006)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Optimal Clustered, Multi-modal CO<sub>2</sub> Transport Considering Non-linear Costs – a Path-planning Approach

Kang Qiu<sup>a,\*</sup>, Sigmund Eggen Holm<sup>a</sup>, Julian Straus<sup>a</sup>, and Simon Roussanaly<sup>a</sup>

<sup>a</sup> SINTEF Energy Research, Department of Gas Technology, Trondheim, Norway

\* Corresponding Author: [kang.qiu@sintef.no](mailto:kang.qiu@sintef.no).

## ABSTRACT

An important measure to achieve global reduction in CO<sub>2</sub> emissions is CO<sub>2</sub> capture, transport, and storage. The deployment of CO<sub>2</sub> capture requires the development of a shared CO<sub>2</sub> transport infrastructure, where CO<sub>2</sub> can be transported with different transport modes. Furthermore, the cost of CO<sub>2</sub> transport can be subject to significant economies of scale effects with respect to the amount of CO<sub>2</sub> transported, also mentioned as clustering effects. Therefore, optimizing the shared infrastructure of multiple CO<sub>2</sub> sources can lead to significant reductions in infrastructure costs. This paper presents a novel formulation of the clustered CO<sub>2</sub> transport network. The Markov Decision Process formulation defined here allows for more detailed modeling of non-linear, discrete transport costs and increased geographical resolution. The clustering effects are modeled through cooperative multi-agent interactions. A multi-agent, reinforcement learning-based algorithm is proposed to optimize the shared transportation network, with examples illustrating the results of the method.

**Keywords:** Carbon Capture, Energy Systems, Artificial Intelligence, Technoeconomic Analysis, Supply Chain.

## INTRODUCTION

Reducing CO<sub>2</sub> emissions is crucial to limit the temperature increase due to global warming [1]. One important measure for reducing CO<sub>2</sub> emissions is CO<sub>2</sub> capture, transport, and storage (CCS). CCS can be coupled with various individual emission sources [2]. Among the most important considered emissions sources are thermal power generation, industrial facilities (*e.g.* cement plants or iron and steel mills [3]), as well as bioenergy production facilities [4] or direct air capture [5] to achieve net removal of CO<sub>2</sub> from the atmosphere. The rollout of CO<sub>2</sub> capture, independent of the source of the CO<sub>2</sub>, requires the concurrent development of a CO<sub>2</sub> transport infrastructure to the targeted CO<sub>2</sub> storage locations. CO<sub>2</sub> can typically be transported via pipelines, barges, ships, trains, and trucks; or combinations of these. The cost of CO<sub>2</sub> transport can experience significant economies of scale with respect to the amount of transported CO<sub>2</sub> [6]. Hence, developing a common CO<sub>2</sub> infrastructure with clustering of CO<sub>2</sub> from different sources can reduce the cost for each individual actor. This is especially important for industries in which smaller

amounts of CO<sub>2</sub> are captured and/or that are located far from the desired storage location.

The implementation of CCS in the different sectors is frequently analyzed using large-scale (mixed integer) linear optimization problem, both individually [7], focusing on a single sector [8], or combined with the analysis of the decarbonization of the complete energy system [9]. These models face in general however two problems; First, it is difficult to include economies of scale due to the linear description of transport cost. Second, all geographical nodes must be defined before the optimization. The latter is especially important when considering the clustering of individual CO<sub>2</sub> sources to utilize economies of scale. In these models, transport of CO<sub>2</sub> and change of transport mode is only possible in predefined geographical nodes. One approach to solving the problem is utilized by Sunny et al. [8], where a regular grid is superimposed on the investigated region. Correspondingly, the model includes the whole region, allowing for both utilization of economies of scale and change of transport mode. However, the inclusion of a regular grid and binary investments in technologies to include economies of scale complicates the problem, resulting in a

reduced geographic resolution.

This study presents a reinforcement learning approach to determining optimal multi-modal transport networks. The novelty of this work is threefold. First, while most of the literature on CCUS for network deployment focuses on onshore CO<sub>2</sub> pipelines as a mean of transport [10][11], the present work includes tanked-based CO<sub>2</sub> transport options (such as trucks, trains, and barges), which are now expected to play a more significant role in CCUS deployment than a few years ago [12]. Secondly, as the costs of pipeline- and tanked-based transport are significantly non-linear despite most of the literature linearizing it, at least by part, a new methodology is proposed to reduce the computational complexity of the large-scale MINLP transportation problem. Thirdly, as such network developments are unlikely to involve only one agent, the methodology incorporates a multi-agent perspective for cooperative agents that minimizes the overall transportation network costs to account for the economies of scale effects. The method utilizes the concept of regular grids on a geographic region but allows for a higher geographical resolution as it focuses exclusively on the potential CO<sub>2</sub> transport network without considering time. It can be utilized for identifying geographical nodes that should be included in large-scale optimization problems.

## METHOD

### Problem Description

Multi-modality and accommodating for detailed geographical resolution increase the solution space significantly and may affect the numerical tractability of the algorithm. Furthermore, transport costs are, in general, not linear with respect to the transport distance and volume. Therefore, non-linear, and discrete costs should be considered in the method. Finally, CO<sub>2</sub> transported from a capture site to a storage location in limited amounts, while only considering a single source and single sink, will result in a high unit transportation cost. Assuming that there will be a significant amount of CO<sub>2</sub> to be transported from future capture sites, clustering multiple CO<sub>2</sub> sources, and designing a shared infrastructure may reduce transport costs drastically. Table 1 summarizes the CO<sub>2</sub> transport network problem addressed in the methodology proposed.

The multi-modality, the discrete, non-linear costs of the formulated transport problem, and the clustering potential will lead to a large-scale MINLP problem, which may lack numerical tractability and be difficult to solve. This paper proposes a multi-agent reinforcement learning-based approach to determine a minimum-cost CO<sub>2</sub> transportation network to circumvent the problem. Truck-based and pipeline-based transports are considered in this paper, though the method can be extended

to more transport modes. Furthermore, information about the amount and location of the transported CO<sub>2</sub> and the location of storage sites is needed. Detailed knowledge of geographical data on existing road networks and terrain can be easily incorporated into the method.

**Table 1:** CO<sub>2</sub> transport network problem description

Problem	Background	Effect
Multi-modality	Consideration of multiple transport modes	Large solution space
Detailed cost modeling	Non-linear and discrete costs for transport and re-conditioning	Complex objective function
Clustering	Considering multiple capture sites with a shared network to lower transport costs.	Multi-agent approach

### Environment

The environment is modeled as a graph  $G = (N, E)$ , where each node  $n \in N$  represents some geographic location and edges  $e \in E$  connect pairs of distinct nodes in the graph. This formulation allows for extendibility in environment representation, as geographical information can be added to each corresponding edge. When an edge connects two nodes corresponding to two different geographic locations, the edge represents some transport mode between them. The edges have data necessary to infer transportation costs, such as distance between nodes, and can incorporate information about terrain. When nodes connected by an edge are at the same location, however, the edge represents the conditioning of CO<sub>2</sub>. Conditioning of CO<sub>2</sub> is needed when a unit of CO<sub>2</sub> transitions from one transport mode to another. Furthermore, a conditioning cost is also applied to a unit of CO<sub>2</sub> when transitioning from a source node to a transport mode or from a transport mode to a sink node.

The nodes and edges in the graph are ordered in different layers. Each layer represents a single transport mode, and the layers are connected by directed edges for conditioning needed to change the transport mode. However, the used algorithm works for a general graph and is not dependent on a layer-structured environment.

### Optimal CO<sub>2</sub> Transport Network as an MDP

This paper proposes to frame the network optimization problem as a Dynamic Programming (DP) problem. DP refers to both a modeling framework and a solution method for sequential decision-making processes. The fundamental DP problem is a Markov Decision Process (MDP), which is described by discrete state transition

dynamics:

$$x_{k+1} = f_k(x_k, u_k), \quad k = 0, 1, \dots, N-1, \quad (1)$$

for a state  $x_k$  at time step  $k$  to transition to the subsequent state  $x_{k+1}$  given an input  $u_k$ . Each state corresponds to a specific node  $n$  in the network, indicating the agent's current location. The action  $u_k$  corresponds to an edge  $e$  that the agent chooses to traverse, representing either CO<sub>2</sub> transport or conditioning modeled as movement from one node to the other. The general MDP formulation allows for stochastic state transition dynamics. This study, however, neglects state transition stochasticity in the transport investment problem.

For every state transition, the agent incurs a stage cost:

$$g_k(x_k, u_k), \quad k = 0, 1, \dots, N-1, \quad (2)$$

and a terminal cost at the terminal state  $g_N(x_N)$ . The stage cost is given as feedback from the environment. As inputs  $u_k$  are applied and stage costs  $g_k$  are observed, the goal of DP is to minimize the cumulative costs over all time steps. These cumulative costs are expressed in the value function, subject to Bellman's principle of optimality [13]:

$$V = g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, u_k). \quad (3)$$

The current problem is solved as steady-state problem. Each agent represents some total annual volume of transported CO<sub>2</sub> moving through the environment describing all possible transport paths from source (capture site) to multiple possible sinks (storage sites), but it is assumed that CO<sub>2</sub> is transported continuously over the year. DP provides a flexible mathematical framework to model sequential decision-making processes under relatively weak assumptions.

While DP algorithms can find the global optimum solution given non-linear and discrete costs, it often struggles with numerical tractability in large state-spaces (so-called "curse of dimensionality"). Approximate Dynamic Programming (ADP) is a family of algorithms derived from DP, which trade off global optimality for numerical tractability. ADP algorithms include reinforcement learning (RL) [14], the A\*-Algorithm [15], Model Predictive Control [16], and Stochastic Programming [17].

Understanding the problem posed in this paper as a DP problem allows for drawing on the rich methodological background to solve such problems. The DP framework can explicitly address modeling a higher geographical resolution, the multi-modality of CO<sub>2</sub> transport, and representing more complex investment costs.

## Multi-agent Approximate Dynamic Programming

The DP framework also offers a solution to model

volume-specific transportation cost decrease when the amount of CO<sub>2</sub> transported increases. These economies of scale effect is often shown in pipeline investments, where building a pipeline with an increased diameter is less costly than building two separate pipelines with smaller diameters to transport the same flow. This effect is also described in minimum concave cost network flow problems which are solved as a network flow optimization [18]. Minimum concave cost network flow problems are shown to be equivalent to the DP problem formulation and can be solved with a DP algorithm known as the send-and-split method [19].

This paper utilizes the MDP formulation and models the economies of scale as multi-agent interactions. The agents are multiple units of CO<sub>2</sub> traversing the graph. When agents invest in pipeline transportation at the same location, the total costs are modelled according to a concave cost function and distributed among the agents contributing to the investment decision. A clustering interaction is therefore defined as multiple agents transitioning over the same transport edge.

Modelling the clustering of CO<sub>2</sub> transport poses another complexity in the transport network investment problem, which increases the computational burden on the solution. The paper explicitly does not consider time-dependency and the rollout of investments. Therefore, agents do not have to traverse the same edge within the same time step.

In an ADP framework, the clustering of agents is incentivized directly through reductions of stage costs to single agents, leading to an overall reduction in system costs (total cost incurred by all agents). The proposed algorithm should be able to utilize discrete and non-linear CO<sub>2</sub> transport network costs and converge to a feasible overall system cost in the transport value chain. We define the pipeline investment problem as a cooperative game with transferable utility, for which the Shapely value [20] is included to distribute the costs fairly among agents [21].

## Algorithm

This paper proposes a modified actor-critic RL algorithm. The critic is formulated as a central evaluator, whose role is to evaluate the current policy followed by the actor [22]. In the multi-agent context, the critic observes the actions of all actors and apply a temporal difference (TD) update [23].

Each actor samples a stochastic policy to determine the next action in a discrete action space, due to the discrete graph structure of the environment. The stochastic policy is a mapping of state-action pair to a real number  $p$  indicating the preference:

$$\Pi(x, u) \rightarrow p. \quad (4)$$

The preference  $p$  is scaled to represent a viable

probability distribution for the stochastic policy with Boltzmann scaling. Each actor will sample a path from their source node to a sink node with their current policy and observe the stage cost feedback from the environment. Because there is no time step dependency defining actor collaboration, the policy update is only applied after each agent completes their path from source to sink.

The centralized critic updates and maintains a value function mapping each state to an expected cost by following the current policy  $V(x)$ . The critic observes the stage costs of a path incurred by all agents and calculates the TD update for the centralized value function.

The actors receive the TD error  $\delta$  to update the actor policy. The TD error is calculated as:

$$\delta(x_k, u_k, x_{k+1}) = g_k(x_k, u_k) + V(x_{k+1}) - V(x_k). \quad (5)$$

The value function is tabular, the actor policy is Boltzmann-sampled. The optimal path of an agent is defined as the trained greedy policy, which maximizes the preference  $p$  over all actions  $u$ .

#### Algorithm 1: Modified Actor-Critic Algorithm

---

```

1. Critic: initialize  $V(x) \leftarrow 0$ , "x
2. For each agent a
3.   Actor: Initialize  $\Pi^a(x, u) \leftarrow 0$ , "x, u
4. For each episode
5.   For each agent a
6.     Initialize  $x \leftarrow x_{source}^a$ 
7.     Initialize path  $p^a$  as empty list
8.     While x is not  $x_{sink}$  do
9.       Actor:  $u \leftarrow \Pi(x)$  get action dictated
           by policy in current state x
10.      Actor: observe state transition  $x \rightarrow x'$ 
11.      Actor: append state-action-state tuple
            $(x, u, x')$  to path
12.      Update state:  $x \leftarrow x'$ 
13. Critic: collect all paths p from agents
14. Critic: determine common state-action-state
           tuples between agents
15. For each agent a
16.   For each step  $(x, u, x')$  in  $p^a$ 
17.     Critic: calculate stage cost g
           considering common state-action-
           state tuples
18.     Critic:  $\delta \leftarrow g + V(x') - V(x)$ 
19.     Critic:  $V(x) \leftarrow V(x) + \alpha\delta$ 
20.     Actor:  $\Pi^a(x, u) \leftarrow V(x) - \alpha\delta$ 

```

---

## EXAMPLES

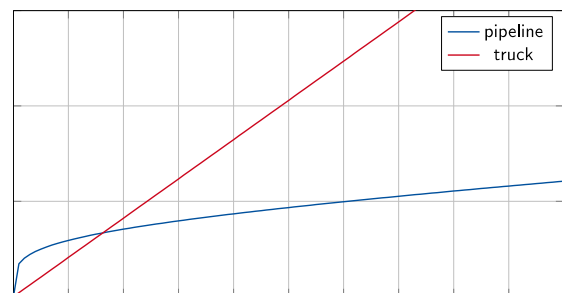
This paper presents a demonstration of the algorithm methodology in simplified examples. As this paper focuses on presenting the CO<sub>2</sub> transport problem formulation and a methodology, geographical data is not used in the examples presented here; however, the method can be extended to incorporate geographical knowledge in the stage costs observed by the agents. The

environment is represented as an equidistant 6-by-6 grid. The graph has two layers representing the two transport modes in the example, truck-based and pipeline-based transports. Every location on the graph is reachable by either transport mode. In total, the graph has 36 transport nodes with additional source and sink nodes.

The transportation costs are dependent on the transport mode, the distance associated with the edges of the environment, and the total transported volume of CO<sub>2</sub> over the edge. For edges representing conditioning of CO<sub>2</sub>, the cost function only depends on the total volume of CO<sub>2</sub> transported over the edge, and what transport modes are available at each of the nodes connected by the edge.

The annualized transport cost (M€/y), including conditioning, is used as a key performance indicator. In order to use representative costs for both truck- and pipeline-based transports, multiple evaluations were carried out with the iCCS tool, an integrated techno-economic and environmental assessment tool for CCS value chains developed by SINTEF Energy Research [24–26]. Based on these evaluations, cost functions of pipeline- and truck-based transports depending on the annual volume and transport distances were regressed for the ranges relevant to the examples.

The Shapley value is computed for the pipeline costs to incentivize agent cooperation. The costs of both truck transport and pipeline transport are plotted in Figure 1 for a set distance of 100 km. Figure 1 shows that pipelines are cheaper than trucks for volumes above 0.16 Mt of CO<sub>2</sub> transported per year.



**Figure 1:** Transport investment cost per volume of CO<sub>2</sub> for pipeline and truck.

### Example 1

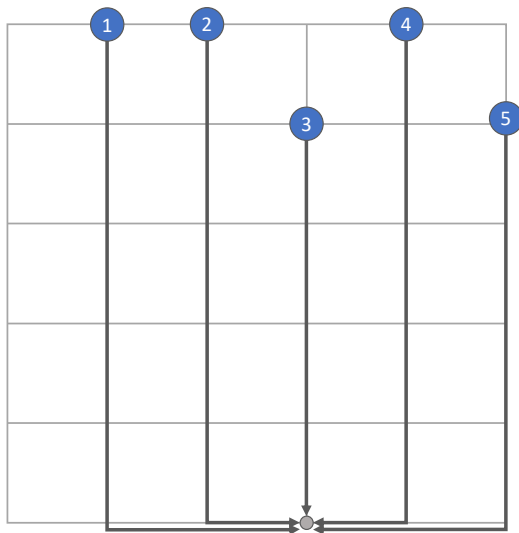
In this example, five agents with relatively large volumes traverse the graph to one sink node located at node (6,4) (shown as a small, light gray circle in the figures, row 6, column 4). The source node locations and volume of each agent are listed in Table 2. The agents are visualized as blue circles with their corresponding numbers in

the figures. Pipeline transport, the preferred transport mode with and without cooperation between the agents, is represented as a solid arrow in the figures.

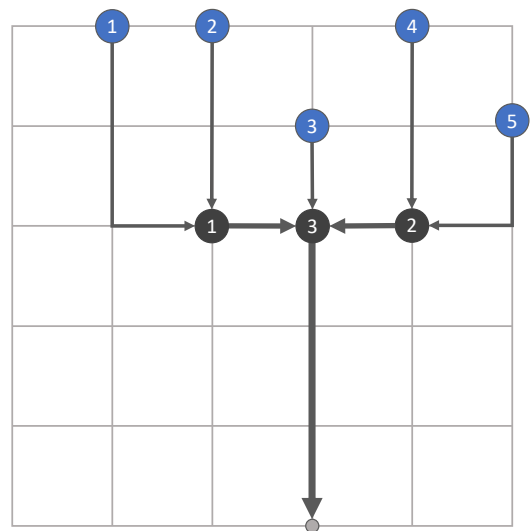
The example is solved in 23.8 s for 4,000 episodes with a learning rate  $\alpha = 10^{-4}$ . Figure 2 shows the optimal paths of each agent traversing the graph without clustering effects. The algorithm identifies three possible clustering points shown as a larger, grey circle. The total system costs are reduced by 32 % by considering clustering of the agents. Figure 3 shows the results of the calculated path of the agents considering that agents traversing the same pipeline transport edge will reduce costs according to Shapley values.

**Table 2:** Agents in example 1

Agent	Volume	Source node
1	1.5 Mt	(1,2)
2	1.5 Mt	(1,3)
3	1.0 Mt	(2,4)
4	2.0 Mt	(1,5)
5	2.5 Mt	(2,6)



**Figure 2:** Results of the algorithm for agents from Example 1 without agent interactions through clustering. Due to the large volumes being transported in Example 1, pipelines are still preferred compared to trucks.



**Figure 3:** Results of the algorithm for agents from Table 2 traversing the grid when clustering is incentivized with cost reduction. Possible clustering points are identified as dark grey, larger circles. The solid arrows represent pipeline investments. Thicker arrows indicate a larger amount of CO<sub>2</sub> transported through the pipeline.

The example is solved in 23.8 s for 4,000 episodes with a learning rate  $\alpha = 10^{-4}$ . Figure 2 shows the results of the calculated path of the agents considering that agents traversing the same pipeline transport edge will reduce costs according to Shapley values. Figure 3 shows the optimal paths of each agent traversing the graph without clustering effects. The algorithm identifies three possible clustering points shown as a larger, grey circle. The total system costs are reduced by 32% by considering the clustering of the agents.

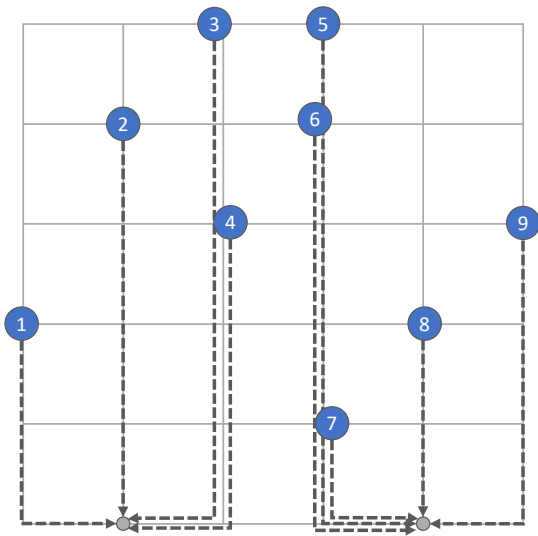
### Example 2

This example models nine agents traversing the graph to two possible sink nodes located at (6,2) and (6,5). Source location and volume of each agent are listed in Table 3.

**Table 3:** Agents in example 2.

Agent	Volume	Source node
1	0.15 Mt	(4,1)
2	0.15 Mt	(2,2)
3	0.15 Mt	(1,3)
4	0.15 Mt	(3,3)
5	0.15 Mt	(1,4)
6	0.15 Mt	(2,4)
7	0.15 Mt	(5,4)
8	0.15 Mt	(4,5)
9	0.15 Mt	(3,6)





**Figure 4:** Results for Example 2 without clustering effects. Due to the small volume of agents, truck transport, represented in dashed arrows, is preferred.

Example 2 is solved in 57.2 s for 8000 episodes and a learning rate of  $\alpha = 10^{-4}$ . Figure 4 shows that the volume of each agent causes truck transport (represented in dashed arrows) to be preferred if agent interactions, i.e., clustering effects, are not considered.

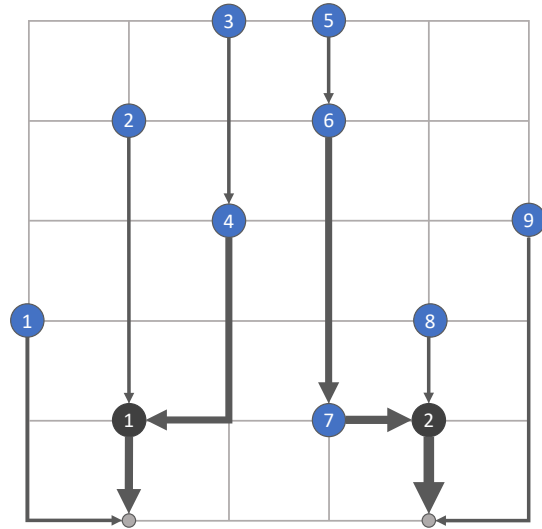
In Figure 5, the paths for each agent considering clustering effects are shown. All agents choose to forego truck transport, as pipeline transport in larger volumes seems to be preferred. Agents 1 and 9 traverse the graph through pipelines from source to sink without clustering with other agents, even though truck transport would yield lesser costs. Pipeline transport exhibits a high variance in costs. Specifically, in the case of cooperation between agents, these costs are shared, which leads the agents to prefer pipeline transport in the multi-agent scenario even though the greedy policy does not present clustering effects. Nonetheless, total system costs are reduced by 28% with cooperative behavior.

## DISCUSSION

The method proposed in this paper is formulated to be able to incorporate detailed geographical data. The DP framework is chosen as a flexible modelling tool which can be extended for multi-modal and geographical representation.

The drawback of DP is the scalability of the solution methods, as they need to accommodate a large search space. Furthermore, the search space may grow rapidly with the number of agents, the complexity of possible agent policies, and the possible interactions between agents. This is especially exacerbated as not only agent interactions within one time step but across all time steps

in a path are considered. The complexity of the algorithm is also increased due to the tabular representation of policy and value function.



**Figure 5:** Results of the algorithm for agents from Table 3 traversing the grid, where clustering is incentivized with cost reduction. The agents traverse the graph to two possible sinks, represented as small light grey circles. Two possible clustering points are identified.

Investments in the CO<sub>2</sub> transport network are the trained policies of agents corresponding to edges of the environment graph. These edges are defined as a “fixed transportation route”, starting in one node and ending in another. This representation is sufficient for geographically fixed investments such as pipelines, flexible resources in the transport network, such as trucks, are not accurately captured.

While the multi-agent approach allows clustering effects as agent interactions to be captured in the first place, convergence to optimality remains an issue. The multi-agent formulation inherently results in learning a policy in a non-stationary environment, where the behavior of other agents constantly influences the learning goal of one agent. These dynamics between agent behavior lead to unclear convergence to a global optimality [27]. In the method proposed, agents communicate indirectly through the centralized critic and the calculated TD error. However, each agent does not consider other agents’ behavior when sampling their policies. Some multi-agent algorithms, therefore, include either an approximation of other agents’ policies in the decision-making of one agent or a communication term [28].

## CONCLUSION AND FURTHER WORK

This paper presents a formulation of the CO<sub>2</sub>

transport network investment as a MDP. The algorithm is formulated to accommodate incorporating detailed geographical modelling and acknowledging the so-called clustering effects or economies of scale in pipeline investments. The transport network investment problem is formulated as a DP problem. This paper focuses on a description of the proposed methodology while presenting simple examples to illustrate the results of such an algorithm.

DP is known to be inhibited with exponentially growing state-action-space complexity and the resulting computational intractability also known as the curse of dimensionality. A multi-agent reinforcement learning-based algorithm proposed to address the state space complexity of DP. While the multi-agent perspective offers a framework where agents make decisions autonomously and collaborate to achieve lower costs which are in turn distributed fairly onto the agents, multi-agent reinforcement learning is subject to rapidly increasing input complexity as the input space grows in dimension with the number of agents and nonstationarity leading to convergence issues [29] which needs to be addressed in further works.

While the illustrative examples presented in this paper are solved quickly, the computational time is expected to grow rapidly with the number of agents, which may make the application of such an algorithm to truly large-scale systems difficult (infeasible). However, the proposed methodology retains its novelty by considering the economies of scale effects that are commonly neglected in single-agent systems. Further work should, therefore, focus on integrating a sequential rollout approach [30], which would cause the computational time to increase linearly with the number of agents while retaining the fundamental cost improvement property for convergence. One issue in this approach is that the economies of scale effects, which are represented in the interactions between the agents, require the agents to act concurrently to calculate the Shapley values for the cooperating agents. Therefore, the convergence under such system properties needs to be addressed.

## ACKNOWLEDGEMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 101022487 (AC-CSESS project).

## REFERENCES

1. Intergovernmental Panel On Climate Change (Ippc): Climate Change 2022 – Impacts, Adaptation and Vulnerability: Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental

- Panel on Climate Change. Cambridge University Press (2023)
2. Bains, P., Psarras, P., Wilcox, J.: CO 2 capture from the industry sector. *Prog. Energy Combust. Sci.* 63, 146–172 (2017).
3. Voldsund, M., Gardarsdottir, S., De Lena, E., Pérez-Calvo, J.-F., Jamali, A., Berstad, D., Fu, C., Romano, M., Roussanaly, S., Anantharaman, R., Hoppe, H., Sutter, D., Mazzotti, M., Gazzani, M., Cinti, G., Jordal, K.: Comparison of Technologies for CO2 Capture from Cement Production—Part 1: Technical Evaluation. *Energies.* 12, 559 (2019).
4. Fridahl, M., Lehtveer, M.: Bioenergy with carbon capture and storage (BECCS): Global potential, investment preferences, and deployment barriers. *Energy Res. Soc. Sci.* 42, 155–165 (2018).
5. Beuttler, C., Charles, L., Wurzbacher, J.: The Role of Direct Air Capture in Mitigation of Anthropogenic Greenhouse Gas Emissions. *Front. Clim.* 1, 10 (2019).
6. Roussanaly, S., Berghout, N., Fout, T., Garcia, M., Gardarsdottir, S., Nazir, S.M., Ramirez, A., Rubin, E.S.: Towards improved cost evaluation of Carbon Capture and Storage from industry. *Int. J. Greenh. Gas Control.* 106, 103263 (2021).
7. Becattini, V., Gabrielli, P., Antonini, C., Campos, J., Acquilino, A., Sansavini, G., Mazzotti, M.: Carbon dioxide capture, transport and storage supply chains: Optimal economic and environmental performance of infrastructure rollout. *Int. J. Greenh. Gas Control.* 117, 103635 (2022).
8. Sunny, N., Mac Dowell, N., Shah, N.: What is needed to deliver carbon-neutral heat using hydrogen and CCS? *Energy Environ. Sci.* 13, 4204–4224 (2020).
9. Seck, G.S., Hache, E., Sabathier, J., Guedes, F., Reigstad, G.A., Straus, J., Wolfgang, O., Ouassou, J.A., Askeland, M., Hjorth, I., Skjelbred, H.I., Andersson, L.E., Douguet, S., Villavicencio, M., Trüby, J., Brauer, J., Cabot, C.: Hydrogen and the decarbonization of the energy system in europe in 2050: A detailed model-based analysis. *Renew. Sustain. Energy Rev.* 167, 112779 (2022).
10. Hasan, M.M.F., First, E.L., Boukouvala, F., Floudas, C.A.: A multi-scale framework for CO2 capture, utilization, and sequestration: CCUS and CCU. *Comput. Chem. Eng.* 81, 2–21 (2015).
11. Morbee, J., Serpa, J., Tzimas, E.: Optimised deployment of a European CO2 transport network. *Int. J. Greenh. Gas Control.* 7, 48–61 (2012).
12. Oeuvray, P., Burger, J., Roussanaly, S., Mazzotti, M., Becattini, V.: Multi-criteria assessment of inland and offshore carbon dioxide transport options. *J. Clean. Prod.* 443, 140781 (2024).
13. Bertsekas, D.P.: Dynamic programming and optimal

- control. Athena scientific, Nashua, NH (2012)
14. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction.
  15. LaValle, S.M.: Planning algorithms. Cambridge University Press, Cambridge ; New York (2006)
  16. Rawlings, J.B., Mayne, D.Q., Diehl, M.: Model predictive control: theory, computation, and design. Nob Hill Publishing, Santa Barbara, California (2020)
  17. Birge, J.R., Louveaux, F.: Introduction to Stochastic Programming. Springer Science & Business Media (2011)
  18. Guisewite, G.M., Pardalos, P.M.: Minimum concave-cost network flow problems: Applications, complexity, and algorithms. *Ann. Oper. Res.* 25, 75–99 (1990).
  19. Erickson, R.E., Monma, C.L., Veinott, A.F.: Send-and-Split Method for Minimum-Concave-Cost Network Flows. *Math. Oper. Res.* 12, 634–664 (1987).
  20. Shapley, L.S.: 17. A Value for n-Person Games. In: Kuhn, H.W. and Tucker, A.W. (eds.) *Contributions to the Theory of Games (AM-28)*, Volume II. pp. 307–318. Princeton University Press (1953)
  21. Van Den Brink, R.: An axiomatization of the Shapley value using a fairness property. *Int. J. Game Theory.* 30, 309–319 (2002).
  22. Grondman, I., Busoniu, L., Lopes, G.A.D., Babuska, R.: A Survey of Actor-Critic Reinforcement Learning: Standard and Natural Policy Gradients. *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* 42, 1291–1307 (2012).
  23. Oroojlooy, A., Hajinezhad, D.: A review of cooperative multi-agent deep reinforcement learning. *Appl. Intell.* 53, 13677–13722 (2023).
  24. Jakobsen, J., Roussanaly, S., Anantharaman, R.: A techno-economic case study of CO<sub>2</sub> capture, transport and storage chain from a cement plant in Norway. *J. Clean. Prod.* 144, 523–539 (2017).
  25. Roussanaly, S., Deng, H., Skaugen, G., Gundersen, T.: At what Pressure Shall CO<sub>2</sub> Be Transported by Ship? An in-Depth Cost Comparison of 7 and 15 Barg Shipping. *Energies.* 14, 5635 (2021).
  26. Kristin Jordal, Øyvind Langørgen, Donghoi Kim, Francesco Finotti, Simon Roussanaly, Nicola Marsh, Mari Voldsund: The CCS Midt-Norge cluster - Industrial CCS collaboration for exploring synergies and common interests. SINTEF Energy Research (2023)
  27. Claus, C., Boutilier, C.: The Dynamics of Reinforcement Learning in Cooperative Multiagent Systems.
  28. Lowe, R., Wu, Y., Tamar, A., Harb, J., Abbeel, O.P., Mordatch, I.: Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments.
  29. Yang, Y., Wang, J.: An Overview of Multi-Agent Reinforcement Learning from Game Theoretical Perspective (2021)
  30. Bertsekas, D.: Multiagent Reinforcement Learning: Rollout and Policy Iteration. *IEEECAA J. Autom. Sin.* 8, 249–272 (2021).

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Promising Opportunities for Improving Round-Trip Efficiencies in Liquid Air Energy Storage (LAES)

Siyue Ren<sup>a,b</sup>, Truls Gundersen<sup>a\*</sup>, Xiao Feng<sup>b</sup>

<sup>a</sup> Department of Energy and Process Engineering, Norwegian University of Science and Technology (NTNU), Kolbjoern Hejes vei 1.A, NO-7491, Trondheim, Norway

<sup>b</sup> School of Chemical Engineering and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi, 710049, China

\* Corresponding Author: truls.gundersen@ntnu.no.

## ABSTRACT

As a promising electricity storage system, Liquid Air Energy Storage (LAES) has the main advantage of being geographically unconstrained. LAES has a considerable potential in energy efficiency improvement by utilizing compression heat and integrating with other systems. In this work, the Stirling Engine (SE) is introduced to improve the energy efficiency of the LAES system. Three LAES-SE systems are modelled in Aspen HYSYS and optimized by the Particle Swarm Optimization (PSO) algorithm. The studied systems include (i) the LAES system with 3 compressors and 3 expanders (3C+3E) using an SE to recover the compression heat, (ii) the 3C+3E LAES system with LNG regasification and SE, and (iii) the 3C+3E LAES system with solar energy and SE. The optimization results show that the Round-Trip Efficiencies (*RTEs*) of the LAES-SE system and the LNG-LAES-SE systems are 68.2% and 73.7%, which are 3.2% and 8.7% points higher than the basic 3C+3E LAES-ORC system with an *RTE* of 65.0%. For the Solar-LAES-SE system, a revised *RTE* and the economic performance with solar energy input are optimized. The traditional *RTE* for the Solar-LAES-SE system, which only accounts for power produced and consumed in the discharging and charging sections, is 189% and 173% respectively, when optimized with respect to energy and economic performances. The revised *RTE* accounts for the integrated external sources, avoiding the confusing result that the *RTE* becomes larger than 100%. The energy and economic performances of the Solar-LAES-SE system are proved to be the best compared with the Solar-LAES-ORC and Solar energy directly heated-LAES systems.

**Keywords:** Modeling and Simulation., Optimization, Solar Energy, Energy Efficiency, Liquid Air Energy Storage, Stirling Engine

## 1 INTRODUCTION

The transition from fossil fuels to renewable energy forms requires substantial investments in energy storage due to the intermittent nature of renewables and fluctuations on the demand side. There is a considerable number of storage technologies, from mature to emerging, all with their inherent advantages and disadvantages. Focusing on storing electricity, the dominating technology is Pumped Hydro Electrical Storage (PHES) with 96% of the world capacity. Another mature, but still developing, technology for storing electricity is different forms of batteries. Emerging energy storage technologies include Compressed Air Energy Storage (CAES) and Liquid Air Energy Storage (LAES).

While PHES and CAES have geographical constraints, standalone LAES can be located wherever it is suitable. When considering integration with external sources of heating and cooling, as the present work will discuss, even LAES has some constraints regarding location. The main disadvantage of LAES has been its relatively low Round-Trip Efficiency (*RTE*). From the first published concepts with an *RTE* of 54.4% [1], recent studies ([2], [3] and [4]) were able to improve the *RTE* for standalone LAES to 68.2% by optimizing the number of compressor and expander stages as well as the cold thermal energy recovery cycles.

In contrast to PHES and CAES, LAES also has substantial improvement potential by utilizing opportunities for integration with external sources of heating and

cooling. Obvious candidates are waste heat streams from neighboring industrial plants and utilization of the cold energy that is released when regasifying LNG. At existing LNG terminals, LNG is regasified using sea water or air; in fact, even combustion of natural gas has been used. This considerable waste of cold energy can easily be converted into power using different types of heat engines such as Organic Rankine Cycle (ORC) and Stirling Engine (SE). In our work, we have even considered solar energy as a heat source for the LAES. These alternatives have been evaluated in the framework of an LAES that operates in three distinct modes: charging, storage and discharging.

Table 1 shows the systems studied and compared in this work. The 3C+3E LAES-ORC system in reference [4] with an  $RTE$  of 65.0% is taken as a base case. In the LAES-SE system, the ORC is replaced by a Stirling engine. In the LNG-LAES-SE system, the heat sink of the LAES-SE system, cooling water, is replaced by the LNG regasification process. In the Solar-LAES-SE system, the heat source and sink of the Stirling engine are solar energy and air in the discharging section, respectively. In this case, all the surplus compression heat is used to drive an ORC.

The energy and economic performances of the Solar-LAES-SE system are compared with the Solar-LAES-ORC system [5] and the Solar directly heated-LAES system [6]. In the first three systems, only the traditional  $RTE$  is studied while in the solar energy integrated systems, both the traditional and the revised  $RTE$  as well as the net income are studied. In the Solar-LAES-SE system and Solar directly heated-LAES system, the charging and discharging sections are decoupled.

The serial numbers for the systems in Table 1 refer to: (1) LAES-ORC system, (2) LAES-SE system, (3) LNG-LAES-SE system, (4) Solar-LAES-SE system, (5) Solar-LAES-ORC system, and (6) Solar directly heated-LAES system.

**Table 1:** Different LAES systems studied.

Sys-tems	Reference	$RTE_{LAES}$	$RTE_{Solar-LAES}$	Net income
1	[4]	✓	×	×
2	This study	✓	×	×
3	This study	✓	×	×
4	This study	✓	✓	✓
5	[5]	✓	✓	✓
6	[6] and this study	✓	✓	✓

## 2 METHODS

Aspen HYSYS has been used to model the different processes while the Particle Swarm Optimization (PSO) algorithm in MATLAB has been used to find the optimal

$RTE$  for standalone LAES as well as LAES integrated with ORC, SE and Solar energy. The economic performance of the combined Solar-LAES-SE system has also been optimized. Equations (1) and (2) show the traditional and revised  $RTE$ , respectively.

$$RTE_{LAES} = \frac{W_{out}t_{dc}}{W_{in}t_{cc}} = \frac{(W_{AT}-W_{CP})t_{dc}}{W_{AC}t_{ch}} \quad (1)$$

$$RTE_{Solar-LAES} = \frac{W_{el,out}t_{dc}}{(W_{el,in}+\eta Q_{th,in})t_{ch}} \quad (2)$$

In these equations,  $w$  is specific work,  $t$  is time,  $dc$  and  $ch$  are for discharging and charging, respectively,  $\eta$  is the power generation efficiency for converting solar heat into power, and  $Q_{th,in}$  is the solar heat to the system.

## 3 SIMULATION MODEL

### 3.1 Basic LAES System

According to reference [4], the LAES system with 2 compressor stages and 3 expander stages (2C+3E) has the highest  $RTE$ . Their study also revealed that surplus compression heat will only be produced when the number of compressor stages is equal to or larger than the number of expansion stages. Unfortunately, the power produced from such surplus compression heat is not enough to compensate for the reduced power production in the discharging section. In summary, they found that (2C+3E) is better than (3C+3E), (3C+4E) is better than (4C+4E), and (4C+5E) is better than (5C+5E). Of these three systems, (2C+3E) has the highest  $RTE$ . However, when the number of compressor stages in the charging section is only 2, the compression ratio will be 13.11 [4], leading to a high investment cost and some concern about lack of maturity. Taking into consideration both energy efficiency and a more practical compression ratio in the charging section, an LAES system with 3 compressor and 3 expander stages is adopted as the base case for the study in this paper. An Organic Rankine Cycle (ORC) is used to recover the surplus compression heat of the 3C+3E LAES system. The optimized  $RTE$  of the base case is 65.0% according to reference [4]. Based on this, different thermal energy recovery systems and external heat sources and sinks are integrated to improve the  $RTE$  of the LAES system.

### 3.2 LAES-Stirling Engine System

The Stirling engine is an efficient heat recovery technology, whose ideal efficiency is as high as the Carnot cycle, which theoretically has the highest possible thermal efficiency [7]. The Stirling engine is proved to be a promising method to improve the  $RTE$  of the LAES system as an alternative to other compression heat recovery cycles including the ORC, Kalina Cycle and Absorption Refrigeration Cycle. Figure 1 shows the flowsheet of the LAES-SE system simulated in Aspen HYSYS. The heat



sink of the Stirling engine is cooling water at ambient conditions.

The Malmo relation [8] is used to model the Stirling engine in the LAES-SE system because the temperature difference between the heat source (thermal oil from the charging part) and sink (cooling water) of the Stirling engine is low. The Malmo relation is shown by Equations (3)-(4).

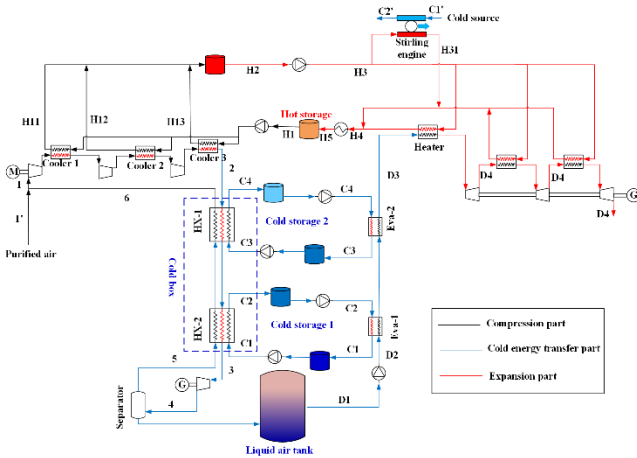
$$P_{\text{Stirling}} = (\eta_H \times \eta_{\text{Mech}} \times \eta_{\text{Thermo}}) \times K_C \times Q_{\text{in}} \quad (3)$$

$$\eta_{\text{Thermo}} = \frac{1-\tau}{1+(1-e) \times (1-\tau) / [(k-1) \times \ln \frac{V_1}{V_2}]} \quad (4)$$

$$\tau = \frac{T_{L,S}}{T_{\text{Heater,wall}}}, T_{L,S} = T_{\text{cooler wall}} + \Delta T_{\text{low}},$$

$$k = \frac{C_P}{C_V}$$

where  $\eta_H$  is the heat source efficiency, which is in the range of 0.85–0.95;  $\eta_{\text{Mech}}$  is the mechanical efficiency, which is in the range of 0.75–0.90;  $K_C$  is the Stirling coefficient, which is in the range of 0.55–0.88; these parameters are taken as 0.9, 0.83 and 0.71 [9].  $e$  refers to the regenerative effectiveness;  $V_1 = V_D + V_P$  is the volume at state 1 and  $V_2 = V_D$  is the volume at state 2 of the compression process;  $V_D$  is displacer swept volume in  $\text{m}^3$  and  $V_P$  is power piston swept volume in  $\text{m}^3$ ;  $\tau$  is the temperature ratio between the heat source and heat sink.  $T_{H,S}$  refers to the high-temperature source and  $T_{L,S}$  refers to the low-temperature source of the Stirling engine.  $k$  is the ratio between specific heat capacity at constant pressure and specific heat capacity at constant volume, and is considered to have a constant value (1.667 for Helium).



**Figure 1.** Flowsheet of the LAES-SE system.

### 3.3 LAES-SE System with LNG regasification

In order to further improve the heat recovery efficiency of the LAES system and thereby improve the  $RTE$  of the entire system, the LNG regasification process is adopted as a heat sink of the Stirling engine with the compression heat as a heat source. Figure 1 is used to

illustrate both the LAES-SE system where cooling water is used as heat sink for the Stirling engine and the LNG-LAES-SE system where LNG regasification replaces cooling water as heat sink.

In the LNG-LAES-SE system, the Pseudo Stirling model [10] is used to simulate the Stirling engine because in this case, the temperature difference between the heat source and heat sink is large. The Pseudo Stirling model is shown in Equations (5)-(10).

$$P_{\text{Stirling}} = \eta_{\text{pcy}}(Q_{\text{high}} - Q_{\text{loss}}) \quad (5)$$

where  $P_{\text{Stirling}}$  refers to the power produced by the Stirling engine,  $\eta_{\text{pcy}}$  refers to the polytropic efficiency as shown in Equation (7),  $Q_{\text{high}}$  is the heat provided by the heat source, and  $Q_{\text{loss}}$  is the heat released to the heat sink by the Stirling engine as shown in Equation (6).

$$Q_{\text{loss}} = Q_{\text{high}}(1 - \eta_{\text{mech}}) \quad (6)$$

where  $\eta_{\text{mech}}$  refers to the mechanical efficiency of the Stirling engine.

$$\eta_{\text{pcy}} = \left[ \frac{(1-RV^{1-k}) - \zeta(RV^{1-k}-1)}{(1-RV^{1-k}) + (1-\zeta)(1-\varepsilon_{ST})} \right] \quad (7)$$

where  $\varepsilon_{ST}$  refers to the efficiency of the Stirling engine regenerator (0.94),  $RV$  refers to the piston compression ratio (1.23), and  $\zeta$  refers to the ratio between minimum and maximum temperature inside the Stirling cycle as shown in Equation (8):

$$\zeta = \frac{T_{L,S}}{T_{H,S}} \quad (8)$$

$$T_{H,S} = T_{\text{heater,wall}} - \Delta T_{\text{high}} \quad (9)$$

$$T_{L,S} = T_{\text{cooler,wall}} + \Delta T_{\text{low}} \quad (10)$$

where  $\Delta T_{\text{high}}$  refers to the temperature change of the heat source and  $\Delta T_{\text{low}}$  refers to the temperature change of the heat sink.

### 3.4 Solar Energy Integrated LAES-SE System

In addition to the case where surplus compression heat is used as a heat source for the Stirling engine, solar energy can also be integrated as a heat source to increase the power generated in the Stirling engine. As shown in Figure 2, solar energy is integrated with the discharging section of the LAES system. The solar heat carried by molten salt is taken as heat source and the air before expanders is taken as heat sink for the Stirling engine. All compression heat from the charging section is used to drive an ORC to generate more electricity. As a result, the charging and discharging sections are decoupled, resulting in improved flexibility and operability of the LAES system compared with previous cases.

The solar energy utilization model is represented by Equations (11)-(13). The efficiency of the collector is defined as the ratio of the energy absorbed by the working fluid in the collector ( $Q_c$ ) to the available energy from the

sun ( $Q_s$ ), as shown by Equation (11). The efficiency of the collector is given by Equation (12), as suggested by Blanco et al. [11]

$$\eta_c \equiv \frac{Q_c}{Q_s} \quad (11)$$

$$\eta_c = 0.75 - 0.000045\Delta T - 0.039 \frac{\Delta T}{G_b} - 0.0003G_b \left(\frac{\Delta T}{G_b}\right)^2 \quad (12)$$

where  $\Delta T$  is the difference between the mean temperature in the solar collector and ambient temperature, as shown by Equation (13).

$$\Delta T = (T_{\text{hot}} + T_{\text{cold}})/2 - T_{\text{amb}} \quad (13)$$

where  $T_{\text{amb}}$  is the ambient temperature, which also has an impact on the system.

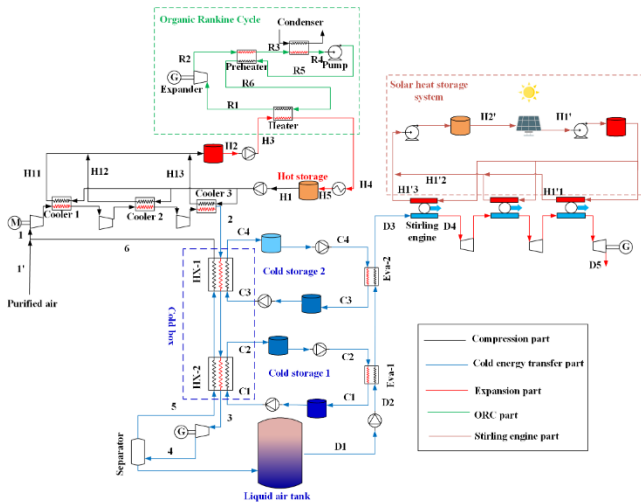


Figure 2. Flowsheet of the Solar-LAES-SE system.

## 4 OPTIMIZATION OF DIFFERENT LAES SYSTEMS

Based on the simulation model of different LAES systems, the Particle Swarm Optimization (PSO) algorithm is used to obtain the optimal  $RTE$  of these systems. The minimum temperature difference of the heat exchangers in the discharging section is fixed as  $10^\circ\text{C}$  while that of the heat exchangers in the charging section is set as a variable changing from  $10^\circ\text{C}$  to  $40^\circ\text{C}$ . The temperature difference of the heat exchangers in the cold box is taken as constraint with a minimum value of  $1^\circ\text{C}$ . Figure 3 shows the framework for the optimization. Besides, for the Solar-LAES-SE system, the  $RTE$  will increase monotonically with the increased input of solar thermal energy, since this heat comes at no expense. Therefore, there is no optimal value for the traditional  $RTE$  given by Equation (1) for the Solar-LAES-SE system. There are two ways to find an optimal condition for this system: (1) consider the solar energy as a heat input in the  $RTE$  calculation, as shown in Equation (2) for the revised  $RTE$ , and (2) consider the economic profit of the solar energy storage

system. Though the energy efficiency could be improved by a large amount of solar energy input, the investment cost of the molten salt and the storage tank would also be increased. Therefore, there will be an optimal value for the investment cost.

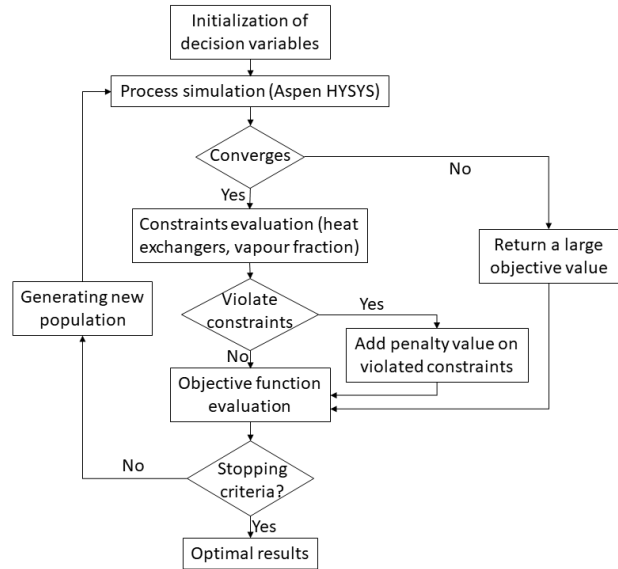


Figure 3. Optimization framework for the LAES systems.

## 5 RESULTS AND DISCUSSION

### 5.1 Energy Efficiency Optimization of the LAES-Stirling Engine System

The optimal  $RTE$  of the LAES-SE system is obtained with compression ratio, expansion ratio, and temperature, pressure and flowrate values for the hot and cold thermal energy recovery sections as decision variables. According to reference [12], the power generation efficiency of the Stirling engine is decreasing with increased outlet temperature of the heat sink. Therefore, the outlet temperature of the heat sink ( $C2'$  in Figure 1) is fixed at  $29^\circ\text{C}$  in order to ensure a reasonable temperature change of the heat sink. The outlet temperature of the heat source ( $H31$  in Figure 1) is fixed at  $30^\circ\text{C}$ , which is the lowest temperature when the minimum temperature difference between the heat source and sink is taken as  $10^\circ\text{C}$ .

For the Stirling engine based compression heat recovery section, the inlet and outlet temperatures of the heat sink and the outlet temperature of the heat source are fixed. Therefore, the power generated in the Stirling engine will mainly be impacted by the inlet temperature and flowrate of the heat source ( $H3$  in Figure 1). These two variables are determined by the compression ratio of the compressors, while the flowrate of the heat source will indirectly be affected by the expansion ratio of the expanders.

The optimization results show that the  $RTE$  of the LAES-SE system will be 68.2%, which is 3.2% points

higher than the LAES-ORC system in reference [4]. This means that the heat recovery efficiency of the Stirling engine is slightly higher than the ORC.

## 5.2 Energy Efficiency Optimization of the LNG-LAES-Stirling Engine System

For the LNG-LAES-SE system, the decision variables are the same as those in the LAES-SE system. For the Stirling engine based compression heat recovery section, the temperature of the heat sink is set to  $-162^{\circ}\text{C}$  since only latent heat from LNG is utilized in this case. The outlet temperature of the heat source is fixed at  $30^{\circ}\text{C}$ . In this case, one part of the thermal oil is used to preheat air before expanders and the other part is used to drive the Stirling engine. Minimum temperature difference of the air preheater is fixed at  $10^{\circ}\text{C}$ . The flowrate of thermal oil to the air preheating part will be determined by (i) the air temperature after evaporation and expansion (D3 and D4 in Figure 1) and (ii) the thermal oil temperature from the charging section. These temperatures depend on (i) the cold storage section and the expansion ratio and (ii) the compression ratio. This means that the flowrate of thermal oil to the Stirling engine also depends on the cold storage section and the expansion and compression ratios, since the total flowrate of thermal oil is fixed by the compression ratio. The optimized *RTE* is 73.7%, which is 8.7% points higher than the basic LAES-ORC system.

## 5.3 Energy Efficiency Optimization of the Solar-LAES-Stirling Engine System

For the Solar-LAES-SE system, as mentioned earlier, there is no optimal *RTE* value for the traditional definition in Equation (1). The *RTE* will be monotonically increasing with the solar energy input since this heat is not included in the traditional *RTE* definition. However, with the solar energy input increasing, the investment cost of the molten salt and the heat storage tank will increase. Therefore, there will be an optimal economic performance for the Solar-LAES-SE system. In addition, when solar energy is included in the *RTE* calculation, there will also be an optimal *RTE* value. Optimal results with respect to economic performance and revised *RTE* value are both studied.

### (1) Optimal revised *RTE* of the Solar-LAES-SE system

In the optimization of the revised *RTE* shown in Equation (2), the power generation efficiency of a Concentrated Solar Power plant (CSP) is taken as  $\eta = 0.214$  [13]. The solar energy input is determined by the cold energy needed in the charging section and the power generation efficiency of the Stirling engine, which is related to the temperature difference between the heat source and sink in the Stirling engine. The revised *RTE* value of the entire LAES-SE system is taken as the optimization objective. The compression and expansion ratios, the

temperature, pressure and flowrate of the hot and cold thermal energy recovery sections, and the molten salt flowrate and outlet temperature from the Stirling engine are taken as decision variables.

The optimization results show that the revised *RTE* will be 76.0% with a molten salt outlet temperature in the Stirling engine of  $500^{\circ}\text{C}$  and the inlet temperature fixed at  $550^{\circ}\text{C}$ . Combined with the density and specific heat capacity of the molten salt represented in Equations (14) and (15) [14], the optimized molten salt mass and volume rates are  $2.83 \times 10^6$  kg/h and  $1160$  m<sup>3</sup>/h. In this optimized condition, the traditional *RTE* is 189%. The optimized condition is related to the power generation efficiency value adopted for the CSP system. A larger power generation efficiency value of CSP will obviously increase the optimal *RTE* value.

$$\rho(\text{kg}/\text{m}^3) = 0.636 \cdot T(^{\circ}\text{C}) + 2089.905 \quad (14)$$

$$cp(\text{kJ}/\text{kg} \cdot \text{K}) = 1.723 \cdot 10^{-4} \cdot T(^{\circ}\text{C}) + 1.443 \quad (15)$$

where  $\rho$  refers to the density of the molten salt,  $cp$  refers to the specific heat capacity of the molten salt, and  $T$  refers to the temperature of the molten salt.

### (2) Optimal economic performance

In this case, the profit of the discharging section is taken as the optimization objective. In addition to the decision variables in the LAES-SE and LNG-LAES-SE systems, the flowrate and temperature of the molten salt are also taken as decision variables. The income from the power generated in the air expanders and Stirling engine and expenses related to the solar heat storage are considered in the profit of the discharging process. The cost of the binary molten salt is 0.8 €/kg and the specific cost of the storage tank is 510 €/m<sup>3</sup> [15]. The price of power is taken as 0.284 €/kWh [16], which was the average electricity price of Europe in 2022. The electricity price will have an obvious impact on the optimization results.

The optimization results show that the traditional and revised *RTE* of the Solar-LAES-SE system will be 173.0% and 72.8% respectively when the economic performance is optimal. In this condition, the outlet temperature of the molten salt in the Stirling engine is  $484^{\circ}\text{C}$  with the inlet temperature fixed at  $550^{\circ}\text{C}$ . The optimal mass and volume flowrates of the molten salt are obtained as  $2.42 \times 10^6$  kg/h and 990 m<sup>3</sup>/h.

### (3) Comparison of optimal revised *RTE* and economic performance

In order to further clarify the difference between the Solar-LAES-SE system under the optimal revised *RTE* and optimal economic performance, the optimal parameters of the corresponding systems are compared, as shown in Table 2. The lower and upper bounds of the air temperature before expanders in the two systems are  $170^{\circ}\text{C}$  and  $270^{\circ}\text{C}$ . The optimization results show that the

optimal air temperature before expanders for the revised  $RTE$  and economic performance is 180.2°C and 180.0°C, respectively. This indicates that a high temperature before the air expanders is not necessary to reach optimal efficiency or economic performance. As the heat sink of the Stirling engine, a higher air temperature before the expander will decrease the power generation efficiency in the Stirling engine due to the reduced temperature difference between the heat source and sink.

Table 2 also shows  $RTE$  values for the Solar-LAES-SE system (both traditional and revised  $RTE$ ). The traditional  $RTE$  ( $RTE_{LAES}$ ) of the entire Solar-LAES-SE system is higher than 100% because the solar energy input is not reflected in the calculation of the traditional  $RTE$ . The optimal revised  $RTE$  ( $RTE_{Solar-LAES}$ ) of the Solar-LAES-SE system is 3.2% points higher than that of the system with optimal economic performance.

**Table 2:** Optimized parameters for the Solar-LAES-SE system.

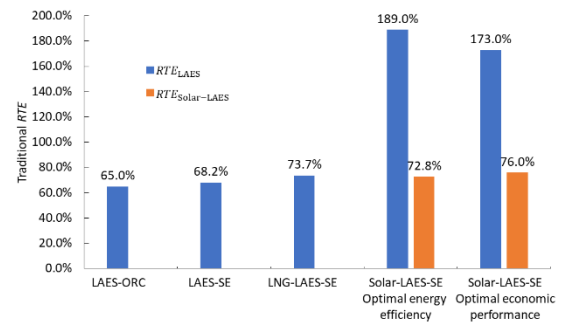
Items	Optimal economic performance	Optimal revised $RTE$	Units
$W_c$	$1.55 \times 10^4$	$1.31 \times 10^4$	kW
$Q_{solar}$	$1.00 \times 10^5$	$9.04 \times 10^4$	kW
$T_{Air}$	180.2	180.0	°C
$m_{Molten\ salt}$	$2.42 \times 10^6$	$2.83 \times 10^6$	kg/h
$V_{Molten\ salt}$	$9.90 \times 10^2$	$1.16 \times 10^3$	m <sup>3</sup> /h
$T_{Molten\ salt}$	484	500	°C
Molten salt cost	$2.71 \times 10^6$	$3.18 \times 10^6$	\$/h
$P_{LAES}$	$1.02 \times 10^4$	$9.43 \times 10^3$	kW
$P_{Stirling}$	$1.42 \times 10^4$	$1.32 \times 10^4$	kW
$P_{ORC}$ (compression heat)	$2.41 \times 10^3$	$1.98 \times 10^3$	kW
$RTE_{LAES}$	173.0	189.0	%
$RTE_{Solar-LAES}$	72.8	76.0	%
Profits	$3.02 \times 10^7$	$2.77 \times 10^7$	\$/h
Net income	$2.75 \times 10^7$	$2.45 \times 10^7$	\$/h

## 5.4 $RTE$ comparison of Different LAES Systems

Figure 4 shows the  $RTE$  comparison of different LAES systems including (i) stand-alone LAES system with ORC or Stirling engine as the surplus compression heat recovery cycles, (ii) LNG integrated LAES system with compression heat and the LNG regasification process as heat source and sink for the Stirling engine, and (iii) Solar energy integrated LAES system with Solar heat and the air discharging process as heat source and sink for the Stirling engine. For the solar energy based LAES system, there are two optimal  $RTE$  values. One for the optimal revised  $RTE$  and another for economic performance.

For the solar energy integrated systems, the traditional and revised  $RTE$ s provide different points of view to evaluate the energy efficiency. In the traditional  $RTE$ , the solar energy is taken as a free resource. For the revised  $RTE$ , the value of the solar energy is accounted for in the  $RTE$  according to the adopted power generation efficiency of the solar energy. In this way, the  $RTE$  value is adjusted to be lower than 100%. However, the definition of the  $RTE$  should be uniform when comparison with other LAES systems is performed, which is why the traditional  $RTE$  is also taken as a comparison index.

For the traditional  $RTE$ , the Solar-LAES-SE system optimizing on energy efficiency has the highest  $RTE$  value, followed by the Solar-LAES-SE system focusing on optimal economic performance, and the LNG-LAES-SE system focusing on energetic performance. This clearly indicates the high efficiency of the solar energy integrated systems. As shown in Table 2, using an economic objective, the net income (27.5 million \$) is 12.2% higher than when using an energetic objective (24.5 million \$).



**Figure 4.** Comparison of  $RTE$  values in LAES systems.

## 5.5 Comparison between Solar-LAES-ORC and Direct Solar Heating

The performance of the proposed Solar-LAES-SE system is compared with the Solar-LAES-ORC [5] and Solar directly heated-LAES [6] systems (with decoupled charging and discharging sections) shown in Figures 5 and 6. In the Solar-LAES-ORC system, both solar heat and compression heat are used to heat the discharging section and drive ORCs. In the Solar directly heated-LAES system, compression heat is used to drive the ORC while solar heat is used to preheat air in the discharging section.

Table 3 shows the economic and energy efficiency optimization results for the Solar-LAES-ORC system. The revised  $RTE$ s of the Solar-LAES-ORC system with optimal economic and energetic efficiency are 58.0% and 61.3% respectively, which are considerably lower than those of the Solar-LAES-SE system. A similar result (i.e. lower values) is obtained for the net income of the Solar-LAES-ORC system. Therefore, the Solar-LAES-SE system is superior to the Solar-LAES-ORC system in terms of both energy and economy.

**Table 3:** Optimized parameters for the Solar-LAES-ORC system.

Items	Optimal economic performance	Optimal revised RTE	Units
$W_c$	$1.26 \times 10^4$	$1.26 \times 10^4$	kW
$Q_{solar}$	$4.16 \times 10^4$	$2.55 \times 10^4$	kW
$T_{Air}$	183.0	183.0	°C
$M_{Thermal\ oil}$	$1.82 \times 10^5$	$1.12 \times 10^5$	kg/h
$V_{Thermal\ oil}$	$2.01 \times 10^2$	$1.24 \times 10^2$	m <sup>3</sup> /h
Thermal oil cost	$5.66 \times 10^5$	$3.48 \times 10^5$	\$/h
$P_{LAES}$	$8.61 \times 10^3$	$8.58 \times 10^3$	kW
$P_{ORC}$ (solar heat)	$2.71 \times 10^3$	$1.35 \times 10^3$	kW
$P_{ORC}$ (compression heat)	$1.14 \times 10^3$	$1.14 \times 10^3$	kW
$RTE_{LAES}$	98.9	87.9	%
$RTE_{Solar-LAES}$	58.0	61.3	%
Profits	$1.40 \times 10^7$	$1.24 \times 10^7$	\$/h
Net income	$1.34 \times 10^7$	$1.21 \times 10^7$	\$/h

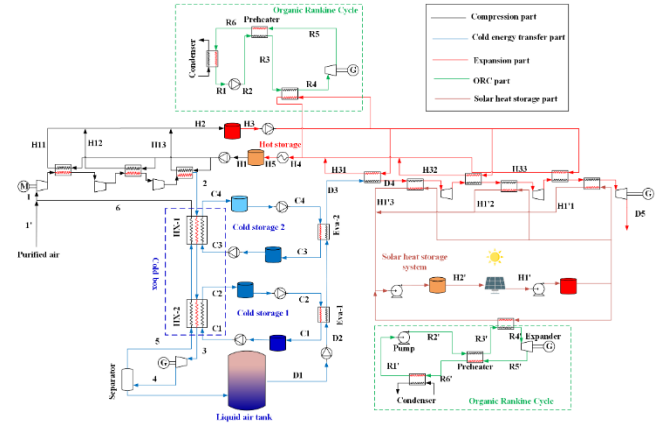
**Table 4:** Optimized parameters for the Solar directly heated ORC system.

Items	Optimal economic performance	Optimal revised RTE	Units
$W_c$	$1.52 \times 10^4$	$1.22 \times 10^4$	kW
$Q_{solar}$	$1.31 \times 10^4$	$1.21 \times 10^4$	kW
$T_{Air}$	183.0	183.0	°C
$M_{Thermal\ oil}$	$5.47 \times 10^4$	$5.06 \times 10^4$	kg/h
$V_{Thermal\ oil}$	60.4	55.8	m <sup>3</sup> /h
Thermal oil cost	$1.70 \times 10^5$	$1.57 \times 10^5$	\$/h
$P_{LAES}$	$9.26 \times 10^3$	$8.39 \times 10^3$	kW
$P_{ORC}$ (compression heat)	$1.72 \times 10^3$	$1.71 \times 10^3$	kW
$RTE_{LAES}$	72.1	82.6	%
$RTE_{Solar-LAES}$	60.9	68.2	%
Profits	$1.23 \times 10^7$	$1.14 \times 10^7$	\$/h
Net income	$1.22 \times 10^7$	$1.12 \times 10^7$	\$/h

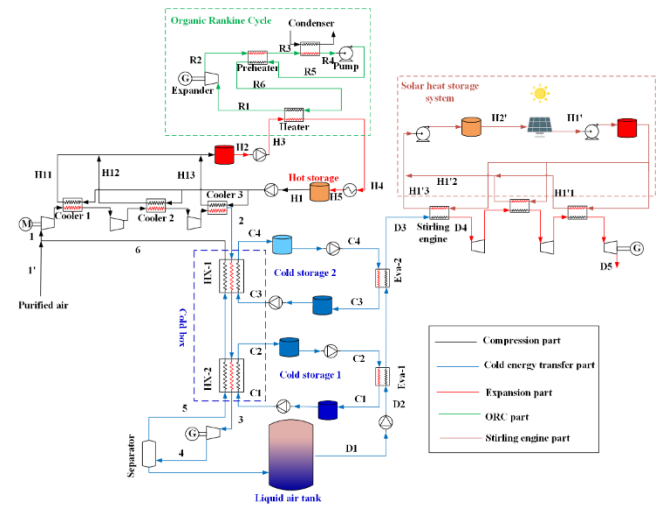
Table 4 shows the optimal results for the solar energy directly heated LAES system. The revised  $RTEs$  of the system with respect to optimal economy and energy efficiency are 60.9% and 68.2%, which are lower than those of the Solar-LAES-SE system but higher than those of the Solar-LAES-ORC system. As for the economic performance, the net income of the Solar energy directly heated-LAES system is the lowest among all solar energy integrated LAES systems.

As a conclusion, for the solar energy integrated

LAES systems, the Solar-LAES-SE system is the best both from the energy efficiency and economic performance point of view. The energy efficiency ranking of the solar energy integrated LAES systems from high to low is Solar-LAES-SE>Solar directly heated LAES>Solar-LAES-ORC, while the economic performance ranking is Solar-LAES-SE>Solar-LAES-ORC>Solar directly heated LAES.



**Figure 5.** Flowsheet of the Solar-LAES-ORC system.



**Figure 6.** Flowsheet of the Solar energy directly heated-LAES system.

## 6 CONCLUSIONS

Different LAES systems integrated with a Stirling engine are modeled and optimized from an energy efficiency and an economic point of view. The following conclusions have been obtained.

- (1) The optimized  $RTE$  of the LAES-SE system is 68.2%, 3.2% points higher than the basic LAES-ORC system. This means that the Stirling engine is more efficient than the ORC in recovering surplus compression heat in the 3C+3E LAES system.



(2) When cooling water is replaced by LNG regasification as a heat sink for the Stirling engine, the optimized *RTE* increases from 68.2% to 73.7%.

(3) For the solar energy integrated system, the traditional *RTE* will be 189% when optimizing with respect to the revised *RTE* and 173% when optimizing with respect to economic performance of the discharging section. Having efficiencies larger than 100% comes as a result of the fact that the basic LAES system (for which the traditional *RTE* is defined) benefits from "free" energy input.

(4) Among the Stirling engine based systems, the Solar-LAES-SE system under the optimal revised *RTE* has the highest energy efficiency, while the LAES-SE system has the lowest energy efficiency. All the studied systems have higher *RTEs* than the basic LAES-ORC system.

(5) For the solar energy integrated LAES systems, Solar-LAES-SE has the highest optimal revised *RTE*, followed by the Solar directly heated LAES-system, while Solar-LAES-ORC has the lowest optimal revised *RTE*. In the case of economic performance, the ranking is Solar-LAES-SE > Solar-LAES-ORC > Solar directly heated LAES.

## ACKNOWLEDGEMENTS

Financial support from the China Scholarship Council (No. 202206280152) and the Research Council of Norway and user partners of HighEFF, an 8-years Research Centre under the FME-scheme (Centre for Environment-friendly Energy Research, 257632), are greatly acknowledged.

## REFERENCES

1. Guizzi GL, Manno M, Tolomei LM, Vitali RM. Thermodynamic analysis of a liquid air energy storage system. *Energy*, 93:1639-1647(2015).
2. Liu Z, Kim D, Gundersen T. Optimal recovery of thermal energy in liquid air energy storage. *Energy*, 240:122810(2022).
3. Liu Z, Gundersen T. Liquid Air Energy Storage - Optimization opportunities. AICHE Annual Meeting, Arizona, 13-18 November 2022.
4. Liu Z, Kim D, Gundersen T. Optimization and analysis of different liquid air energy storage configurations. *Computers & Chemical Engineering*, 169:108087(2023).
5. Ding X, Duan L, Zhou Y, Gao C, Bao Y. Energy, exergy, and economic analyses of a new liquid air energy storage system coupled with solar heat and organic Rankine cycle. *Energy Conversion and Management*, 266(2022).
6. Yang M, Duan L, Tong Y, Jiang Y. Study on design optimization of new liquified air energy storage (LAES) system coupled with solar energy. *Journal of Energy Storage*, 51(2022).
7. Ding H, Li J, Heydarian D. Energy, exergy, exergoeconomic, and environmental analysis of a new biomass-driven cogeneration system. *Sustainable Energy Technologies and Assessments*, 45:101044(2021).
8. Kongtragool B, Wongwiset S. Investigation on power output of the gamma-configuration low temperature differential Stirling engines. *Renewable Energy*, 30(3):465-476(2005).
9. Mehrpooya M, Sayyad S, Zonouz MJ. Energy, exergy and sensitivity analyses of a hybrid combined cooling, heating and power (CCHP) plant with molten carbonate fuel cell (MCFC) and Stirling engine. *Journal of Cleaner Production*, 148: 283-294 (2017).
10. Ansarinassab H, Hajabdollahi H. Multi-objective optimization of a geothermal-based multigeneration system for heating, power and purified water production purpose using evolutionary algorithm. *Energy Conversion and Management*, 223:113476(2020).
11. Blanco J, Alarcón D, Sánchez B, Malato S, Maldonado MI. Technical comparison of different solar-assisted heat supply systems for a multi-effect seawater distillation unit. *Solar Energy for a Sustainable Future*:14-19(2003).
12. Ren S, Gundersen T, Liu Z, Feng X. Performance improvement of liquid air energy storage: Introducing Stirling engine and solar energy. *Energy Conversion and Management*, 296(2023).
13. Awan AB, Zubair M, Memon ZA, Ghaleb N, Tlili I. Comparative analysis of dish Stirling engine and photovoltaic technologies: Energy and economic perspective. *Sustainable Energy Technologies and Assessments*, 44(2021).
14. Peiró G, Gasia J, Miró L, Prieto C, Cabeza LF. Influence of the heat transfer fluid in a CSP plant molten salts charging process. *Renewable Energy*, 113:148-158(2017).
15. Sau S, Corsaro N, Crescenzi T, D'Ottavi C, Liberatore R, Licoccia S, et al. Techno-economic comparison between CSP plants presenting two different heat transfer fluids. *Applied Energy*, 168:96-109(2016).  
Electricity prices for household consumers. In: Eurostat, 2022.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# The Impact of Electrified Process Heating on Process Design, Control and Operations

Jong Hyun Rho<sup>a</sup>, Michael Baldea<sup>a,b\*</sup>, Elizabeth E. Endler<sup>c</sup>, Monica A. Heredia<sup>c</sup>, Vesna Bojovic<sup>c</sup> and Pejman Pajand<sup>c</sup>

<sup>a</sup> The University of Texas at Austin, McKetta Department of Chemical Engineering, Austin, TX, USA

<sup>b</sup> The University of Texas at Austin, Oden Institute for Computational Engineering and Sciences, Austin, TX, USA

<sup>c</sup> Shell International Exploration and Production, Houston, TX, USA

\* Corresponding Author: mbaldea@che.utexas.edu.

---

## ABSTRACT

We study the impact of switching from combustion heating to electric heating in processes comprising high temperature reaction/separation sequences, where the heat supporting the reaction(s) is substantially provided by combusting a reaction byproduct (fuel gas). A canonical process structure is defined. It is shown that the conventional combustion-based process presents significant interactions. An asymptotic analysis is utilized to investigate and compare the dynamic responses of the conventional and electric process configurations. It is demonstrated that the dynamic behavior of the two processes exhibits two timescales, with the faster corresponding to the evolution of the temperatures of the units with high heat duty, and the slow time scale capturing the variables involved in the material balance. A simplified ethylene cracking process example is used to demonstrate these findings.

---

Keywords: Energy Systems, Process Design, Process Electrification

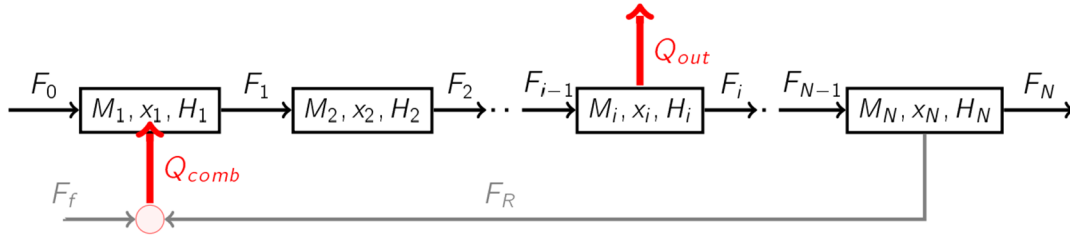
## INTRODUCTION

Process electrification using electricity generated by renewable sources has emerged as one of the routes for decarbonizing manufacturing processes [1,2]. Electrification may involve a complete change of processing technology (e.g., utilizing electrochemical reactions) [3] or upgrades and modifications of existing technology. Of particular interest in the latter area is the electrification of process heating, whereby existing combustion heaters are replaced with electric heating technologies [1,2].

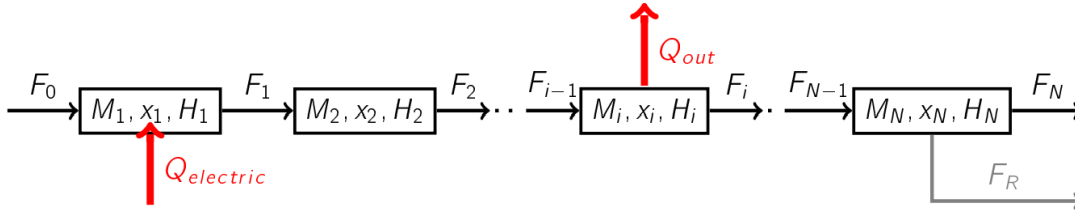
Electrification can in principle eliminate combustion-heating-related CO<sub>2</sub> emissions and has the potential co-benefits of affording more precise/localized heating and providing additional degrees of freedom for operation and control [4]. Nevertheless, it presents several challenges. From a design perspective, electric heating may disrupt existing process integration structures, which rely on, e.g., combusting process-generated waste streams ("tail gas" or "fuel gas") to generate heat [5]. From an operation and control perspective, electric heating must account for the availability of renewable electricity, which

may fluctuate during the day [6]. Motivated by the above, in this work, we provide a rigorous analysis of the design and process dynamics implications of electrified process heating in the context of integrated process systems. We note that there are other sources of CO<sub>2</sub> emissions in chemical/petrochemical processing (including scope 2) emissions, that are not accounted for in this work.

A prototype process structure with reaction, separation, and recycle is defined. It is shown that the conventional combustion-based process presents significant interactions due to the impact of the downstream units (via the heating value of the fuel gas) on the upstream units. An asymptotic analysis is utilized to investigate and compare the dynamic responses of the two process configurations. It is demonstrated that the dynamic behavior of the two processes exhibits two time scales, with the faster corresponding to the evolution of the temperatures of the units with high heat duty, and the slow time scale capturing the variables involved in the material balance. A simplified ethylene cracking process example is used to demonstrate these findings.



**Figure 1:** Structure of integrated process using combustion of fuel gas as the heat source. Black lines denote material streams connecting process units, grey lines denote streams of material used as fuel, red lines denote energy flow.



**Figure 2:** Structure of process with electric heating.

## PROTOTYPE PROCESS

A system of  $N$  process units, each with mass holdup  $M$ , mass fraction  $x$ , and mass specific enthalpy  $H$ , connected in series is considered. The feed stream, of mass flow rate  $F_0$  has mass specific enthalpy  $H_0$  and contains a reactant with mass fraction  $x_0$ . The reactant undergoes a high-temperature, endothermic transformation in the first unit. Processing units  $2, \dots, i-1$  follow, and heat is eventually removed/recovered from the resulting process stream in unit  $i$ . Finally, the product stream is obtained in unit  $N$ , along with a waste stream of flow rate  $F_R$ . The waste (“off gas”, “tail gas” or “fuel gas”) stream can be used as a fuel but does not have any economic value. Two process configurations are considered:

- A conventional, integrated design (Figure 1), whereby the majority of the heat required for the endothermic transformation occurring in Unit 1 is provided by combusting the fuel stream (possibly combined with a fresh fuel stream of mass flow rate  $F_f$ ). The rate of heat generation by combustion is given by

$$Q_{comb} = \lambda_f F_f + \lambda_R F_R \quad (1)$$

where  $\lambda_f$  and  $\lambda_R$  refer to the heating values of the fresh fuel and the fuel gas, respectively. Combustion of fuel gas occurs in air, and a flue gas stream is generated (not shown in the figure).

- An electrified process configuration (Figure 2), whereby the heat required for the endothermic transformation in Unit 1 is provided by electric

heating. In this case, other means to process the waste gas stream must be identified.

The structure of the two process configurations shown in Figures 1 and 2 is the same as far as material processing is concerned. From a steady state design point of view, it is important to ensure that the product streams resulting from the two configurations are the same (in terms of composition  $x$  and enthalpy  $H$ ), and that the amount of heat recovered  $Q_{out}$  is the same. To this end, the rates of energy input to both processes must be the same, and the following condition must be met:

$$\bar{Q}_{comb} = \bar{Q}_{elec} \quad (2)$$

where the overbar denotes steady-state values. Note that the above condition assumes that heating is 100% efficient, that is, all the heat generated by combustion (as defined in equation (1)) or by electricity is transferred to the process. Heat transfer efficiency is lower in practice.

## DYNAMIC ANALYSIS AND TIME SCALE DECOMPOSITION

The dynamic model of the process is based on the following assumptions: each unit is well-mixed, has constant holdups, constant physical properties, and has no heat loss. With these assumptions, the model of the conventional process shown in Figure 1 can be written as:

Unit 1 :

$$\frac{dx_1}{dt} = \frac{1}{M_1} [F_0 x_0 - F_1 x_1 + f_1(x_1, H_1)]$$

$$\frac{dH_1}{dt} = \frac{1}{M_1} [F_0 H_0 - F_1 H_1 + \lambda_f F_f + \lambda_R F_R + g_1(x_1, H_1)]$$

Unit 2 :

$$\frac{dx_2}{dt} = \frac{1}{M_2} [F_1 x_1 - F_2 x_2 + f_2(x_2, H_2)]$$

$$\frac{dH_2}{dt} = \frac{1}{M_2} [F_1 H_1 - F_2 H_2 + g_2(x_2, H_2)]$$

⋮ (3)

Unit  $i$  :

$$\frac{dx_i}{dt} = \frac{1}{M_i} [F_{i-1} x_{i-1} - F_i x_i + f_i(x_i, H_i)]$$

$$\frac{dH_i}{dt} = \frac{1}{M_i} [F_{i-1} H_{i-1} - F_i H_i - Q_{out} + g_i(x_i, H_i)]$$

Unit  $N$  :

$$\frac{dx_N}{dt} = \frac{1}{M_N} [F_{N-1} x_{N-1} - F_N x_N - F_R x_R + f_N(x_N, H_N)]$$

$$\frac{dH_N}{dt} = \frac{1}{M_N} [F_{N-1} H_{N-1} - F_N H_N - F_R H_R + g_N(x_N, H_N)]$$

The model of the electrified process in Figure 2 can be written in a similar fashion, replacing the sum  $\lambda_f F_f + \lambda_R F_R$  (which represents  $Q_{comb}$ ) in the energy balance of Unit 1 in equation (3) with  $Q_{elec}$ .

In order to proceed with the analysis, the following scaled entities are defined:  $u_j = \frac{F_j H_j}{F_j \bar{H}_j}$ ,  $j = 1, \dots, N$ ,  $u_r = \frac{F_R H_R}{F_R \bar{H}_R}$ ,  $u_q = \frac{Q_{out}}{Q_{out}}$ .

Then, the following assumptions are made, that pertain to steady state operation (the overbar denotes again steady state values):

- Let  $k_1 = \frac{F_1 \bar{H}_1}{F_R \bar{\lambda}_R}$ , and assume that  $k_1 = \mathcal{O}(1)$ , meaning that the amount of energy leaving unit 1 via flow is of comparable magnitude to the energy provided by combusting the waste gas.
- Let  $k_0 = \frac{\bar{F}_f \bar{\lambda}_f}{\bar{F}_0 \bar{H}_0}$ , and assume that  $k_0 = \mathcal{O}(1)$ , meaning that the amount of heat provided by combusting fresh fuel is small and comparable with the amount of heat provided by convection by the feed stream.
- Let  $\varepsilon = \frac{F_0 \bar{H}_0}{\bar{F}_R \bar{\lambda}_R} \ll 1$ , meaning that the energy input from the feed stream is small compared with the amount of energy provided by combusting the fuel gas.
- Let  $k_j = \frac{F_j \bar{H}_j}{F_R \bar{\lambda}_R}$ ,  $j = 1, \dots, i-1$ , and assume that  $k_j = \mathcal{O}(1)$ , meaning that the amount energy leaving unit  $j = 1, \dots, i-1$  via flow is of comparable magnitude to the energy provided by combusting the waste gas stream.
- Let  $I_j = \frac{F_j \bar{H}_j}{\bar{F}_f \bar{\lambda}_f}$ ,  $j = i, \dots, N$ , and assume that  $I_j = \mathcal{O}(1)$ , meaning that the amount energy leaving unit  $j = i, \dots, N$  via flow is of comparable magnitude to the

energy provided by combusting fresh fuel feed.

- Let  $I_p = \frac{\bar{F}_R \bar{H}_N}{\bar{F}_f \bar{\lambda}_f}$ , and assume that  $I_p = \mathcal{O}(1)$ , meaning that the amount of energy contained in the waste stream is of comparable magnitude to the amount of heat provided by combusting fresh fuel feed.

With these definitions and assumptions, the model becomes:

Unit 1 :

$$\frac{dx_1}{dt} = \frac{1}{M_1} [F_0 x_0 - F_1 x_1 + f_1(x_1, H_1)]$$

$$\frac{M_1}{\bar{F}_0 \bar{H}_0} \frac{dH_1}{dt} = u_0 + u_r k_0 + \frac{g_1(x_1, H_1)}{\bar{F}_0 \bar{H}_0} + \frac{1}{\varepsilon} (u_r - k_1 u_1)$$

Unit 2 :

$$\frac{dx_2}{dt} = \frac{1}{M_2} [F_1 x_1 - F_2 x_2 + f_2(x_2, H_2)]$$

$$\frac{M_2}{\bar{F}_0 \bar{H}_0} \frac{dH_2}{dt} = \frac{g_2(x_2, H_2)}{\bar{F}_0 \bar{H}_0} + \frac{1}{\varepsilon} (k_1 u_1 - k_2 u_2)$$

⋮ (4)

Unit  $i$  :

$$\frac{dx_i}{dt} = \frac{1}{M_i} [F_{i-1} x_{i-1} - F_i x_i + f_i(x_i, H_i)]$$

$$\frac{M_i}{\bar{F}_0 \bar{H}_0} \frac{dH_i}{dt} = -I_i k_0 u_i + \frac{g_i(x_i, H_i)}{\bar{F}_0 \bar{H}_0} + \frac{1}{\varepsilon} (k_{i-1} u_{i-1} - k_i u_i)$$

Unit  $N$  :

$$\frac{dx_N}{dt} = \frac{1}{M_N} [F_{N-1} x_{N-1} - F_N x_N - F_R x_R + f_N(x_N, H_N)]$$

$$\frac{M_N}{\bar{F}_0 \bar{H}_0} \frac{dH_N}{dt} = I_{N-1} k_0 u_{N-1} - I_N k_0 u_N - I_p k_0 u_p + \frac{g_N(x_N, H_N)}{\bar{F}_0 \bar{H}_0}$$

This model is a singularly perturbed system of ordinary differential equations in standard form, and it is expected to have a dynamic behavior featuring two time-scales<sup>7</sup>. The framework presented by Baldea and Daoutidis [7] is employed to study the dynamic behavior via an asymptotic analysis using singular perturbation arguments.

Its dynamic behavior is analyzed below, starting from the fast timescale. To this end, a new "stretched" time variable is defined as:

$$\tau = \frac{1}{\varepsilon} \quad (5)$$

With this, the model (4) becomes:

Unit 1 :

$$\frac{dx_1}{d\tau} = \varepsilon \frac{1}{M_1} [F_0 x_0 - F_1 x_1 + f_1(x_1, H_1)]$$

$$\frac{M_1}{\bar{F}_0 \bar{H}_0} \frac{dH_1}{d\tau} = \varepsilon [u_0 + u_r k_0 + \frac{g_1(x_1, H_1)}{\bar{F}_0 \bar{H}_0}] + u_r - k_1 u_1$$

Unit 2 :

$$\frac{dx_2}{d\tau} = \varepsilon \frac{1}{M_2} [F_1 x_1 - F_2 x_2 + f_2(x_2, H_2)]$$

$$\begin{aligned} \frac{M_2}{F_0 \bar{H}_0} \frac{dH_2}{d\tau} &= \varepsilon \frac{g_2(x_2, H_2)}{F_0 \bar{H}_0} + k_1 u_1 - k_2 u_2 \\ &\vdots \end{aligned} \quad (6)$$

Unit  $i$  :

$$\begin{aligned} \frac{dx_i}{d\tau} &= \varepsilon \frac{1}{M_i} [F_{i-1} x_{i-1} - F_i x_i + f_i(x_i, H_i)] \\ \frac{M_i}{F_0 \bar{H}_0} \frac{dH_i}{d\tau} &= \varepsilon [-I_i k_0 u_i + \frac{g_i(x_i, H_i)}{F_0 \bar{H}_0}] + k_{i-1} u_{i-1} - k_q u_q \end{aligned}$$

Unit  $N$  :

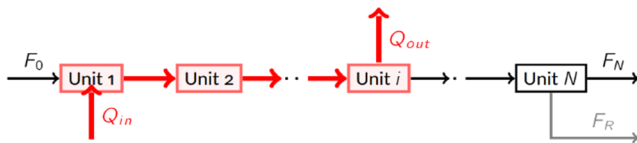
$$\begin{aligned} \frac{dx_N}{dt} &= \varepsilon \frac{1}{M_N} [F_{N-1} x_{N-1} - F_N x_N - F_R x_R + f_N(x_N, H_N)] \\ \frac{M_N}{F_0 \bar{H}_0} \frac{dH_N}{dt} &= \varepsilon [I_{N-1} k_0 u_{N-1} - I_N k_0 u_N - I_\rho k_0 u_\rho + \frac{g_N(x_N, H_N)}{F_0 \bar{H}_0}] \end{aligned}$$

An expression of the dynamics in the fast time scale can be obtained by considering the limit case  $\varepsilon \rightarrow 0$ , which corresponds to the limit where the amount of heat obtained from combusting the tail gas is infinitely higher than the amount of heat derived from the combustion of fresh fuel. The fast dynamics thus take the form:

$$\begin{aligned} \text{Unit 1 :} \quad \frac{dx_1}{d\tau} &= 0 \\ \frac{M_1}{F_0 \bar{H}_0} \frac{dH_1}{d\tau} &= u_r - k_1 u_1 \\ \text{Unit 2 :} \quad \frac{dx_2}{d\tau} &= 0 \\ \frac{M_2}{F_0 \bar{H}_0} \frac{dH_2}{d\tau} &= k_1 u_1 - k_2 u_2 \\ &\vdots \end{aligned} \quad (7)$$

$$\begin{aligned} \text{Unit } i : \quad \frac{dx_i}{d\tau} &= 0 \\ \frac{M_i}{F_0 \bar{H}_0} \frac{dH_i}{d\tau} &= k_{i-1} u_{i-1} - k_q u_q \\ \text{Unit } N : \quad \frac{dx_N}{d\tau} &= 0 \\ \frac{M_N}{F_0 \bar{H}_0} \frac{dH_N}{d\tau} &= 0 \end{aligned}$$

The system of equations in (7) suggests that the variables whose dynamic response exhibits a fast component are the mass specific enthalpies (equivalently, temperatures)  $H_1, \dots, H_i$  of units  $1, \dots, i$ . Based on the assumptions made above, Unit 1 has a significant rate of heat input (via either combustion or electric heating). Heat is then conveyed to units  $2, \dots, i$  until it is removed from Unit  $i$  at a rate  $Q_{out}$ . Thus, units  $1, \dots, i$  represent a high energy throughput path through the process (Figure 3) [8].



**Figure 3** The fast component of the dynamics captures the energy balance of the units located in the high energy

throughput pathway of the process.  $Q_{in}$  represents either  $Q_{comb}$  or  $Q_{elec}$

The above arguments suggest that  $\bar{Q}_{out}$  is of comparable magnitude to  $\bar{Q}_{in}$ ; in the case where endothermic reactions occur in units  $1, \dots, i$ , the overall energy balance of the process suggests that the two quantities would differ by the heat consumed by the reactions (as reflected in the terms  $g_i(x_i, H_i)$ ).

The dynamics in the fast time scale are influenced by the rate of energy input to the process (as reflected by the term  $u_r$  in equation (7)), the rate of energy removal (term  $k_q u_q$  in equation (7)), and are not influenced by the rate of raw material input, the production rate and the rate at which fuel is provided from external sources.

It is of note that, even though the dynamics of the variables in the energy balance of units  $1, \dots, i$  are fast, the dynamic response of the individual units may be different; in other words, the time constants  $\frac{M_j}{F_0 \bar{H}_0}$ ,  $j = 1, \dots, i$  in equations (7) may be quite different.

The slow component of the dynamics evolves in time scale  $t$ . Considering the same limit  $\varepsilon \rightarrow 0$  in the original (slow) time scale  $t$  gives rise to the following constraints:

$$0 = u_r - k_1 u_1 \quad (8a)$$

$$0 = k_1 u_1 - k_2 u_2 \quad (8b)$$

$\vdots$

$$0 = k_{i-1} u_{i-1} - k_q u_q \quad (8c)$$

These constraints correspond to a quasi-steady state of the fast dynamics and describe a manifold/subspace where the slow dynamics of the process evolve. By considering these constraints and substituting (8) in equation (6), the slow component of the dynamics can be described as:

Unit 1 :

$$\frac{dx_1}{dt} = \frac{1}{M_1} [F_0 x_0 - F_1 x_1 + f_1(x_1, H_1)]$$

$$\frac{M_1}{F_0 \bar{H}_0} \frac{dH_1}{dt} = u_0 + u_r k_0 + \frac{g_1(x_1, H_1)}{F_0 \bar{H}_0}$$

Unit 2 :

$$\frac{dx_2}{dt} = \frac{1}{M_2} [F_1 x_1 - F_2 x_2 + f_2(x_2, H_2)]$$

$$\frac{M_2}{F_0 \bar{H}_0} \frac{dH_2}{dt} = \frac{g_2(x_2, H_2)}{F_0 \bar{H}_0}$$

$\vdots$

(9)

Unit  $i$  :

$$\frac{dx_i}{dt} = \frac{1}{M_i} [F_{i-1} x_{i-1} - F_i x_i + f_i(x_i, H_i)]$$

$$\frac{M_i}{F_0 \bar{H}_0} \frac{dH_i}{dt} = -I_i k_0 u_i + \frac{g_i(x_i, H_i)}{F_0 \bar{H}_0}$$

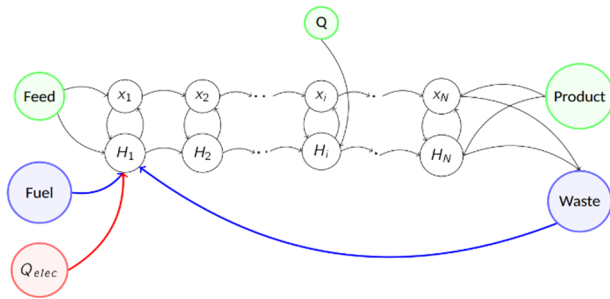


Unit  $N$  :

$$\frac{dx_N}{dt} = \frac{1}{M_N} [F_{N-1}x_{N-1} - F_Nx_N - F_Rx_N + f_N(x_N, H_N)]$$

$$\frac{M_N}{\bar{F}_0\bar{H}_0} \frac{dH_N}{dt} = I_{N-1}k_0u_{N-1} - I_Nk_0u_N - I_\rho k_0u_\rho + \frac{g_N(x_N, H_N)}{\bar{F}_0\bar{H}_0}$$

The model of the slow dynamics describes the evolution of the variables in the material balance (notably the product composition  $x_N$ , influenced by the product flow rate  $F_0$ ), as well as the evolution of the slow component of the energy balance of the process.



**Figure 4** Directed graph of the dynamic model of the prototype process. Nodes and edges specific to the integrated process using combustion of fuel gas as the heat source are shown in blue. Nodes and edges specific to electric heating are shown in red. The green and black nodes/edges are common to the two digraphs.

## OPERATIONAL IMPLICATIONS OF ELECTRIFICATION

The implications of the analysis presented earlier are discussed in this section. The presence of a two time scale behavior suggests that these processes lend themselves quite naturally to a two-tiered control and operational decision-making structure.<sup>7</sup> The management of energy use (including temperature control of Units 1, ...,  $i$ ) should be pursued in the fast time scale, while the control and management of production (in terms of production rate, product purity) should be pursued in the slow time scale [7]. The latter can also involve production scheduling.

The fast dynamics (7) of Units 1, ...,  $i$  means that changes in the rate of heat input to Unit 1 will be quickly reflected in the temperatures and compositions of units 2, ...,  $i$ . The impact on units  $i + 1, \dots, N$  will be apparent more slowly. In addition to deliberate operator (or control intervention), changes in the heat rate to Unit 1 can be caused by disturbances and constraints. The origin of these may be exogenous (from outside the process) or endogenous (from within).

To further investigate the impact of these disturbances and constraints on the heat rate, the structure of the two prototype processes is considered. Figure 4

shows the directed graph (digraph) of the system model (3). The fundamental difference between the conventional and electrified processes lies in the presence of a cycle in the graph of the conventional process, reflecting the fact that the heat rate provided to the Unit 1 depends on the composition  $x_N$  of the waste stream. Consider equation (7), and specifically the term  $u_r$ , that reflects the energy input due to combustion of the tail gas. The definition of  $u_r$ , i.e.,  $u_r = \frac{F_R H_R}{F_R \lambda_R}$ , indicates that rate of energy input is a function of  $\lambda_R$ , the lower heating value of the waste gas. The lower heating value is a function of the composition of the waste gas, and hence of the variable  $x_R$ .  $x_R$  is, in turn, affected by (effectively any and all) disturbances affecting the process, from Unit 1 to Unit  $N$ . Thus, a feedback effect is present, whereby variables in the *downstream* sections of the process impact the *upstream* units. Additionally, the dynamics of  $x_R$  evolve exclusively in the slow time scale, meaning that this feedback effect will occur over a long-time horizon.

In the process with electric heating, the heat rate to Unit 1 is completely independent of the downstream units of the process and constitutes an additional degree of freedom for controlling/optimizing the process operation. On the other hand, given the goal of using *renewable* electricity for process heating implies that the rate of heat input may be subject to upper bounds that are lower than the nominal steady state value.

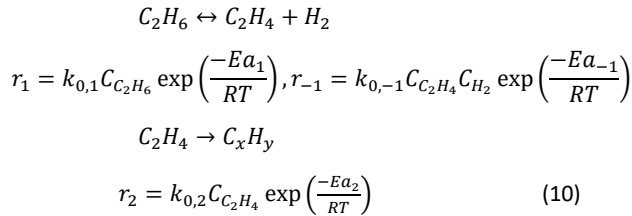
## CASE STUDY: ETHYLENE CRACKER

Thermal cracking of hydrocarbon feedstock to obtain ethylene is one of the most energy-intensive chemical processes and one of the largest carbon emitters in the chemical industry. Its decarbonization potential lies in replacing the conventional combustion-based furnace with an electric furnace [9]. Relevant research efforts are underway in both industry and academia [10,11].

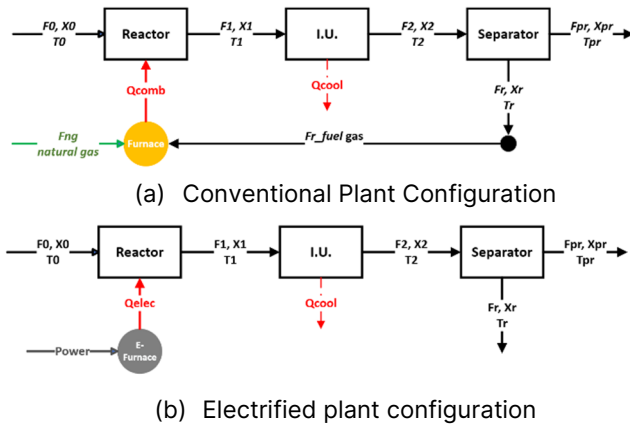
Ethylene crackers are highly integrated, and the by-products of the cracking reactions (methane, hydrogen) constitute an important fraction of the fuel used to heat the cracking reactor (which operates at high temperatures, of around 900°C) [5,12], and the heating value of the fuel gas stream depends on its composition (and ultimately on the operation of the reactor, as argued above).

Ethylene plants are complex and comprise extensive separation sections. We use a simplified representation for the purpose of investigating the dynamic implications of electrified process heating. The model consists of three units: reactor, intermediate unit, and separator, with the latter two serving as a proxy for the separation section. The dynamics of each unit are represented by the material and energy balances, under similar assumptions as listed in developing the model (3), with the exception that holdups are not constant.

The reactor unit is fed with pure ethane. The complex reactions occurring in industrial furnaces are approximated using a simplified two reaction scheme, a reversible reaction emulating the cracking transformation, and the degradation of product ethylene; we utilize methane as a proxy of the hydrocarbon degradation products:



with  $r_m$  referring to reaction  $m = 1, 2$  and  $-1$ .



**Figure 5.** Structure of the ethylene cracker prototype model

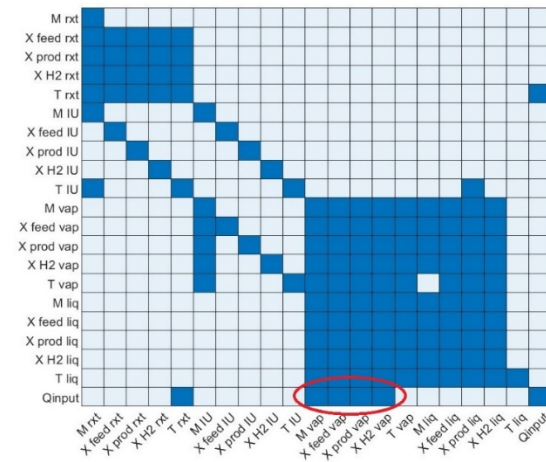
The product stream consists of unreacted ethane, product ethylene, and by-products methane and hydrogen. The reactor is modeled as a CSTR with the appropriate material and energy balances.

The intermediate units (I.U.) are intended to capture the (largely physical) transformations occurring post cracking, including the product quench step (represented as heat exiting the block  $Q_{cool}$ ). The material and energy balances follow first order dynamics. A significantly larger holdup is used in the unit compared to that of the reactor, to reflect the large time constant of the separation section of an ethylene plant.

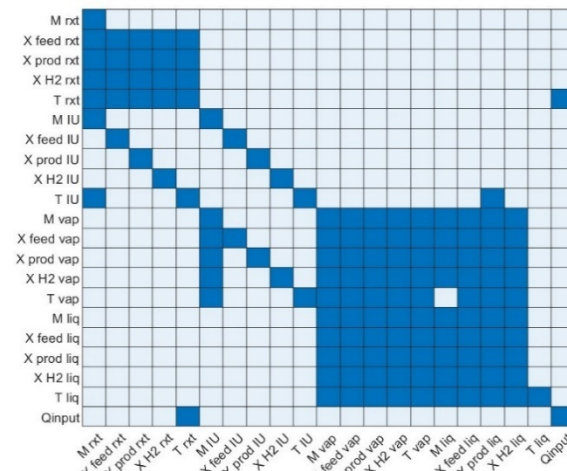
In the separation block, by-products hydrogen and methane are separated from the product (in this case, ethane and ethylene) and recovered to use as fuel. The unit is modeled as a rate-based separation with vapor and liquid holdups, with by-product fuel gas exiting as the vapor stream and ethylene and unreacted ethane exiting as the liquid product stream.

In the conventional system, the fuel gas is recycled to be combusted in the furnace, providing heat for the

cracking reaction. The fuel gas composition, which determines the lower heating value (computed as a linear combination of individual heat values multiplied by their compositions in the fuel gas,  $LHV = \sum_j X_{vap,j} LHV_j$ ), and the corresponding energy input to the reactor, depends on the performance of the reactor and separator. On the other hand, in the electrified system case, the energy is supplied to the reactor from a source outside the process and thus independent of the performance of the process itself.



Structure of matrix A for the integrated process



Structure of matrix A for the electrified process

**Figure 6.** Structure of system matrix A of the linearized process model

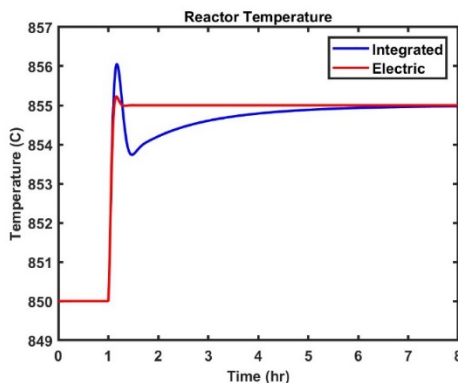
Holdup and temperature controllers were implemented in each unit to stabilize holdups and temperatures. Simple P controllers that manipulate outlet flowrates were implemented for holdup control. For temperature control, PI controllers were implemented to manipulate the heat supply to the units. An additional purity controller is implemented in the liquid phase of the

separator to control the mass fraction of the product ethylene in the product stream by manipulating the temperature setpoint in the intermediate units.

An initial analysis focused on model structure; the two models (for the conventional and electrified processes) were linearized around their nominal steady states. The structure of the system matrices  $A$  in both cases is shown in Figures 6. As shown in Figure 6a, the matrix for the linearized model of the integrated process features additional non-zero entries that correspond to the impact of the states related to the composition of the fuel gas on the heat input rate to the reactor, as discussed above and confirming the findings of the graph-theoretical analysis and the structure of the graph in Figure 4.

Subsequent simulations focused on process dynamics. As an example, a  $5^{\circ}\text{C}$  increase in the reactor temperature setpoint, imposed at  $t = 1$  h is shown and discussed.

Figure 7 shows the behavior of reactor temperature in the two systems. In both systems, the controller settings and the tuning parameters are identical, and the difference in the behaviors would arise solely due to the process structure. The reactor corresponds to Unit 1 in the theoretical analysis and is therefore a unit in the high energy-throughput path. Two-time scale behavior is expected in the integrated model with high energy recycle throughput via combustion of by-product fuel gas. In the electrified case, the setpoint is tracked accurately and quickly, corresponding to the fast dynamics of the temperature variable. On the other hand, in the conventional case, the dynamic behavior is more complex, with the fast dynamics being followed by a slower evolution of the temperature towards the set-point, which, as explained earlier, is due to the (slow) feedback effect linking the composition changes in the reactor and separator to the heating value of the fuel gas.



**Figure 7.** Reactor temperature profile subject to reactor temperature setpoint increase of  $5^{\circ}\text{C}$ .

## CONCLUSIONS

In conclusion, the effects of electric heating on process structures and dynamics were studied theoretically and were demonstrated using simulation results. Through singular perturbation analysis, the multiple time-scale behavior was predicted in conventional systems with mass- energy recycling via combustion of waste products, whereas such behavior was expected to disappear in new electrified heating systems. The theoretical results were proven via simulations depicting a conventional ethylene cracker model with tail gas recycle and an electric cracker model.

## DIGITAL SUPPLEMENTARY MATERIAL

The Matlab implementation of the model is freely available on GitHub at <https://github.com/Baldea-Group/EthylenePlant/>.

## ACKNOWLEDGEMENTS

Financial support from Shell International Exploration and Production, Houston, TX is acknowledged with gratitude.

## NOMENCLATURE

### Acronym

LHV Lower heating value [kJ/kg]

### Uppercase

$A$	Linearized system matrix
$C$	Concentration [ $\text{mol}/\text{m}^3$ ]
$E_a$	Activation Energy [kJ/mol]
$F$	Mass flowrate [kg/s]
$H$	Mass enthalpies [kJ/kg]
$I$	Scaled variable
$M$	Mass holdup [kg]
$N$	Total number of units in the prototype process model
$Q$	Heating / cooling rate [kW]
$T$	Temperature [ $^{\circ}\text{C}$ ]

### Lowercase

$f$	Transformation in mass fraction in each unit
$g$	Transformation in mass enthalpy in each unit
$i$	Unit at which cooling occurs
$u$	Scaled input
$k$	Scaled energy flow
$k_0$	Rate constant
$x$	Mass fraction

### Greek Letter

$\varepsilon$	Singular perturbation parameter
$\lambda$	Heating value [kJ/kg]
$\tau$	Stretched time variable for the fast dynamics [s]

## Subscript

<i>comb</i>	Combustion
<i>elec</i>	Electric heating
<i>f</i>	Fresh fuel stream
<i>j</i>	Units, streams [1,2,... <i>i</i> ,... <i>N</i> ]
<i>m</i>	Reactions [1,2,-1]
<i>q</i>	Heating / cooling flow
<i>R</i>	Recycle fuel gas flow

## Superscript

$\square$	Steady state
-----------	--------------

## REFERENCES

1. McMillan C, Schoeneberger C, Zhang J, Kurup P, Masanet E, Margolis R, Meyers S, Bannister M, Rosenlieb E, and Xi W. Opportunities for solar industrial process heat in the United States. Technical Report NREL/TP-6A20- 77760, National Renewable Energy Laboratory, Golden, CO, (2021)
2. Schoeneberger C, McMillan CA, Kurup P, Akar S, Margolis R, and Masanet E. Solar for industrial process heat: A review of technologies, analysis approaches, and potential applications in the United States. *Energy*, 206:118083, (2020)
3. Schiffer ZJ and Manthiram K. Electrification and decarbonization of the chemical industry. *Joule*, 1(1):10–14, (2017)
4. Electric Power Research Institute (EPRI). Program on technology innovation: Industrial electrotechnology development opportunities. Technical Report 1019416, EPRI, Palo Alto, CA, (2009)
5. Layritz LS, Dolganova I, Finkbeiner M, Luderer G, Penteado AT, Ueckerdt F, and Repke JU. The potential of direct steam cracker electrification and carbon capture & utilization via oxidative coupling of methane as decarbonization strategies for ethylene production. *Applied Energy*, 296:117049, (2021)
6. Flexibility Resources Task Force. Increasing electric power system flexibility: The role of industrial electrification and green hydrogen production. (2022)
7. Baldea M and Daoutidis P. Dynamics and Control of Integrated Process Systems. Cambridge University Press (2012)
8. Baldea M and Daoutidis P. Model reduction and control of reactor–heat exchanger networks. *Journal of Process Control*, 16(3):265–274, (2006)
9. Amghizar I, Dedeyne JN, Brown DJ, Marin GB, and Van Geem KM. Sustainable innovations in steam cracking: CO<sub>2</sub> neutral olefin production. *React. Chem. Eng.*, 5:239–257 (2020)
10. Tijani MEH, Zondag H, and Van Delft Y. Review of

electric cracking of hydrocarbons. *ACS Sustainable Chemistry & Engineering*, 10(49):16070–16089, (2022)

11. Dow Corporate. Dow and Shell to develop electric cracking. URL <https://corporate.dow.com/en-us/news/press-releases/dow-shell-electric-cracking-technology.html>. Accessed: 10/25/2023.
12. Ren T et al. Olefins from conventional and heavy feedstocks: Energy use in steam cracking and alternative processes. *Energy*, 31(4):425–451, March (2006)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Biogas Valorization from a Process Synthesis Perspective: Heat and Work Integration to Maximize CO<sub>2</sub> Conversion

Baraka C. Sempuga\*, Selusiwe Ncube

University of South Africa, Institute for Catalysis and Energy Solutions, Johannesburg, Gauteng, South Africa.

\* Corresponding Author: sempubc@unisa.ac.za

## ABSTRACT

Biogas is often considered as a source of renewable energy, for heat and power production. However, biogas has greater promise as a source of concentrated CO<sub>2</sub> in addition to methane, making it a rich supply of carbon and hydrogen for the generation of fuel and chemicals. In this work, we use the concept of attainable region in the enthalpy-Gibbs free energy space to identify opportunities for effective biogas valorization that maximizes the conversion of CO<sub>2</sub>. The AR concept allows us to study a chemical process without knowing the exact reaction mechanism that the species in the process use. Deriving Material Balance equations that relate a reactive process's output species to its input species is sufficient to identify process limits and explore opportunities to optimize its performance in terms of material, energy, and work. The conversion of biogas to valuable products is currently done in two steps; the high temperature and endothermic reformer step, followed by the low temperature exothermic synthesis step. We demonstrate, using Aspen Simulation, that energy integration, both heat and work, between the two steps is crucial to achieving a substantial amount of CO<sub>2</sub> conversion. We also show how a heat pump configuration can be utilized to integrate energy between the reformer and synthesis steps against the temperature gradient by integrating external renewable energy.

**Keywords:** Carbon Dioxide, Process Synthesis, Energy, Optimization, Methane Reforming, Target Material Balance, Minimizing CO<sub>2</sub> Emissions, Work Analysis, Entropy Analysis

## INTRODUCTION

Biogas is not only a source of green energy [1] but also a source of concentrated green CO<sub>2</sub>. The composition of CO<sub>2</sub> in biogas typically ranges between 25-50% by volume, with the balance primarily consisting of methane, varying depending on the feedstock [2,3]. This characteristic makes biogas a suitable feed for CO<sub>2</sub> conversion processes. Several authors have shown different ways of upgrading and valorizing biogas by either extracting the energy by direct combustion or converting it to chemicals such as methanol [1,4-6]. Several authors have also demonstrated different techniques to converting the CO<sub>2</sub> in biogas along with methane to chemicals, including dry reforming where CO<sub>2</sub> and CH<sub>4</sub> are converted to synthesis gas followed by methanol synthesis [4]. Another technique starts by removing the CO<sub>2</sub> from CH<sub>4</sub> before converting to hydrogen via steam reforming and water gas shift reaction followed by CO<sub>2</sub> hydrogenation to methanol

[6]. While these studies have shown these conversion techniques to be feasible, one major drawback is the intense energy demand of the reforming step which could result in substantially lower net CO<sub>2</sub> conversion depending on where the energy is sourced from. For instance, if the energy is supplied by burning a carbon-based fuel, which is commonly practiced, then one needs to consider the CO<sub>2</sub> emitted during the combustion of the fuel. Ideally, the energy for reforming should come from a renewable source in the form of electricity to achieve a net negative CO<sub>2</sub> emission. Another drawback is the inability to integrate the energy between the synthesis and reformer sections due to temperature gradients. Ncube et. al. (2023) [7] have shown that heat integration between the reformer and synthesis section is essential to achieving higher energy efficiency and increased CO<sub>2</sub> conversion and they suggested that this can be done by integrating a heat pump that runs on renewable energy to the system. In this work, we use the attainable region concept



to identify more effective ways for converting biogas from a material, energy, and work balance perspective. To optimize the conversion of CO<sub>2</sub>, we place a strong emphasis on the product selection and energy integration.

## THE APPROACH

In this section we would like to determine the attainable region for biogas conversion into various products. The attainable region will provide information on the performance limits of the process from the material and energy perspectives. The limits of performance can be understood as the extreme points that any biogas conversion process can achieve for a given feed composition. Thus, for the given feed composition and a given set of species involved in the process, we can determine the attainable region of the process by finding all possible composition states that can result from the chemical conversion of the biogas into the specified species. We will explore how this is done by using the notation and terminology found in Smith & Missen (1982) [8].

For a *chemical system* containing  $N$  species and  $M$  elements, denoted as:  $\{(A_1, A_2, \dots, A_i, \dots, A_N), (E_1, E_2, \dots, E_k, \dots, E_M)\}$  where  $A_i$  is the elemental formula of species  $i$  and  $E_k$  is the  $k$ th element in the system. The order in which the species and elements are presented in the system is not important for the attainable region, however, once decided upon, it must be kept consistent during the derivation as most the equations are written in matrix form. Thus for the biogas conversion chemical system with 9 species ( $N = 9$ ) and 3 elements ( $M = 3$ ) we have

$\{(CH_4, CO_2, H_2O, O_2, H_2, CO, CH_2, CH_3OH, C), (C, H, O)\}$ .

The *element-abundance* constraint is one of the constraints that must always be satisfied as it constitutes the basis for the material conservation within the system.

$$\mathbf{A}\mathbf{n} = \mathbf{b} \quad (1)$$

-  $\mathbf{A}$  is the *formula matrix* an  $M \times N$  matrix containing the amount of each element in each species where  $M$  is the number of elements and  $N$  is the number of species in the system.

-  $\mathbf{n}$  is the *species-abundance* vector ( $N \times 1$ ) containing the number of moles of each species in the system.

-  $\mathbf{b}$  is the *element-abundance* vector containing the total amount of each element available in the system. Thus  $\mathbf{b}$  is fixed for a given feed composition.

Given the initial number of moles of all the species (or the initial species-abundance vector)  $\mathbf{n}^0$ , the element-abundance vector  $\mathbf{b}$  is fixed and is given by  $\mathbf{b} = \mathbf{A}\mathbf{n}^0$ . Therefore, the element-abundance constraint can also be written as follows:

$$\mathbf{A}\mathbf{n} = \mathbf{A}\mathbf{n}^0 \quad (2)$$

The compositional state  $\mathbf{n}$  of a system can be

expressed in terms of an initial compositional state  $\mathbf{n}^0$  as follows:

$$\mathbf{n} = \mathbf{n}^0 + \sum_{j=1}^R \mathbf{v}_j \xi_j \quad (3)$$

where  $\mathbf{v}_j$  is the stoichiometric vector such that:

$$\mathbf{A}\mathbf{v}_j = \mathbf{0}; (\mathbf{v}_j \neq \mathbf{0}); \quad j = 1, 2, \dots, R \quad (4)$$

Thus  $\mathbf{v}_j$  is any set of  $R$  independent solutions that satisfies eq.4. Therefore,  $R$  is the maximum number of linearly independent solutions for eq.4. It is given by:

$$R = N - C \quad (5)$$

$$C = \text{rank}(\mathbf{A}) \quad (6)$$

For the biogas system the formula matrix  $\mathbf{A}$  is derived from the elemental composition of the species as shown in the table below

**Table 1:** Derivation of the formula matrix from the elemental composition of the species

$N \setminus M$	$C$	$H$	$O$
$CH_4$	1	4	0
$CO_2$	1	0	2
$H_2O$	0	2	1
$O_2$	0	0	2
$H_2$	0	2	0
$CO$	1	0	1
$C_8H_{18}$	8	18	0
$CH_3OH$	1	4	1
$C$	1	0	0

Therefore  $\mathbf{A}(M \times N)$  is:

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 1 & 8 & 1 & 1 \\ 4 & 0 & 2 & 0 & 2 & 0 & 18 & 4 & 0 \\ 0 & 2 & 1 & 2 & 0 & 1 & 0 & 1 & 0 \end{pmatrix} \quad (7)$$

It follows that  $C = \text{rank}(\mathbf{A}) = 3$ ; therefore  $R = N - C = 9 - 3 = 6$ , where  $R$  is the maximum number of independent chemical equations that can be derived from the system.

We therefore define the matrix  $\mathbf{N}$ , an  $N \times R$  matrix, as the *complete stoichiometric matrix* whose  $R$  columns are linearly independent vectors  $\mathbf{v}_j$ . And therefore eq. 4 can be written as a single matrix as follows:

$$\mathbf{A}\mathbf{N} = \mathbf{0} \quad (8)$$

where  $\mathbf{N} = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_R)$ .

The vector  $\mathbf{v}_j$  contains the stoichiometric coefficients of the  $j$ th chemical equation. Therefore, for any chemical system there are only  $R$  independent chemical equations that represents all possible compositional states of the system.

For the biogas conversion process, we adopt the method described in Smith & Missen (1982) [8] and use Matlab® to derive the stoichiometric matrix  $\mathbf{N}(N \times R)$ . In this case we derive a set of the 6 non-unique independent chemical equations for this system, where negative

values stand for the coefficient on the left side (reactant) and positive are those on the right side (product) of the equation.

$$\mathbf{N} = \begin{pmatrix} 0.50 & -0.25 & -0.25 & -6.25 & -0.75 & -0.5 \\ -0.50 & 0.25 & -0.75 & -1.75 & -0.25 & -0.5 \\ -1.00 & -0.50 & 0.50 & 3.50 & -0.5 & 1 \\ 1.00 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.00 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.00 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.00 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1.00 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

From eq.4, a complete set of  $\mathbf{v}_j$  (also written as  $\mathbf{N}$ ) is not unique but can be any set that satisfies eq.4 (or eq.8). Therefore, the set of independent chemical equations derived from the stoichiometric vectors do not represent any *chemical reactions* occurring in the system but are simply algebraic solutions of all possible composition states that any chemical reactions path can achieve. Thus, to distinguish a *chemical equation* from a *chemical reaction*, we use the equal sign (=) instead of an arrow ( $\rightarrow$ ).

The quantities  $\xi_j$  in eq.3 are a set of parameters, which determine the linear combination of the coefficient vectors  $\mathbf{v}_j$  required to achieve a particular compositional state  $\mathbf{n}$ . Thus, the number of moles of each species at a compositional state is given by:

$$n_i = n_i^o + \sum_{j=1}^R v_{ij} \xi_j \quad (9)$$

where  $v_{ij}$  is the *stoichiometric coefficient* of  $i$ th species in the  $j$ th stoichiometric vector, and  $\xi_j$  is the parameter for the  $j$ th stoichiometric vector.

The *Attainable Region* (AR) for the chemical system is defined by the sets of  $\xi_j$ 's for which the number of moles of each species is positive. Thus, the AR is obtained by using eq 9 as an additional constraint when written as an inequality equation as follows:

$$n_i = n_i^o + \sum_{j=1}^R v_{ij} \xi_j \geq 0 \quad (10)$$

One can also formulate a linear programming problem to determine the attainable region as follows:

$$\text{find all } \xi_j, \text{ subject to } \begin{cases} \mathbf{n} = \mathbf{n}^o + \sum_{j=1}^R \mathbf{v}_j \xi_j \geq 0 \\ \mathbf{A}\mathbf{n} = \mathbf{A}\mathbf{n}^o \end{cases} \quad (11)$$

Eq.11 is a linear programming problem whose vertices constitute a convex connected region in the  $\xi$  space of dimension  $R$ . It can also be transformed into the  $n_i$  space of dimension  $N$ .

If there are  $K$  extreme points in the attainable region, we can denote the set of these vertices in the  $\xi$  space as  $\xi_V$  ( $\xi_{V_1}, \xi_{V_2}, \dots, \xi_{V_K}$ ), where  $\xi_{V_k}$  is the parameter vector at the  $k$ th vertex. Similarly, in the  $n_i$  space as we denote  $\mathbf{n}_V$  ( $\mathbf{n}_{V_1}, \mathbf{n}_{V_2}, \dots, \mathbf{n}_{V_K}$ ), where  $\mathbf{n}_{V_k}$  is the composition vector at the  $k$ th vertex of the AR.

Another way of visualizing the AR is to transform its vertices into a two-dimensional space of Gibbs Free Energy and enthalpy as follows:

$$\left. \begin{aligned} \Delta G_{V_k} &= \sum_{j=1}^R \xi_{jV_k} \mathbf{v}_j \hat{\mathbf{G}}^T \\ \Delta H_{V_k} &= \sum_{j=1}^R \xi_{jV_k} \mathbf{v}_j \hat{\mathbf{H}}^T \end{aligned} \right\} \quad (12)$$

Where  $\hat{\mathbf{G}}$  and  $\hat{\mathbf{H}}$  are the molar Gibbs free energy and molar enthalpy vector for the species.

Figure 1 is the attainable region of biogas conversion in the Gibbs Free Energy and enthalpy space (GH-AR) with a feed composition of  $\text{CH}_4$ ,  $\text{CO}_2$  and  $\text{H}_2\text{O}$  at a ratio of 4:4:1. It shows the theoretical limits of all possible product compositions that can be obtained from the given feed composition, while simultaneously showing the minimum energy requirement ( $\Delta H$ ) and minimum work (or energy quality) requirement ( $\Delta G$ ), both taken at reference temperature  $T_o = 25^\circ$  and reference pressure  $P_o = 1\text{bar}$ , associated with the conversion of the given feed to products. Each point of the GH-AR can be considered as an *overall process* target for energy, exergy (here represented by  $\Delta G$  since the change in exergy at  $T_o$  and  $P_o$  is equal to  $\Delta G$  at  $T_o$  and  $P_o$ ), and material conversion, and as such can serve as a benchmark for real processes. A more elaborate discussion on the GH-AR can be found in Fernandez-Torres et al. (2023) [9], however, we will discuss a few important insights into the process. Using the GH-AR in Figure 1. The vertices in the GH-AR represent the points where one or more of the species is zero thus represent the limits for biogas conversion. Thus, points of complete and incomplete overall conversion of  $\text{CH}_4$  and  $\text{CO}_2$  can be seen on GH-AR where either  $s_1$  or  $s_2$ , representing  $\text{CH}_4$  and  $\text{CO}_2$  respectively, or both appear. Some of the points contain undesirable products such as carbon C ( $s_9$ ), CO ( $s_6$ ) and  $\text{H}_2$  ( $s_5$ ) as final products.

Another important insight from the GH-AR is regarding biogas conversion to methanol. The observation in Figure 1 is that every point in the GH-AR representing methanol ( $s_8$ ) as a product is accompanied by other less desirable products. An example of such a point is point M ( $s_2, s_8, s_9$ ) representing a process target that produces methanol and the less desirable products  $\text{CO}_2$  and C. Since  $\text{CO}_2$  is one of the feed components, the production of  $\text{CO}_2$  at M can be thought of as incomplete overall conversion of  $\text{CO}_2$ . Another example is at point N ( $s_4, s_6, s_8$ ) representing  $\text{O}_2$ , CO and methanol as final product. While  $\text{O}_2$  is not necessarily an undesirable product but CO is less desirable as a final product. This observation suggests that undesirable by-products may be inevitable when converting biogas to methanol. Thus, the process seems to be less optimal although it receives increasing interest among several authors [4-6].

On the other hand, Fisher Tropsch fuels (FT-fuels) production can be more effective in the conversion of  $\text{CO}_2$  without producing undesired by-products. This can be seen at point F ( $s_4, s_8$ ) in Figure 1 representing a

process target to produce octane (s8) and oxygen (s4) only as final products. Furthermore, point F represents a 100% carbon and hydrogen conversion efficiency where all the carbon and hydrogen in the feed report to the desired product, octane. As such, this point must be considered the ultimate target for biogas conversion for the given feed composition. We will explore with the aid of an Aspen Plus simulation how the target at F can be achieved by integrating both the material and energy within the process.

The GH-AR not only provides information regarding the material conversion limit but also the energy ( $\Delta H$ ) and the work ( $\Delta G$ ) requirements to achieve a specific product composition [8]. One important observation is that for biogas conversion almost all targets are in the Region A and B with positive  $\Delta H$  and  $\Delta G$ , which shows that external energy supply is necessary to achieve the targets where  $\text{CO}_2$  is consumed. Points located at the bottom section of the GH-AR have less  $\text{CO}_2$  conversion thus need less energy compared to those in the upper part where more  $\text{CO}_2$  is consumed and therefore need more energy at higher quality as  $\Delta G$  is higher.

Furthermore, the  $45^\circ$  line represents processes that require the highest energy quality (pure work, or electricity). The processes in Region A, below the  $45^\circ$  line, are relatively easy to achieve in one step by utilizing a source of thermal energy at an appropriate temperature. However, those in region B, above the  $45^\circ$  line, are not feasible if done in one step from feed to final product but require several intermediate steps that are fully integrated in terms of energy and materials and would require a supply of high-quality energy. This is the case to produce FT-fuels, with its ultimate target located in region B. Thus, to achieve this target, one must run the process in at least 2 steps, notably, the reforming step where biogas is converted to synthesis gas followed by the FT synthesis section where syngas is converted to liquid fuels.

One of the major challenges in achieving this target is the inability to integrate energy between the reforming section, an endothermic process that runs at higher temperatures, and the synthesis section, an exothermic section that runs at lower temperature. Ncube et. al. (2023) [7] have also shown that a heat pump is needed to integrate the energy between these two sections. In this work we explore how a heat engine configuration can be used to achieve both the material and energy integration within the system, by means of water electrolysis as shown in the Aspen Plus simulation in Figure 2.

The feed to the process flowsheet denoted as Biogas enters the process at a ratio of 4:4:1 of  $\text{CH}_4$ :  $\text{CO}_2$ :  $\text{H}_2\text{O}$ . The feed is preheated before entering the reformer where the feed is converted to syngas, a mixture of mainly  $\text{CO}$  and  $\text{H}_2$ . The reformer is modelled using a Gibbs reactor block (RGibbs), which uses Gibbs free energy minimization to determine the equilibrium composition of specified

components in the reactor at given conditions of temperature and pressure. The components that were considered are  $\text{CH}_4$ ,  $\text{CO}_2$ ,  $\text{H}_2\text{O}$ ,  $\text{O}_2$ ,  $\text{H}_2$ ,  $\text{CO}$ , and  $\text{C}$  (as pure solid). A temperature of  $950^\circ\text{C}$  in the reformer was used to achieve a high conversion of  $\text{CH}_4$  and  $\text{CO}_2$  at 1 bar. The syngas from the reformer is cooled down using the incoming stream to the reformer, and then goes to the Fischer Tropsch (FT) synthesis reactor (FT-RCT) where  $\text{CO}$  and  $\text{H}_2$  are converted FT-fuels (represented by  $\text{C}_8\text{H}_{18}$ ) and  $\text{H}_2\text{O}$ . The FT-synthesis reactor is modelled using a stoichiometric reactor block (RStoic) which is used when the reactions and conversions are known. In this simulation we consider only one reaction in the FT-RCT to model the production of FT-fuels and we assume complete conversion of  $\text{CO}$  to  $\text{C}_8\text{H}_{18}$  via the reaction;  $8\text{CO} + 17\text{H}_2 \rightarrow 8\text{H}_2\text{O} + \text{C}_8\text{H}_{18}$  at  $250^\circ\text{C}$  and 1 bar. The product stream from the FT-RCT is cooled down and separated using a Sep block (PR-SEP1) which assumes perfect separation based on split fractions or mole flows of pure components. The PR-SEP1 produces a pure stream of FT-fuel, water and a gas stream containing unreacted  $\text{H}_2$ ,  $\text{CO}$ ,  $\text{CO}_2$  and  $\text{CH}_4$ . The water is sent to the electrolyzer reactor (ELECT-RC) modelled using the RStoic block to produce  $\text{H}_2$  and  $\text{O}_2$  at  $250^\circ\text{C}$ . The temperature of the electrolyzer was chosen to allow heat integration with the FT-RCT. The product from the electrolyzer is separated using the Sep block to achieve a perfect separation of  $\text{H}_2$  and  $\text{O}_2$ . Part of the  $\text{H}_2$  produced is used to supply the energy needed at the reformer using a portion of the  $\text{O}_2$  produced from the electrolyser, and the remaining  $\text{H}_2$  is fed to the reformer to enhance the conversion of  $\text{CO}_2$ . On the hand the remaining  $\text{O}_2$  from the electrolyser is taken out of the process as a product. The combustion of  $\text{H}_2$  with  $\text{O}_2$  is modelled using the RStoic block with the combustion temperature set at  $955^\circ\text{C}$  to allow for at least a temperature approach of  $5^\circ\text{C}$  for heat exchanged between the combustion unit at the reformer.

## SIMULATION RESULTS

The electrolyzer and  $\text{H}_2$  combustion configuration enables achieving three main goals:

- 1) Heat transfer from the synthesis reactor at  $250^\circ\text{C}$  to the reformer at  $950^\circ\text{C}$  is achieved by using the energy released from the FT-synthesis to provide heat to the electrolyzer and thus lowering the electrical requirements of the electrolyzer. This energy is chemically stored in  $\text{H}_2$  part of which is converted to high temperature heat required by the reformer.



2) The recycled hydrogen to the reformer provides a hydrogen rich environment that enhances the conversion of CO<sub>2</sub> via the reverse water gas shift reaction to produce CO and H<sub>2</sub>O, carbon formation is also eliminated at temperatures above 750°C. The reformer temperature is set at 950°C leading to a CO<sub>2</sub> conversion across the reactor of approximately 94% and a CH<sub>4</sub> conversion of 99% (See Table 2). Similar results were also obtained by Rosha et.al. (2021) with slight deviations due to the differences in CH<sub>4</sub> and CO<sub>2</sub> ratios in the feed and in the operating pressure. The ratio H<sub>2</sub>/CO obtained in approximately 2.16.

3) The overall hydrogen and carbon efficiency ( $\eta_{H_2}$  and  $\eta_C$ ) in Eq. 13,

$$\eta_i = \frac{\text{moles of atoms of } i \text{ coming into the process}}{\text{moles of atoms of } i \text{ in the desired product}} \times 100 \quad (13)$$

is the fraction of the total hydrogen and carbon atoms in the feed that reports to the desired product. Assuming a high CO conversion in the FT-reactor, a 97% hydrogen efficiency and a 99% carbon efficiency were obtained. This results in almost no tail gas, leading to FT-fuels and O<sub>2</sub> as the main products across the process.

4) The energy efficiency is evaluated in terms thermal efficiency expressed as

$$\eta_{thermal} = \frac{m_{FT-fuels} HHV_{FT-fuels}}{m_{Biogas} HHV_{Biogas} + W_{electricity}} \quad (14)$$

where *HHV* is the higher heating value, *m* is the molar flowrate and *W<sub>electricity</sub>* is the electrical power input to the process. The electrical power supplied to the electrolyzer is the only point of entry of high-quality energy across the process. This meets the target set out in the attainable region as needed by the overall process as discussed in the previous section. The process requires 135kW of electrical energy. This translates to a thermal efficiency of 96%, a significant improvement compared to other reported results such as those by Guares et al. (2021) [10] who reported 41-46% thermal efficiency in a biogas to power process. Rosha et al. (2021) [4] reported 63.7% thermal efficiency in a biogas to methanol process.

**Table 2:** Material flows in selected major streams of the Aspen Plus simulation in Figure 2.

Kmol/h	Biogas	Syn-gas-1	FT-Fuel	Oxy-gen	Tail Gas
CH <sub>4</sub>	1	0.005	0	0	0.005
CO <sub>2</sub>	1	0.062	0	0	0.062
H <sub>2</sub> O	0.25	0.193	0	0	0
O <sub>2</sub>	0	0	0	1.06	0
H <sub>2</sub>	0	4.174	0	0	0.065
CO	0	1.933	0	0	0
C <sub>8</sub> H <sub>18</sub>	0	0	0.242	0	0
CH <sub>3</sub> OH	0	0	0	0	0
C	0		0	0	0

## CONCLUSION

The efficiency of biogas conversion to valuable products depends on the choice of the products and the efficacy of the material and energy integration within the process. The GH attainable region is a useful tool in identifying the process material balance targets with high CO<sub>2</sub> conversion and with less or no undesirable by products. Conversion of biogas to FT-fuel is one such a process with an ultimate target of 100% CO<sub>2</sub> and CH<sub>4</sub> conversion to produce FT-fuels and oxygen. This, however, requires the supply of high-quality energy, such as electricity or high temperature heat, from an external source, coupled with adequate material and energy integration within the process. We have shown, using an Aspen plus simulation, that electrical power supply to the overall process via water electrolysis not only enables supplying high quality energy to the process but also facilitates the integration of energy from the synthesis section to the reformer section against the temperature gradient hence acting as heat pump. The Aspen Plus simulation shows that with this process configuration 97% carbon and 99% hydrogen conversion is possible.

The present work is based on a specific feed composition. Changing this parameter will change the results observed in the attainable region in Figure 1 which could lead to a different ultimate process target and would require a different process configuration to achieve the target. For instance, a higher ratio of CH<sub>4</sub> to CO<sub>2</sub> in the feed could lead to the production of FT-fuels and Carbon as the ultimate process target rather than the production of FT-fuels and oxygen. To avoid this one may require modifying the amount of water to be fed to the process to ensure that a target for the production of FT-fuels and O<sub>2</sub> is maintained. This suggests that the ratios of carbon, hydrogen and oxygen in the feed is a significant factor that needs to be investigated.

## REFERENCES

- Caetano B C, Santos N D, Hanriot V. M, Sandoval O R, Huebner R, Energy Conversion of Biogas from Livestock Manure to Electricity Energy Using a Stirling Engine. *Energy Conversion and Management*. 15, 100224 (2022).
- Da Costa Gomez, C.1- Biogas as an Energy Option: An Overview. In *The Biogas Handbook*; Welling, A., Murphy, J., Baxter, D., Eds., Woodhead Publishing, 1-16 (2013).
- Aghel B., Behaein S., Wongwises S., Shadloo M. S. A, Review of Recent Progress in Biogas Upgrading: With Emphasis on Carbon Capture. *Biomass and Bioenergy*. Elsevier Ltd May 1, (2022).
- Rosha P., Kumar S., Ibrahim H, A Thermodynamic Analysis of Biogas-to-Methanol Conversion with



- CH<sub>4</sub> recycling and CO<sub>2</sub> utilization using Aspen HYSYS. *Sustain Energy Fuels* 5 (17), 4336–4345 (2021).
5. Moiola E., Schildhauer T, Eco-Techno-Economic Analysis of Methanol Production from Biogas and Power-to-X. *Ind Eng Chem Res*, 61 (21), 7335–7348 (2022).
  6. Ghosh S., Uday V., Giri A., Srinivas S, Biogas to Methanol: A Comparison of Conversion Processes Involving Direct Carbon Dioxide Hydrogenation and via Reverse Water Gas Shift Reaction. *J Clean Prod* 217, 615–626 (2019).
  7. Ncube S., Sempuga B. C., Liu X., Biogas Valorization to Liquid Fuels: Modelling and Setting up Process Targets. In *33rd European Symposium on Computer Aided Process Engineering*, Kokossis, Computer Aided Chemical Engineering; Elsevier, 52, 2827–2832 (2023).
  8. Smith W. R., Missen R. W, *Chemical Reaction Equilibrium Analysis: Theory and Algorithms*; John Wiley & Sons: New York, (1982).
  9. Fernandez-Torres M, Hildebrandt D, Glasser D, Sempuga B, Thermodynamic Constraints on the Catalytic Reduction of Nitrates in Drinking Water. *Ind Eng Chem Res* 62 (12), 5305–5314 (2023).
  10. Guares, S. A.; Lima, J. D. de; Oliveira, G. A. Techno-Economic Model to Appraise the Use of Cattle Manure in Biodigesters in the Generation of Electrical Energy and Biofertilizer. *Biomass Bioenergy* **2021**, 150

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Towards Designing Sector-Coupled Energy Systems Within Planetary Boundaries

David Y. Shu<sup>a,†</sup>, Jan Hartmann<sup>b,†</sup>, Christian Zibunas<sup>b</sup>, Nils Baumgärtner<sup>c</sup>, Niklas von der Assen<sup>b</sup>, André Bardow<sup>a,\*</sup>

<sup>a</sup> ETH Zurich, Energy and Process Systems Engineering, Zurich, Switzerland

<sup>b</sup> RWTH Aachen University, Institute of Technical Thermodynamics, Aachen, Germany

<sup>c</sup> Currenta GmbH & Co. OHG, Energy Trading and Dispatching, Leverkusen, Germany

\* Corresponding Author: [abardow@ethz.ch](mailto:abardow@ethz.ch).

† These authors contributed equally.

---

## ABSTRACT

The transition to net-zero greenhouse gas emissions requires a rapid redesign of energy systems. However, the redesign may shift environmental impacts to other categories than climate change. To assess the sustainability of the resulting impacts, the planetary boundaries framework provides absolute limits for environmental sustainability. This study uses the planetary boundaries framework to assess net-zero sector-coupled energy system designs for absolute environmental sustainability. Considering Germany as a case study, we extend the common focus on climate change in sustainable energy system design to seven additional Earth-system processes crucial for maintaining conditions favorable to human well-being. Our assessment reveals that transitioning to net-zero greenhouse gas emissions reduces many environmental impacts but is not equivalent to sustainability, as all net-zero designs transgress at least one planetary boundary. However, the environmental impacts vary substantially between net-zero designs, highlighting that design choices exist to address transgressions of planetary boundaries.

---

**Keywords:** Energy Systems, Life Cycle Assessment, Modelling, Optimization, Carbon Capture, Environment, Sector-coupling

## INTRODUCTION

The energy system needs to be redesigned to reduce greenhouse gas emissions to net-zero. The redesign is commonly guided by energy systems models in optimization studies [1]. A common finding in energy system optimization studies is the rise of sector coupling to integrate low-carbon electricity into sectors such as mobility and heating [2–5]. Combined with environmental life-cycle assessment (LCA) [6], energy system modeling and optimization can account for climate change and additional environmental impacts.

In LCAs of the energy system transition, reducing the climate change impact of energy systems has been shown to result in burden-shifting, i.e., environmental impacts shift from climate change to other categories, such as in land use, resource depletion, toxicity, and ecosystem diversity [7, 2]. However, traditional LCA commonly adopts a comparative approach, where the

environmental impacts of systems are assessed in relation to a reference system [8]. While such traditional LCAs show relative differences in impacts, they do not quantify the severity of shifting environmental impacts.

Recently, metrics were introduced to assess environmental trade-offs and provide critical limits by so-called absolute environmental sustainability assessment [8]. Such methods connect life-cycle assessment with absolute environmental sustainability assessment (for a review, see [8]). A popular example is the application of the planetary boundary framework [9–11] to life-cycle assessment [12, 13]. The planetary boundaries framework defines safe operating spaces for climate change and 8 additional Earth-system processes critical to maintaining an Earth-system state that is beneficial for humans. A recent assessment [11] finds that the planetary boundaries are transgressed for 6 Earth-system processes: climate change, change in biosphere integrity, biogeochemical flows of phosphate and nitrogen, land-system change,

freshwater change, and novel entities. Only 3 processes remain within boundaries: stratospheric ozone depletion, ocean acidification, and atmospheric aerosol loading. Hence, it is crucial to consider planetary boundaries in the sustainable design of future energy systems.

In pioneering work, absolute environmental sustainability assessments going beyond climate change have been applied to energy systems [14] but typically consider only one sector, such as power [15–17] or building heat systems [18]. Hence, the impact of sector coupling on absolute sustainability is poorly understood. Here, we conduct an absolute environmental sustainability assessment for net-zero sector-coupled energy systems via the planetary boundaries framework.

## PLANETARY BOUNDARIES IN ENERGY SYSTEM MODELING

Energy system modeling and optimization frameworks vary in model complexity depending on the application. For (inter)national sector-coupled energy system models, linear or mixed-integer linear programming formulations are commonly selected to represent techno-economic constraints [19].

In addition to techno-economic constraints, environmental impacts are included via LCA in models of sector-coupled energy systems from international [3, 4] to national scale [5, 2]. Previous LCA studies [2–5] found environmental burden-shifting resulting from the energy transition, e.g., increasing the use of land, water, and resources.

In previous work, we considered the environmental impacts of the German sector-coupled energy transition to net-zero operational greenhouse gas emissions [20]. Our work revealed increases in up to 7 of 16 impact categories compared to the status quo, e.g., resource depletion of minerals and metals may increase up to four times. However, the degree of burden-shifting can be reduced by design choices: Carbon capture and storage is found to be a lever to steer environmental impacts.

While the identified burden-shifting highlights potential areas of concern, the relative increase in impacts does not reveal if the increase contributes to a transgression of limits of absolute sustainability. Absolute assessments via planetary boundaries aim to overcome this limitation and have been conducted for a single sector of the energy system [14–17]. As these studies are limited to a sector or geographical region and the safe operating space applies to all human activities, downscaling is required, where a share of the safe operating space is allocated to the assessed system [21]. Typically, the safe operating space is allocated by one or a combination of the following principles: egalitarian, utilitarian, or acquired

rights. All of the energy-related studies [15–18] apply downscaling by population to account for the geographical scope, and some further apply downscaling by economic principles or acquired rights.

## METHODS

### Absolute environmental sustainability assessment via planetary boundaries

Here, we assess the absolute sustainability of the net-zero designs for the German sector-coupled energy system identified in previous work [20]. We apply the planetary boundaries framework using the impact assessment method provided in [13]. Note that novel entities are excluded from our assessment, as they are not quantified in [13]. However, in [11], the planetary boundary for novel entities is considered transgressed if any synthetic chemical is released into the environment without adequate safety testing. Hence, the boundary is likely transgressed for the energy system.

As the geographical and sectorial scope is limited in the case study, we allocate a share of the global safe operating space to the German energy system. We first apply downscaling by population based on egalitarian principles to account for the geographical scope. Thus, a share of the total safe operating space is allocated to Germany based on its share of the global population in 2021 [22]. We subsequently apply downscaling by grandfathering to account for the sectorial scope. For the grandfathering, we limit the share of the German safe operating space that the energy system can occupy to the share of environmental impacts caused by the energy system in the reference year (2016). The share of environmental impacts in the reference year is estimated compared to the total German impacts determined using a global input-output database [23]. As an additional reference, we include the share of safe operating space based on gross domestic product in 2021 [24] instead of population. We thus identify environmental impacts exceeding absolute limits for sustainability.

### Energy system model description

The net-zero sector-coupled energy systems are designed via a modeling and optimization framework with integrated LCA<sup>1</sup> [25] based on the life-cycle inventory database ecoinvent 3.5 (APOS) [26].

The system boundary of the energy system includes the electricity sector, the private mobility sector, and the heating sector for buildings and for industry on three temperature levels. In addition, we include CCS technologies and a direct air capture technology to enable CO<sub>2</sub> emission avoidance and CO<sub>2</sub> removal (Table 1).

As the functional unit, we select the supply of all

<sup>1</sup> [git-ce.rwth-aachen.de/secmod](https://git-ce.rwth-aachen.de/secmod)

exogenous end-use demands for electricity, mobility, and heat. Additionally, we constrain operational greenhouse gas emissions to reach net-negative emissions of -29 Mt CO<sub>2</sub>-eq. in 2045, assuming that the energy system contributes to balancing hard-to-abate emissions, e.g., in agriculture. A detailed description is available in [20].

**Table 1:** Technologies considered in the energy system case study based on [2, 20].

<b>electricity</b>	<b>heating</b>
biogas-to-power	<b>building</b>
geothermal	natural gas boiler (district)
hard coal	natural gas boiler
hydrogen fuel cell	electrode boiler
lignite	energetic rehabilitation
natural gas combined cycle (NGCC)	heat pump
natural gas turbine	oil boiler
nuclear	<b>industry, low-temp.</b>
oil	natural gas boiler (district)
photovoltaics	natural gas boiler
run-of-river	electrode boiler
waste-incineration	heat pump
wind, offshore	<b>industry, medium-temp.</b>
wind, onshore	natural gas boiler (district)
other non-renewables	natural gas boiler
lithium-ion battery	electrode boiler
pumped hydro storage	<b>industry, high-temp.</b>
	natural gas boiler (district)
	natural gas boiler
<b>private mobility</b>	<b>carbon capture &amp; storage</b>
battery electric	direct air capture
compressed natural gas	cement industry
diesel	CO <sub>2</sub> pipeline
gasoline	NGCC
hydrogen fuel cell	steel industry
plug-in hybrid	geological storage CO <sub>2</sub>
<b>power-to-X</b>	<b>transmission</b>
power-to-diesel	220 kV power line
power-to-hydrogen	380 kV power line
power-to-methane	upgrade 220 to 380 kV

The multi-period investment decisions are determined in a rolling-horizon optimization, minimizing total annualized cost with investments every 5 years for a foresight horizon of 10 years. We apply a linear programming formulation for the design problem, assuming linear input-output relationships in energy conversion, continuous equipment sizing, and linear investment and operating costs.

## CASE STUDY RESULTS

Transition pathways depend on key technology options, such as the availability of green electricity imports or DAC as a carbon dioxide removal technology. DAC opens a design space with solutions spanning between the cost-optimal and the minimally-required deployment of CO<sub>2</sub> sequestration that still meets greenhouse gas emission targets.

Here, we assess the absolute environmental sustainability assessment of a conservative scenario from [20] that excludes electricity imports into the energy system. Further, we consider 3 sub-scenarios where CO<sub>2</sub> storage is 1) unconstrained (min-TAC), 2) constrained to a minimum (min-storage), and 3) constrained to an intermediate value (compromise).

The net-zero energy systems outperform the fossil system from the initial year of the transition horizon in at least 6 of 9 impact categories (Figure 1). Only for the nitrogen cycle, impacts increase beyond the level of 2016 in all net-zero designs. Additional burden-shifting occurs only in scenarios with minimal CO<sub>2</sub> sequestration, where the impacts increase in atmospheric aerosol loading and in freshwater use compared to the reference in 2016. However, the absolute sustainability assessment reveals that the burden-shifting for freshwater use does not result in a transgression of the safe operating space.

In the other Earth system processes, the net-zero designs reduce impacts, sometimes substantially, e.g., in ocean acidification (-91 %), climate change (-90 %), the phosphorus cycle (-86 %), change in biosphere integrity (-82 %), and land system change (-57 %).

However, no net-zero energy system stays within all planetary boundaries when downscaling is applied based on population. In particular, all energy systems transgress boundaries for the nitrogen cycle and atmospheric aerosol loading, while some further exceed boundaries for climate change, change of biosphere integrity, and ocean acidification. The designs obtained for the compromise scenario and the scenario with minimal CO<sub>2</sub> sequestration exceed boundaries for climate change despite reaching net-zero operational greenhouse gas emissions due to greater infrastructure intensity with embedded emissions.

The transgression is particularly large for the nitrogen cycle, where the boundaries are exceeded by a factor of 4.2 on average across the three net-zero designs due to massive investments in battery electric vehicles, power-to-methane, and insulation material for energetic rehabilitation.

While all three net-zero designs transgress at least 2 planetary boundaries and are therefore unsustainable, the designs differ substantially in their environmental impact (Figure 1): On average, the energy system occupies 77 % of the safe operating space for the case minimizing

total annualized cost with unconstrained CO<sub>2</sub> sequestration (min-TAC) but by up to 190 % for the case with minimal CO<sub>2</sub> sequestration (min-stor). The results indicate that design choices, such as the availability of flexible negative emission technologies, can reduce transgressions.

The results vary depending on the choice of downscaling methods, which involves distributive justice considerations and introduces subjectivity into the assessment [21]. Therefore, we include downscaling by gross domestic product instead of population as an additional reference.

Downscaling by gross domestic product instead of population quadruples the safe operating space due to Germany's large per-capita gross domestic product. For downscaling by gross domestic product, an energy system design within the modeled planetary boundaries seems possible if CO<sub>2</sub> sequestration is unconstrained (Figure 1, min-TAC). However, downscaling by economic indicators is controversial [21]. In general, the

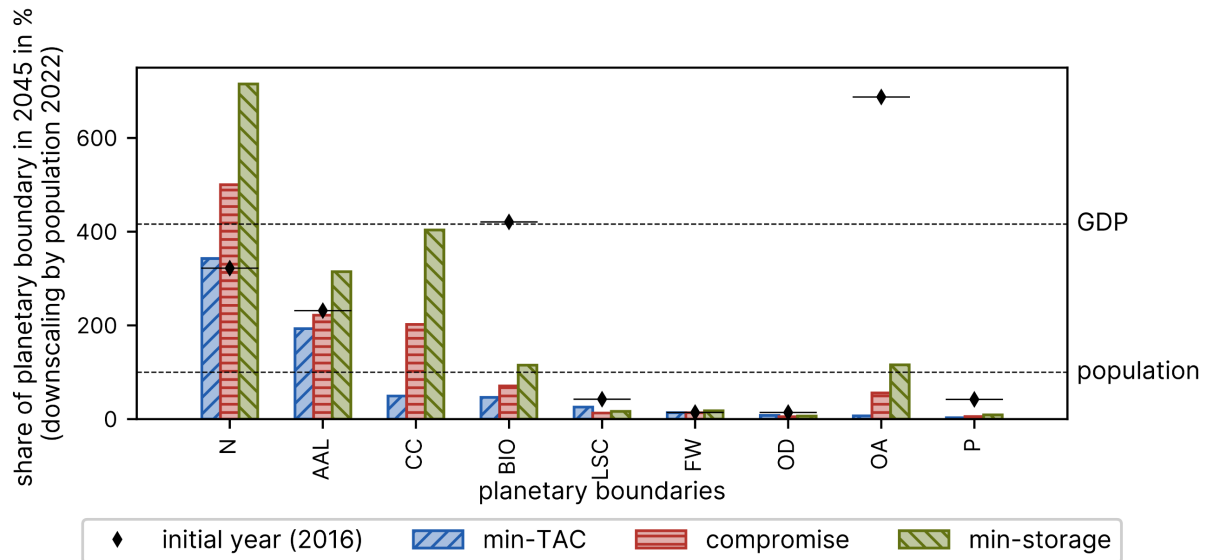
downscaling method requires careful consideration in the interpretation of results.

## CONCLUSIONS AND OUTLOOK

Global greenhouse gas emissions must decline rapidly to limit human-induced climate change. In addition to climate change, other sustainability challenges must be addressed simultaneously.

The planetary boundaries framework defines a safe operating space for human activities for 8 Earth-system processes in addition to climate change. The planetary boundaries thus impose additional constraints on the design space of sustainable energy systems that are commonly neglected.

Here, we evaluate the environmental impacts of net-zero energy system designs considering the planetary boundaries. In particular, we consider net-zero designs of the sector-coupled energy system of Germany, a representative industrial economy. Our case study reveals a



**Figure 1.** Share of safe operating space occupied by sector-coupled energy system designs with net-zero operational greenhouse gas emissions for the German energy transition in 2045 aiming for minimal cost (min-TAC) or minimal CO<sub>2</sub> sequestered (min-storage), and an intermediate solution (compromise). Downscaling of the global safe operating space to the share of safe operating space allocated to the German energy system is based on 1) the population of Germany in 2021 and 2) on the energy system's share of the total environmental impacts in Germany in the original year of the transition horizon (2016) (population). As a reference, downscaling based on gross domestic product in 2021 instead of population is indicated as well (GDP).

The share of safe operating space occupied by the energy system in the reference year (2016) is marked (♦) for comparison with the energy systems in 2045.

abbreviations: nitrogen cycle (N), atmospheric aerosol loading (AAL), climate change (CC), change in biosphere integrity (BIO), land-system change (LSC), freshwater use (FW), stratospheric ozone depletion (OD), ocean acidification (OA), phosphorous cycle (P)

Novel entities are excluded, as they are not quantified in [13]. However, in [11] the planetary boundary for novel entities is considered transgressed if any synthetic chemical is released to the environment without adequate safety testing. Hence, the boundary is likely transgressed for the energy system.



transgression of at least 2 planetary boundaries for all net-zero designs. At the same time, the transgressions vary substantially across designs, indicating opportunities to address transgressions via design choices.

The present work demonstrates the need to include all planetary boundaries in the design of sustainable energy systems. This perspective leads to multiobjective optimization for design space exploration to determine technological barriers and supply chain contributions to environmental impacts. We thus aim to identify enablers of energy systems within planetary boundaries, which will be presented in future work.

## ACKNOWLEDGEMENTS

DS has received funding from the pilot project DemoUpCARMA, funded and supported by the Swiss Federal Office of Energy and the Federal Office for the Environment. AB acknowledges funding by the Swiss Federal Office of Energy's SWEET programme as part of the project PATHFNDR.

## REFERENCES

1. Pfenninger S, Hawkes A, and Keirstead J. Energy systems modeling for twenty-first century energy challenges. *Renew Sust Energ Rev* 33:74–86 (2014).
2. Baumgärtner N, Deutz S, Reinert C, Nolzen N, Kuepper LE, Hennen M, Hollermann DE, and Bardow A. Life-Cycle Assessment of Sector-Coupled National Energy Systems: Environmental Impacts of Electricity, Heat, and Transportation in Germany Till 2050. *Front Energy Res* 9 (2021).
3. Volkart K, Mutel CL, and Panos E. Integrating life cycle assessment and energy system modelling: Methodology and application to the world energy scenarios. *Sustain Prod Consum* 16:121–133 (2018).
4. McDowall W, Solano Rodriguez B, Usubiaga A, and Acosta Fernández J. Is the optimal decarbonization pathway influenced by indirect emissions? Incorporating indirect life-cycle carbon dioxide emissions into a European TIMES model. *J Clean Prod* 170:260–268 (2018).
5. Vandepaer L, Panos E, Bauer C, and Amor B. Energy System Pathways with Low Environmental Impacts and Limited Costs: Minimizing Climate Change Impacts Produces Environmental Cobenefits and Challenges in Toxicity and Metal Depletion Categories. *Environ Science Technol* 54:5081–5092 (2020).
6. DIN Deutsches Institut für Normung e.V. Environmental management – Life cycle assessment – Principles and framework (Beuth Verlag GmbH) (2/2021).
7. Algunaibet IM, and Guillén-Gosálbez G. Life cycle burden-shifting in energy systems designed to minimize greenhouse gas emissions: Novel analytical method and application to the United States. *J Clean Prod* 229:886–901 (2019).
8. Bjørn A, Chandrakumar C, Boulay A-M, Doka G, Fang K, Gondran N, Hauschild MZ, Kerkhof A, King H, and Margni M, et al. Review of life-cycle based methods for absolute environmental sustainability assessment and their applications. *Environ Res Lett* 15:83001 (2020).
9. Rockström J, Steffen W, Noone K, Persson A, Chapin FS, Lambin EF, Lenton TM, Scheffer M, Folke C, and Schellnhuber HJ, et al. A safe operating space for humanity. *Nature* 461:472–475 (2009).
10. Steffen W, Richardson K, Rockström J, Cornell SE, Fetzer I, Bennett EM, Biggs R, Carpenter SR, Vries W de, and Wit CA de, et al. Sustainability. Planetary boundaries: guiding human development on a changing planet. *Science* 347:1259855 (2015).
11. Richardson K, Steffen W, Lucht W, Bendtsen J, Cornell SE, Donges JF, Drüke M, Fetzer I, Bala G, and Bloh W von, et al. Earth beyond six of nine planetary boundaries. *Sci Adv* 9:eadh2458 (2023).
12. Ryberg MW, Owsianiak M, Richardson K, and Hauschild MZ. Development of a life-cycle impact assessment methodology linked to the Planetary Boundaries framework. *Ecol Indic* 88:250–262 (2018).
13. Bachmann M, Zibunas C, Hartmann J, Tulus V, Suh S, Guillén-Gosálbez G, and Bardow A. Towards circular plastics within planetary boundaries. *Nat Sustain*:599–610 (2023).
14. Weidner T, Galán-Martín Á, Ryberg MW, and Guillén-Gosálbez G. Energy systems modeling and optimization for absolute environmental sustainability: current landscape and opportunities. *Comput Chem Eng* 164:107883 (2022).
15. Algunaibet IM, Pozo C, Galán-Martín Á, Huijbregts MAJ, Mac Dowell N, and Guillén-Gosálbez G. Powering sustainable development within planetary boundaries. *Energy Environ Sci* 12:1890–1900 (2019).
16. Stranddorf LK, Clavreul J, Prieur-Vernat A, and Ryberg MW. Evaluation of life cycle impacts of European electricity generation in relation to the Planetary Boundaries. *Sustain Prod Consum* (2023).
17. Negri V, Klukowski SH, and Vázquez D. Absolute life cycle optimization of the CDR-power nexus. In Proceedings of the Foundations of Computer-Aided Process Operations - Chemical Process Control conference, January 8–12, San Antonio, USA (FOCAPO-CPC 2023) (2023).
18. Weidner T, and Guillén-Gosálbez G. Planetary boundaries assessment of deep decarbonisation

- options for building heating in the European Union. *Energy Convers Manag* 278:116602 (2023).
19. Langiu M, Shu DY, Baader FJ, Hering D, Bau U, Xhonneux A, Müller D, Bardow A, Mitsos A, and Dahmen M. COMANDO: A Next-Generation Open-Source Framework for Energy Systems Optimization. *Comput Chem Eng* 152:107366 (2021).
  20. Shu DY, Deutz S, Winter BA, Baumgärtner N, Leenders L, and Bardow A. The role of carbon capture and storage to achieve net-zero energy systems: Trade-offs between economics and the environment. *Renew Sust Energ Rev* 178:113246 (2023).
  21. Ryberg MW, Andersen MM, Owsianiak M, and Hauschild MZ. Downscaling the planetary boundaries in absolute environmental sustainability assessments – A review. *J Clean Prod* 276:123287 (2020).
  22. World Population Prospects 2022. Summary of results (United Nations) (2023).
  23. Hartmann J, and Assen N von der. The Planetary Footprint of Nations: An Absolute Environmental Sustainability Assessment of National Economies. (in preparation) (2024).
  24. National Accounts Section of the United Nations Statistics Division (2023). GDP/breakdown at current prices in US Dollars (all countries), <https://unstats.un.org/unsd/amaapi/api/file/2>.
  25. Reinert C, Schellhas L, Mannhardt J, Shu DY, Kämper A, Baumgärtner N, Deutz S, and Bardow A. SecMOD: An Open-Source Modular Framework Combining Multi-Sector System Optimization and Life-Cycle Assessment. *Front Energy Res* 10 (2022).
  26. Wernet G, Bauer C, Steubing B, Reinhard J, Moreno-Ruiz E, and Weidema B. The ecoinvent database version 3 (part I): overview and methodology. *Int J Life Cycle Assess* 21:1218–1230 (2016).

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Sustainable Green Hydrogen Transport: A Systematic Framework for the Design of the whole Supply Chain

Elvira Spatolisano, Laura A. Pellegrini

<sup>a</sup> GASP - Group on Advanced Separation Processes & GAS Processing, Dipartimento di Chimica, Materiali e Ingegneria Chimica "G. Natta", Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy

\* Corresponding Author: [elvira.spatolisano@polimi.it](mailto:elvira.spatolisano@polimi.it).

## ABSTRACT

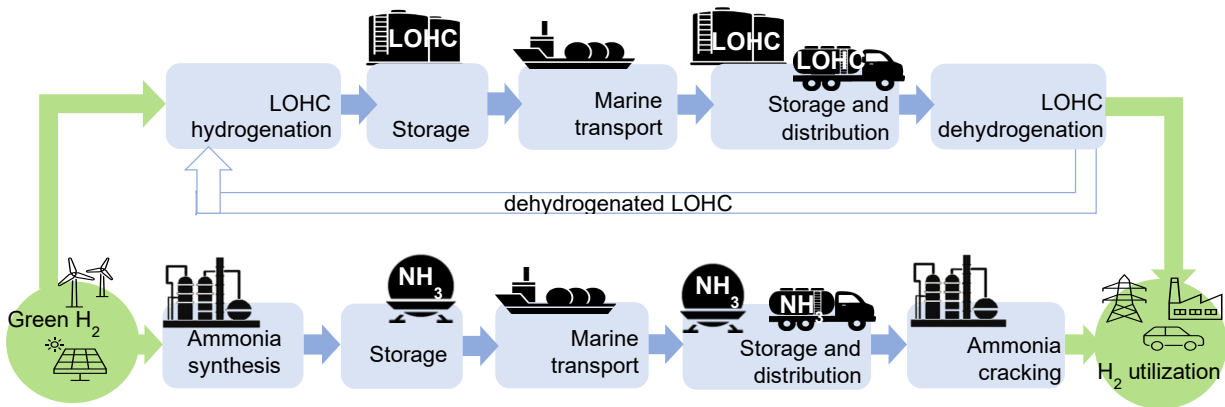
In view of achieving the decarbonization target, green hydrogen is commonly regarded as the alternative capable of reducing the share of fossil fuels. Despite its wide application as a chemical on industrial scale, hydrogen utilization as an energy vector still suffers from unfavorable economics, mainly due to its high cost of production, storage and transportation. To overcome the last two of these issues, different hydrogen carriers have been proposed. Hydrogen storage and transportation through these carriers involve: 1. the carrier hydrogenation, exploiting green hydrogen produced at the loading terminal, where renewable sources are easily accessible, 2. the storage and transportation of the hydrogenated species and 3. its subsequent dehydrogenation at the unloading terminal, to favour H<sub>2</sub> release. Although there is a number of studies in literature on the economic feasibility of hydrogen transport through different H<sub>2</sub> vectors, very few of them delve into the technical evaluation of the hydrogen value chain. From the process design point of view, the hydrogenation and dehydrogenation stages are of paramount importance, considering that they are the cost drivers of the whole system. This work aims to address this gap by presenting a systematic methodology to technically analyse different hydrogen vectors. For the sake of example, ammonia and dibenzyltoluene are considered. Weaknesses of the overall value chain are pointed out, to understand where to focus research efforts for future process intensification.

**Keywords:** H<sub>2</sub> transport, H<sub>2</sub> carriers, sustainable energy, techno-economic assessment, computer-aided process design.

## INTRODUCTION

In the transition towards sustainable energy, green hydrogen has emerged as a promising low-emission alternative. Nonetheless, its transportation is hampered by its low volumetric density. To address this problem, different H<sub>2</sub> carriers have been proposed as a reliable solution. The typical H<sub>2</sub> value chain is depicted in Figure 1, considering ammonia (NH<sub>3</sub>) and liquid organic hydrogen carriers (LOHCs) as hydrogen vectors. NH<sub>3</sub> and LOHCs are commonly perceived as the most encouraging choices, due to the easiness adaptability to the existing infrastructures. These hydrogen-bearing molecules can be hydrogenated at the loading terminal, exploiting green hydrogen produced through renewable sources, more easily stored, transported and, upon arrival, dehydrogenated for H<sub>2</sub> release. The released hydrogen can serve either the mobility or the industrial sectors [1].

NH<sub>3</sub> is a global commodity, already produced in large scale facilities and distributed worldwide. On the other hand, a variety of LOHCs have been considered in literature [2-3]. While some chemical structures are more susceptible to hydrogenation/dehydrogenation than others, certain rules have been identified for selecting favourable compounds. The optimal LOHC has to show low melting point and high boiling point, to avoid solidification and volatilization issues; high H<sub>2</sub> storage capacity; low



**Figure 1:** Hydrogen value chain exploiting LOHCs and  $\text{NH}_3$  as  $\text{H}_2$  carriers.

dehydrogenation enthalpy; low toxicity and low cost [2]. Selected the best candidate as hydrogen carrier, the essential starting point for the feasibility study of the hydrogen value chain is the identification of the basis of the design, including:

1. production rate of green  $\text{H}_2$ , to be fed to the system. Specifically, tuning the plant size according to the land footprint of renewables is crucial for the assessment of realistic scenarios. Renewable sources are characterized by a power density several orders of magnitude lower than fossil fuels. The transition pathways towards sustainability must consider both the limitations of available land and the specific geophysical conditions [4]. Understanding the extent of land needs can put the feasible scale of green hydrogen production into perspective [5]. In this analysis, flat  $\text{H}_2$  production of  $20000 \text{ Nm}^3/\text{h}$  is supposed via 100 MW alkaline electrolyzers, available at 20 bar and  $25^\circ\text{C}$ .
2. Loading and unloading terminal location and, consequently, distance to be covered for  $\text{H}_2$  transport. Different scenarios can be inferred, as the long-distance harbour-to-harbour hydrogen transport, which involves the  $\text{H}_2$  seaborne transport or the short distance hydrogen transport, that implies the road or pipeline hydrogen transport. In the hydrogen value chain of Figure 1, the seaborne  $\text{H}_2$  transport is considered.
3.  $\text{H}_2$  utilization and its centralized or decentralized application. According to the centralized scenario, the  $\text{H}_2$  produced is conveyed into a power plant for green electricity production. In this case, less stringent specifications on  $\text{H}_2$  purity are needed, likely. On the other hand, for the decentralized  $\text{H}_2$  utilization,  $\text{H}_2$  has to be distributed to several hydrogen refuelling stations, to serve the mobility

sector. Thus, high  $\text{H}_2$  purity is necessary. The selection of the centralized/decentralized scenario also affects the process design of the whole value chain and, consequently, the operating conditions of the delivered hydrogen (*i.e.*, temperature and pressure of discharge at the end user, together with required purity). In the present study,  $\text{H}_2$  centralized application at a  $\text{H}_2$  valley is considered, such that the hydrogen product is released at 30 bar and with a purity of 99.9 mol%.

Defined the basis of design, both technical and economic assessments can be carried out. However, it is worth noticing that the economic evaluations are, generally, strongly dependent on the hypotheses introduced and on the methodology adopted. Variable results can be obtained, such that it is difficult to draw general conclusions. For this reason, from the process engineering point of view, it is more relevant to focus on the design of the whole value chain and its technical assessment, with a particular attention to the cost-driving stages, rather than demonstrating the economic viability of hydrogen transport. In this respect, this work aims at presenting a systematic methodology to analyse different hydrogen value chains. Ammonia and dibenzyltoluene (H0-DBT) are selected as the representative carriers, due to their promising features. Hydrogenation and dehydrogenation stages have been designed in Aspen Plus V11<sup>®</sup>. Based on the simulations presented, a detailed technical analysis is discussed, to pave the way for future process intensification.

## SIMULATION OF THE COST-DRIVING PROCESSES

### Ammonia ( $\text{NH}_3$ )

Ammonia is a global commodity, already distributed worldwide and used in the chemical industry. Recently, it



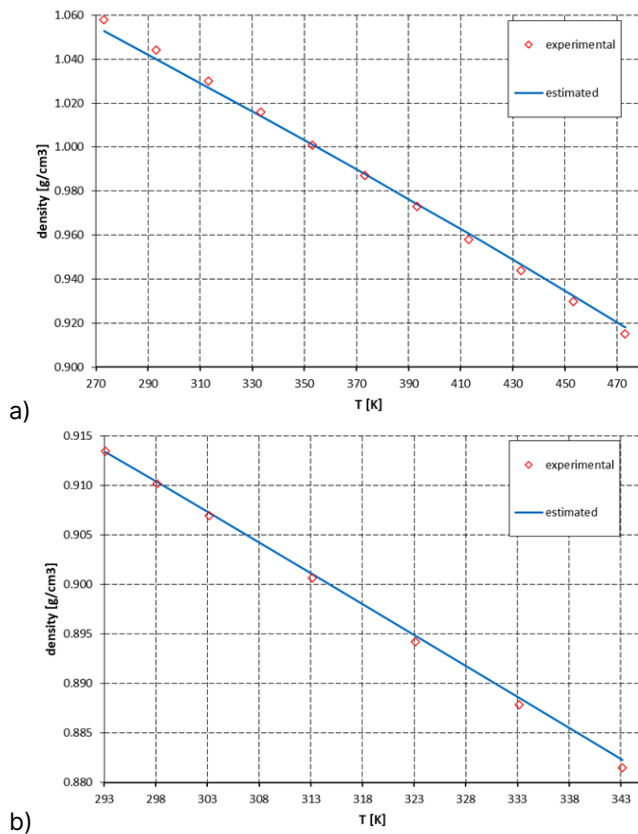


including the new pseudo-components in Aspen Plus simulation software. Once these properties are introduced, the compounds physical-chemical characterization occurs in Aspen Plus® with the aid of NIST TDE (ThermoData Engine) tool and the Aspen property estimation system.

**Table 1:** Dibenzyltoluene (DBT) and perhydrodibenzyltoluene (H18-DBT) properties [10].

Property	Value	
	DBT	H18-DBT
TB [°C]	390	371
MW [kg/kmol]	272.4	290.4
SG* @ 60°F [-]	1.047	0.916
$\Delta H_{\text{form}}$ [kJ/kmol]	225700	-387400

\*SG: Specific Gravity.



**Figure 3:** Comparison between experimental and predicted density values for: a) H0-DBT and b) H18-DBT.

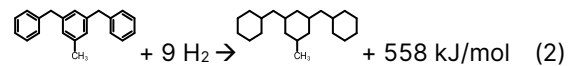
In addition to what provided in Table 1, supplying experimental molar volume and ideal gas heat capacity can enable the development of a more robust model. Implemented all the available experimental datasets, in view

of assessing the reliability of the models, the predicted physical properties can be compared with the experimental ones. Results are reported in Figure 3 considering the density of the hydrogenated and dehydrogenated species for the sake of example. A satisfactory agreement is obtained between the experimental and calculated values.

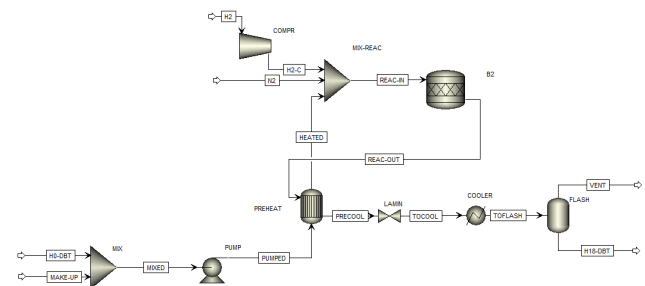
### H0/H18 DBT hydrogenation and dehydrogenation

Once the process simulator has been set up, both hydrogenation and dehydrogenation stages have been designed in Aspen Plus V11®.

The hydrogenation section is represented in Figure 4. The process receives at its battery limits the hydrogen produced via the alkaline electrolyzers, at 50°C and 20 bar. This stream undergoes compression up to 35 bar and is fed to the reactor (*B2* in Figure 4), together with a nitrogen stream that acts as a thermal diluent. Also dibenzyltoluene coming both from the make up stream and the dehydrogenation process is fed to the hydrogenation reactor, after preheating in a process-process heat exchanger (*PREHEAT* in Figure 4). The reactor operates at a pressure of 35 bara and a temperature of 210°C for the catalytic hydrogenation reaction (2) to occur. While the exothermic nature of the hydrogenation reaction would suggest lower temperatures for enhanced reaction equilibrium, the operating temperature also considers the catalytic activity requirements [11]. Based on what available in literature [12], the reactor is simulated through the *RStoic* module of Aspen Plus at quantitative conversion and selectivity.



Downstream the reaction section, the outlet stream, which contains the perhydro-dibenzyltoluene along with unreacted hydrogen and nitrogen, is cooled and expanded to favour the separation of light gases from the heavier H18-DBT, which can be sent for storage and then transported to the spot of hydrogen delivery.



**Figure 4:** H0-DBT hydrogenation section. Simulation in Aspen Plus V11®.



As for the pair H0-DBT/H18-DBT, the process simulation suffers from the not detailed reactor modeling, such that fundamental research is needed to identify a well-established kinetic expression to be included in the Aspen Plus scheme. Despite the drawback of low maturity, the dibenzyltoluene shows the intrinsic advantage of a reduced volatility, which allows the production of high purity hydrogen after dehydrogenation, facilitating the H<sub>2</sub> separation from unconverted gases downstream the reaction section.

The hydrogenation section electricity consumption amounts to 615.96 kW, for the compression and pumping of reactants at the battery limits, while only cooling water is needed as external utility, to provide a cooling duty of 482.3 kW. The high endothermicity of the hydrogenation reaction (16462.3 kW) could represent a plus point in the overall energy balance. When designing the whole value chain, strategies for heat integration have to be figured out to understand how to maximise revenues.

Concerning the dehydrogenation section, 3758.97 kW of electricity consumption are registered for the H18-DBT dehydrogenation process. Despite this number could seem high at a first glance, most of the electric energy required by the process is due to the hydrogen compression up to the delivery pressure, which cannot be reduced. Thermal energy consumption of the process is related to the cooling duty by means of cooling water, which accounts for a total of 4822 kW, together with the heat necessary for the hydrogenation reaction to occur. Despite no external utility is required, the high reaction endothermicity (19887.3 kW) is responsible for a dramatic reduction of the hydrogen flow rate exiting the battery limits. For 103.83 kmol/h of H18-DBT entering the process, 575.65 kmol/h of hydrogen are produced. The choice of coping the reaction endothermicity by burning part of the hydrogen produced is for sustainability purposes. It is true that H<sub>2</sub> is a high value-added product, probably too valuable to be burnt in a combustion chamber, but, on the other hand, the utilization of any fossil-based fuels does not make sense environmentally, if CO<sub>2</sub> is emitted upon combustion. Research efforts for process intensification should focus, primarily, on the identification of a suitable organic molecule to be used as H<sub>2</sub> vector, to reduce the heat requirements of the dehydrogenation section as much as possible. Selected the best candidate as LOHC, heat integration in the whole hydrogen value chain, together with the selection of a suitable clean burning fuel, compliant with plant location and needs, could be strategies for energy savings and, ultimately, operating cost reduction.

## CONCLUSIONS

Green hydrogen transport through different H<sub>2</sub> carriers still suffers from unfavourable economics. With the

aim of understanding where to focus research efforts for process intensification, this work offers a systematic methodology for the analysis of the value chain cost drivers, *i.e.*, hydrogenation and dehydrogenation stages. Ammonia and the pair dibenzyltoluene/perhydrodibenzyltoluene are selected as representative carriers.

For ammonia value chain, focal points of research to be investigated for cost reductions are:

- the NH<sub>3</sub> synthesis process intensification on small scale, to enable green ammonia production at milder operating conditions;
- the NH<sub>3</sub> thermocatalytic cracking process intensification, to enable the reaction to occur at milder operating conditions and to optimize the N<sub>2</sub>-H<sub>2</sub> separation downstream the reactor as much as possible.

For LOHC value chain, active and fundamental research is still needed to identify the most promising candidate to be used as the hydrogen carrier, to assess its physical-chemical properties and to investigate the hydrogenation and dehydrogenation reaction kinetics at the laboratory scale. Furthermore, points to be investigated are:

- for the hydrogenation section, the maximisation of the heat integration, taking into account the reaction exothermicity;
- for the dehydrogenation section, the identification of alternative methods for heat supply to cope the reaction endothermicity.

## REFERENCES

1. Roland Berger. Hydrogen transportation. The key to unlocking the clean hydrogen economy (2021).
2. Aakko-Saksa PT, Cook C, Kiviahio J, Repo T. Liquid organic hydrogen carriers for transportation and storing of renewable energy – Review and discussion. *J Power Sources* 396: 803-823 (2018).
3. Díaz E, Rapado-Gallego P, Ordóñez S. Systematic evaluation of physicochemical properties for the selection of alternative liquid organic hydrogen carriers. *J Energy Storage* 59: 106511 (2023).
4. Tran TH, Egermann M. Land-use implications of energy transition pathways towards decarbonisation – Comparing the footprints of Vietnam, New Zealand and Finland. *Energy Policy* 166: 112951 (2022).
5. Spatolisano E, Pellegrini LA. Haber-Bosch process intensification: a first step towards small-scale distributed ammonia production. *Chem Eng Res Des* 195: 651-661 (2023).
6. Restelli F, Spatolisano E, Pellegrini LA, de Angelis AR, Cattaneo S, Roccaro E. Detailed techno-economic assessment of ammonia as green H<sub>2</sub>



- carrier. *Int J Hydrogen Energy* 52: 532-547 (2024).
7. Lucentini I, Garcia X, Vendrell X, Llorca J. Review of the decomposition of ammonia to generate hydrogen. *Ind Eng Chem Res* 60: 18560-18611 (2021).
  8. Spatolisano E, Pellegrini LA, De Angelis AR, Cattaneo S, Roccaro E. Ammonia as a carbon-free energy carrier: NH<sub>3</sub> cracking to H<sub>2</sub>. *Ind Eng Chem Res* 62: 10813-10827 (2023).
  9. Hydrogenious LOHC Technologies. <https://hydrogenious.net/>
  10. Müller K, Stark K, Emel'yanenko VN, Varfolomeev MA, Zaitsau DH, Shoifet E, et al. Liquid Organic Hydrogen Carriers: Thermophysical and Thermochemical Studies of Benzyl- and Dibenzyl-toluene Derivatives. *Ind Eng Chem Res* 54: 7967-7976 (2015).
  11. Spatolisano E, Restelli F, Matichecchia A, Pellegrini LA, de Angelis AR, Cattaneo S, Roccaro E. Assessing opportunities and weaknesses of green hydrogen transport via LOHC through a detailed techno-economic analysis. *Int J Hydrogen Energy* 52: 703-717 (2024).
  12. Shi L, Qi S, Qu J, Che T, Yi C, Yang B. Integration of hydrogenation and dehydrogenation based on dibenzyltoluene as liquid organic hydrogen energy carrier. *Int J Hydrogen Energy* 44: 5345-5354 (2019).
  13. Modisha P, Gqogqa P, Garidzirai R, Ouma CNM, Bessarabov D. Evaluation of catalyst activity for release of hydrogen from liquid organic hydrogen carriers. *Int J Hydrogen Energy* 44: 21926-21935 (2019).
  14. Palys MJ, Daoutidis P. Using hydrogen and ammonia for renewable energy storage: A geographically comprehensive techno-economic study. *Comput Chem Eng* 136: 106785 (2020).
  15. Humphreys J., Lan R, Tao S. Development and Recent Progress on Ammonia Synthesis Catalysts for Haber-Bosch Process. *Adv Energy Sus Res* 2: 2000043 (2021).
  16. Kale MJ, Ojha DK, Biswas S, Militti JI, McCormick AV, Schott JH, Dauenhauer PJ, Cussler EL. Optimizing Ammonia Separation via Reactive Absorption for Sustainable Ammonia Synthesis. *ACS Appl Energy Mater* 3: 2576-2584 (2020).
  17. International Energy Agency. The Future of Hydrogen. Seizing today's opportunities. Report prepared by the IEA for the G20, Japan (2019).

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>

# A mathematical programming optimization framework for wind farm design considering multi-directional wake effect

Javiera Vergara-Zambrano<sup>a</sup>, Styliani Avraamidou<sup>a\*</sup>

<sup>a</sup> University of Wisconsin-Madison, Department of Chemical and Biological Engineering, Madison, Wisconsin, USA

\* Corresponding Author: [avraamidou@wisc.edu](mailto:avraamidou@wisc.edu).

---

## ABSTRACT

The placement of wind turbines is a crucial design element in wind farms, given the energy losses resulting from the wake effect. Despite numerous studies addressing the Wind Farm Layout Optimization (WFLO) problem, considering multiple directions to determine wind turbine spacing and layout remains limited. However, relying solely on one predominant direction may lead to overestimating energy production, and loss of energy generation. This work introduces a novel mathematical programming optimization framework to solve the WFLO problem, emphasizing the wind energy's nonlinear characteristics and wake effect losses. Comparisons with traditional layout approaches demonstrate the importance of optimizing wind farm layouts during the design phase. By providing valuable insights into the renewable energy sector, this research aims to guide future wind farm projects towards layouts that balance economic considerations with maximizing energy production.

---

**Keywords:** Wind, Turbines, Energy Systems, Renewable and Sustainable Energy, Optimization

## INTRODUCTION

Due to the escalating energy crisis, there has been a growing inclination towards generating more energy from renewable sources. Consequently, aside from constructing new power generation facilities such as wind or solar plants, studying how to optimize these installations is crucial [1]. Among the various renewable energy options available, wind energy stands out due to its efficient power generation capacity and the ability to produce energy on a large scale, making it an attractive option for expanding electricity generation capacity [2-5].

Wind power is produced by converting the kinetic energy of air in motion using a turbine. Usually, wind turbines are arranged in groups known as wind farms to increase power production and minimize costs. Several factors influence the energy production of a wind farm, including wind speed and direction, and various meteorological conditions. Specifically, energy losses occur due to the wake effect. As wind turbines extract energy from the wind, a wake forms downstream, reducing the wind speed. Consequently, the placement of wind turbines

significantly impacts the efficiency of a wind farm [6].

Traditionally, wind farms have followed a rule of thumb of placing wind turbines in rows with 8–12 rotor diameters of spacing parallel to the prevailing wind direction and columns spaced 3–5 rotor diameters apart perpendicular to the wind direction [7]. More recently, several studies have been conducted to determine the optimal positioning of wind turbines within a designated land area [8-11]. The primary objective of these studies has been to minimize wake effects and, consequently, maximize expected power production. These studies consider dividing the domain into a grid that defines possible turbine locations. However, only a limited number of studies have extended their focus to include determining the optimal inter-turbine distance. This problem, known as the wind farm layout optimization (WFLO) problem, has garnered substantial research interest.

Given the nonlinearity inherent in wind energy's characteristics, addressing this issue poses a noteworthy difficulty. Proposed approaches to tackle the WFLO problem predominantly involve using data-driven or metaheuristic algorithms such as genetic algorithms,



Abbreviations		Variables	
1D	one wind direction	$A$	land area used ( $m^2$ )
2D	two wind directions	$AC$	annual costs (USD/year)
WFLO	wind farm layout optimization	$C_{inv}$	annual investment costs (USD/year)
<b>Indices</b>		$C_{OM}$	annual operation & maintenance costs (USD/year)
$c$	columns	$N_{tot}$	total number of turbines (-)
$d$	day	$P_{r,c,t,d}$	energy produced at row $r$ , column $c$ at time $t$ of day $d$ (MWh)
$j$	wind direction number	$P_{r,c,t,d}^{curve}$	energy produced according to power curve (MWh)
$r$	rows	$P_{nom}$	nominal capacity of a wind turbine (MW)
$t$	hour time step	$v$	wind velocity considering wake effect (m/s)
<b>Parameters</b>		$v_{r,c,t,d}^{curve}$	wind velocity at row $r$ , column $c$ at time $t$ of day $d$ (m/s)
$\alpha$	shift coefficient (-)	$v_{wd_j,r,c,t,d}$	wind velocity in direction $wd_j$ at row $r$ , column $c$ at time $t$ of day $d$ (m/s)
$A_{max}$	land area available ( $m^2$ )	$v^0$	undisturbed velocity at hub height (m/s)
$C_{inv}^{cap}$	Investment costs per capacity (USD/MW)	$v_{t,d}^0$	undisturbed velocity at hub height time $t$ of day $d$ (m/s)
$C_{OM}^{cap}$	operation & maintenance costs per capacity (USD/MW)	$w_{wd_j,t,d}$	wind direction
$CRF$	capital recovery factor (%)	$WL$	wake effect losses (-)
$Ct$	trusted coefficient (-)	$x$	x-axis distance (m)
$Di$	rotor diameter (m)	$y$	y-axis distance (m)
$f_d$	frequency of representative day $d$ (-)	$y_{r,c}^b$	wind turbine purchase at row $r$ , column $c$
$H$	hub height (m)	$z$	characteristic distance (m)
$H_i$	altimeter height (m)		
$i$	interest rate (%)		
$k$	rate of wake expansion (-)		
$L$	wind turbines lifetime (years)		
$M$	big M (-)		
$v_{t,d}^m$	wind velocity at altimeter height at time $t$ of day $d$ (m/s)		

random search, and particle swarm optimization. While these algorithms are practical for providing near-optimal solutions, they often do not supply guarantees of optimality [12]. Furthermore, to our knowledge, none of the studies have considered more than one dominant wind direction to determine the spacing and layout of wind turbines, even though promising locations for wind farm installations, such as the Texas Panhandle, have at least two dominant wind directions, and some studies have suggested that ignoring wind direction could lead to an overestimation of the wind energy production [13].

This study introduces a novel mathematical programming optimization framework designed to determine the optimal position and spacing of wind turbines. It explores various design objective criteria, including costs and energy production. The study considers the intrinsic nonlinear attributes of wind energy, incorporates modeling of wake effect losses, and, for the first time, considers both single and multiple dominant wind directions. To illustrate its efficacy, the proposed model is applied in a case study focusing on the energy transition of the University of Wisconsin-Madison campus.

## PROBLEM FORMULATION

### Wake effect model

The wind losses due to the wake effect are model using the Jensen model, which is one of the most widely used wake model [14-15]. It assumes that the diameter of the wake increases linearly in proportion to the downstream distance,  $z$ . The speed downstream can be calculated as

$$v(z) = v^0 \cdot [1 - (1 - \sqrt{1 - Ct}) \left( \frac{Di}{Di + 2kz} \right)^2] \quad (1)$$

, where  $v^0$  is the undisturbed incoming velocity,  $k$  is the rate of the wake expansion and have a value of 0.075 for onshore wind,  $Di$  is the diameter of the rotor of the wind turbine and  $Ct$  is the trust coefficient, which as a value of 0.8 [16-17].

In this study, the wake effect model was adapted to consider multiple wind directions. For each wind direction  $wd_j$  the velocity is calculated considering in characteristic distance  $z_j$ . The calculation of the wind velocity for the two directions  $wd_1$  and  $wd_2$  is presented below, while an illustration of the two chosen directions is presented in Figure 1. The main assumption is that the wind velocity reaching the turbines ( $v_{wd_j,r,c,t,d}$ ) that are in the same row ( $r$ ) or column ( $c$ ) are not affected by each other. The

losses due to the wake effect will be denoted as  $WL$ , rewriting equation (1) as the following:

$$v(z) = v^0 \cdot [1 - WL(z)] \quad (2)$$

For wind direction  $wd_1$ , the velocity at each position at time  $t$  of day  $d$  ( $v_{wd_1,r,c,t,d}$ ) can be calculated as:

- For the first row ( $r = 1$ )

$$v_{wd_1,1,c,t,d} = v_{t,d}^0 \quad \forall c, t, d \quad (3)$$

- For any other row ( $\forall r \neq 1, c, t, d$ )

$$v_{wd_1,r,c,t,d} = v_{wd_1,r,c-1,t,d} \cdot [1 - y_{r,c}^b \cdot WL(y)] \quad (4)$$

Similarly, for wind direction  $wd_2$ , the velocity ( $v_{wd_2,r,c,t,d}$ ) is calculated as:

- For the first column ( $c = 1$ )

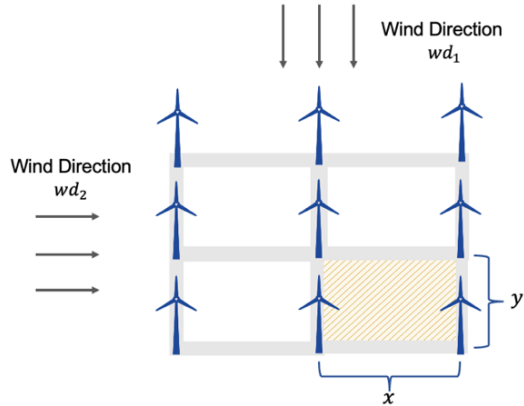
$$v_{wd_2,r,1,t,d} = v_{t,d}^0 \quad \forall r, t, d \quad (5)$$

- For any other column ( $\forall r, c \neq 1, t, d$ )

$$v_{wd_2,r,c,t,d} = v_{wd_2,r,c-1,t,d} \cdot [1 - y_{r,c}^b \cdot WL(x)] \quad (6)$$

The binary variable  $y_{r,c}^b$  denotes if the turbine in row  $r$ , and column  $c$  is installed (1 if it is installed, 0 if not). Lastly, it is assumed that the wind only blows in one direction at each time step, given by equation (6). We denote the parameter  $w_{wd,t,d}$  which has a value of 1 if the wind blows in direction  $wd$  at time  $t$  of day  $d$ , and 0 if not.

$$v_{r,c,t,d}^{curve} = \sum_{wd} v_{wd,r,c,t,d} \cdot w_{wd,t,d} \quad (7)$$



**Figure 1.** Wind farm layout and turbine spacing.

## Wind farm layout optimization model

The power output of the wind turbines ( $P_{r,c,t,d}$ ) was calculated using piecewise linear approximation of the power curve of a wind turbine Vestas V112-3.08, selected considering the average size of wind turbine used in the market [18]. Figure 2 shows the results of the linearization done. Regarding the operation of the wind farm, the turbines only produce energy if they are installed.

$$P_{r,c,t,d} \leq M \cdot y_{r,c}^b \quad \forall r, c, t, d \quad (8)$$

$$P_{r,c,t,d} \leq P_{r,c,t,d}^{curve}(v_{r,c,t,d}^{curve}) \quad \forall r, c, t, d \quad (9)$$

The area used by each turbine is considered to be a rectangle corresponding to the shaded area in Figure 1, and the total area available is limited (Equation 10 and 11).

$$A = N_{tot} \cdot x \cdot y \quad (10)$$

$$A \leq A_{max} \quad (11)$$

Two objective functions are considered in this study: minimization of annual costs, and maximization of the annual energy produced. The annual costs ( $AC$ ) include the annual investment costs ( $C_{inv}$ ) and the annual operation and maintenance costs ( $C_{OM}$ ), as expressed by equations (12).

$$AC = C_{inv} + C_{OM} \quad (12)$$

The initial investment costs ( $C_{inv}^{cap}$ ) are converted into annual investment cost per capacity using the Capital Recovery Factor (CRF) [19], defined as

$$CRF = \frac{i}{1 - (1+i)^{-L}} \quad (13)$$

, where  $i$  is the interest rate and  $L$  the lifetime of the wind turbines. In this study, an interest rate of 7% and a lifetime of 30 years were assumed [20]. The annual investment and operation and maintenance costs can be expressed by equations (14) and (15), respectively.

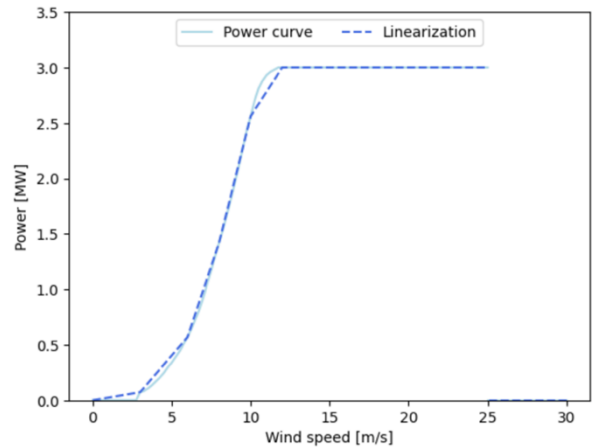
$$C_{inv} = N_{tot} \cdot P_{nom} \cdot C_{inv}^{cap} \cdot CRF \quad (14)$$

$$C_{OM} = N_{tot} \cdot P_{nom} \cdot C_{OM}^{cap} \quad (15)$$

The annual energy produced ( $AEP$ ) is calculated as the sum of the energy produced at each hour and day represented by equation (16).

$$AEP = \sum_{r,c,t,d} P_{r,c,t,d} \cdot f_d \quad (16)$$

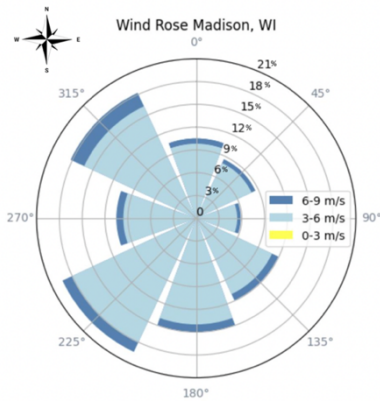
, where  $f_d$  is the frequency of the representative day  $d$ .



**Figure 2.** Power curve of the turbine Vestas V112-3.08 [21].

## CASE OF STUDY

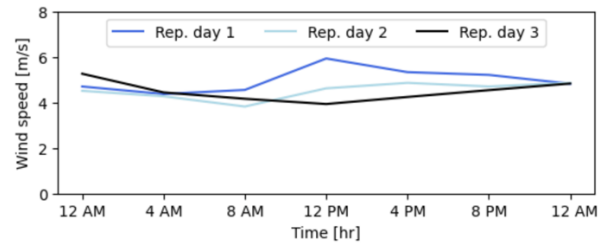
To demonstrate the applicability of the developed model formulation and framework, a case study was conducted to design a wind farm in Madison, Wisconsin as an option for energy transition in the University of Wisconsin-Madison. The wind data used were obtained from the AOSS Tower located at the University of Wisconsin-Madison [22]. A typical meteorological year was constructed using data from 2013 to 2023. The annual wind source distribution can be observed in Figure 3. There are two predominant wind directions: SW and NW, with an average wind speed of 4.8 m/s at 30 m (altimeter height). It is noteworthy that the wind speed is below 3 m/s only 1% of the year (yellow label), indicating an area with potential for the installation of a wind farm.



**Figure 3.** Wind speed and direction distribution in Madison, Wisconsin.

The proposed model is classified as a mixed integer nonlinear problem (MINLP) problem, where the nonlinear terms are associated with the wake effect. To account for the variability and intermittency of the wind resource, and at the same time make the problem tractable, the year wind data was reduced to three representative days using hierarchical clustering. Hierarchical clustering is a machine learning algorithm that groups similar data points into nested clusters based on their proximity, forming a tree-like structure [23]. The three representative days selected are presented in Figure 4. Before using the data in the model, the wind speed inputs were adjusted from the measured height ( $H_i$ ) to the hub height of the turbine ( $H$ ) using equation (17) and considering a shift coefficient ( $\alpha$ ) of 0.14 [21]. The model was implemented in GAMS [24] and solved using the BARON solver [25].

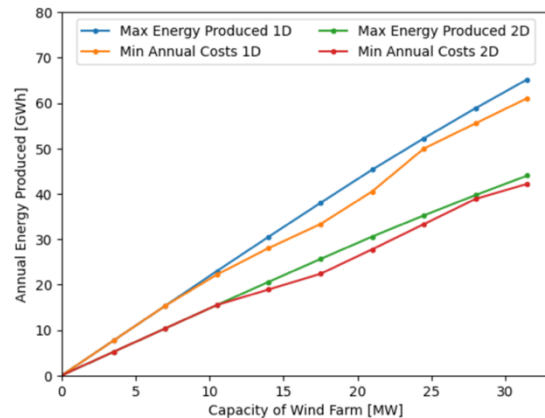
$$v_{t,d}^0 = v_{t,d}^m \cdot \left(\frac{H}{H_i}\right)^\alpha \quad \forall t, d \quad (17)$$



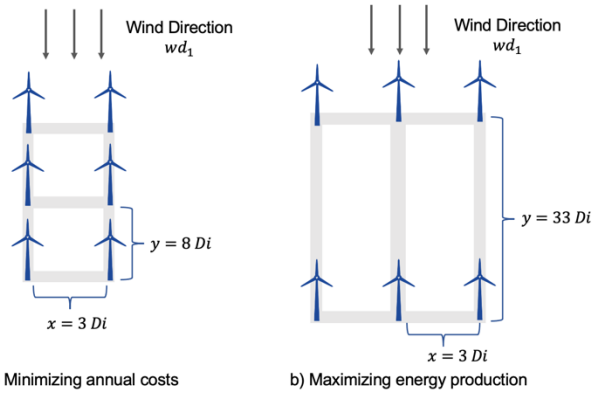
**Figure 4.** Representative days for wind speed using hierarchical clustering.

## RESULTS & DISCUSSION

In this study, we approached the problem by examining one and two wind directions ( $wd_1$  and  $wd_2$ ) for up to nine wind turbines. Figure 5 presents the results for minimizing annual costs and maximizing yearly energy production. Utilizing costs as the primary design criterion for a wind farm may not be ideal, as it does not optimize turbine placement and overall area utilization, consequently reducing the energy output. The findings reveal a potential reduction in energy production of up to 13% for identical-capacity wind farms when costs are minimized compared to the maximization of energy production. Maximizing energy output appears more effective in optimizing the wind farm layout, as it reduces energy losses. However, it is essential to evaluate the land area used. The model tends to position turbines further apart to maximize energy, which might require a more extensive area. Figure 6 illustrates different arrangements leading to varying energy production for the installation of six turbines for reference. In both cases, the annual costs are the same, but the configurations differ. Minimizing costs results in layouts with more closely spaced rows and a higher number of rows, increasing wake effect energy losses. On the other hand, maximizing energy production leads to a larger distance between turbines, utilizing more land.



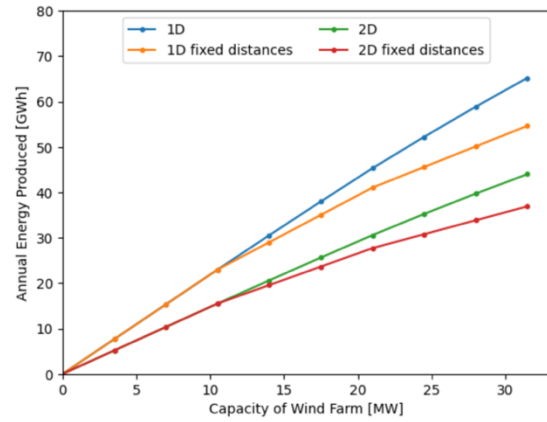
**Figure 5.** Annual energy produced considering the minimization of annual costs and maximization of energy produced.



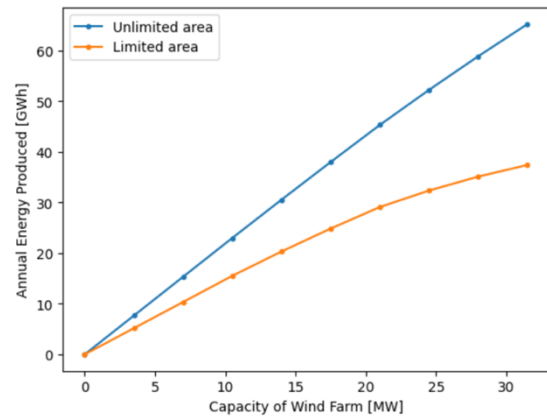
**Figure 6.** Layout of installing 6 wind turbines considering a) minimization of annual costs b) maximization of annual energy production.

Figure 7 compares energy production when considering the model with one or two wind directions and using Patel's rule of thumb. The findings indicate that adhering to the rule of thumb for turbine placement could lead to lower energy production; therefore, optimizing the layout is a crucial aspect of wind farms and should be evaluated during design. Considering only the predominant wind direction, i.e., one direction, could lead to overestimating energy production, which can affect the project's financial aspects. Taking into consideration multiple wind directions represents the system more accurately, as the wind blows in multiple directions in many locations.

A sensitivity analysis was conducted regarding the land area available. The results, depicted in Figure 8, indicate that the model is sensitive to this value. It is important to note, that in practical scenarios, the land availability will be either already bought or modifying the existing land may involve associated costs or logistical constraints, imposing limitation for its modification. Considering an unlimited area can lead to an overestimated energy production by up to 15%, according to the model developed. Moreover, assuming an unlimited or vast area suggests a linear relationship between capacity and energy produced, neglecting wake effects and energy losses. As the number of turbines increases, so does the wake effect, resulting in a nonlinear increase in energy production.



**Figure 7.** Annual energy production considering different number of wind speed directions and distances between turbines.



**Figure 8.** Annual energy produced considering different values for land area available.

## CONCLUSION

This study introduced a comprehensive optimization framework to design an optimal layout for a wind farm, accounting for the wake effect in various directions. The energy system was formulated as a MINLP problem with two objective functions: minimizing annual costs and maximizing energy production, showing a difference of up to 13% in the energy produced between both criteria. The results indicate that considering only the predominant wind direction could lead to an overestimation of the energy produced. The model was solved using up to two wind directions, but future work should extend this to evaluate the ideal number of wind directions to consider when designing wind farms in different locations, considering solving time and model accuracy. Additionally, coping with the uncertainties associated with wind farm parameters is an important area for future work. The following steps for this work will also focus on extending the analysis to higher wind farm capacities and evaluate extreme weather scenarios to refine the model's accuracy and applicability.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge financial support from the University of Wisconsin-Madison.

## REFERENCES

1. Hoxha B, Shesho I, Filkoski R. Analysis of Wind Turbine Distances Using a Novel Techno-Spatial Approach in Complex Wind Farm Terrains. *Sustainability* 14: 1–16 (2022)
2. Balasubramanian K, Babu T, Subramaniam U, Sudhakar N, Sichilalu S. A novel review on Optimization Techniques used in Wind Farm Modelling. *Renew. Energy Focus* (2020)
3. Kakodkar R, He G, Demirhan CD, Arbabzadeh M, Baratsas SG, Avraamidou S, Mallapragada D, Miller I, Allen RC, Gençer E, Pistikopoulos EN. A review of analytical and optimization methodologies for transitions in multi-scale energy systems. *Renewable and Sustainable Energy Reviews*, 160: 112277 (2022)
4. Allen RC, Nie Y, Avraamidou S, Pistikopoulos EN. Infrastructure planning and operational scheduling for power generating systems: An energy-water nexus approach. In *Computer Aided Chemical Engineering* 47: 233-238 (2019)
5. Chen Z, Avraamidou S, Liu P, Li Z, Ni W, Pistikopoulos EN. Optimal design of integrated urban energy systems under uncertainty and sustainability requirements. *Computers & Chemical Engineering* 155 (2021)
6. Manousakis N, Psomopoulos C, Ioannidis G, and Kaminaris, S. Optimal placement of wind turbines using semidefinite programming. *IET Conf. Publ.*, 711: 12–15 (2016)
7. Patel MR. Wind and Solar Power System. Taylor & Francis (2006)
8. Cook J, Di Martino M, Allen R, Pistikopoulos E, Avraamidou S. A decision-making framework for the optimal design of renewable energy systems under energy-water-land nexus considerations. *Sci. Total Environ* (2022)
9. Ulku I, Alabas-Uslu C. A new mathematical programming approach to wind farm layout problem under multiple wake effects. *Renew. Energy* (2018)
10. Archer R, Nates G, Donovan S, Waterer H. Wind Turbine Interference in a Wind Farm Layout Optimization Mixed Integer Linear Programming Model. *Wind Eng* 35:2 (2011)
11. Turner S, Romero D, Zhang P, Amon C, Chan T. A new mathematical programming approach to optimize wind farm layouts. *Renew. Energy* (2018)
12. Stanley A, Roberts O, King J, Bay, C. Objective and algorithm considerations when optimizing the number and placement of turbines in a wind power plant. *Wind Energy Sci* (2021)
13. Narain A, Srivastava S, Singh S. The impact of wind direction on wind farm power output calculation considering the wake effects of wind turbines. *Wind Engineering* 47:1 (2023)
14. Jensen NO. A note on wind generator interaction. Technical. Report Risø National Laboratory (1983)
15. Katic I, Højstrup J, Jensen NO. A simple model for cluster efficiency. *Eur. Wind Energy Assoc. Conf. Exhib* (1986)
16. Troen I, Lundtang Petersen E. European wind atlas. Risø National Laboratory (1989)
17. Adams A, Keith D. Are global wind power resource estimates overstated?. *Environ. Res. Lett.* 8 (2013)
18. Wiser R, Bolinger M, Hoen B, Millstein D, Rand J, Barbose G, Darghouth N, Gorman W, Jeong S, O'Shaughnessy E, Paulos B. *Land-Based Wind Report: 2023 Edition* (2023)
19. Maleki A, Askarzadeh A. Comparative study of artificial intelligence techniques for sizing of a hydrogen-based stand-alone photovoltaic/wind hybrid system. *Int. J. Hydrogen Energy* 39:19 (2014)
20. International Renewable Energy Agency (IRENA). *Renewable Power Generation Costs in 2022*. Tech Report (2022)
21. National Renewable Energy Laboratory (NREL). *System Advisor Model Version 2022.11.29 (SAM) Power Curve Library*.
22. Space Science & Engineer Center UW-Madison. AOSS Tower. <https://metobs.ssec.wisc.edu/aoss/tower/>
23. Murtagh F, Contreras P. Algorithms for hierarchical clustering: an overview II. *WIREs Data Mining Knowl Discov* (2017)
24. Bussieck M.R., Meeraus A. General algebraic modeling system (GAMS). *Modeling Languages in Mathematical Optimization*. In: Applied Optimization Springer US (2004)
25. Sahinidis N.V. Baron: a general purpose global optimization software package. *J. Glob. Optim.* 8 (1996)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# RiNSES<sup>4</sup>: Rigorous Nonlinear Synthesis of Energy Systems for Seasonal Energy Supply and Storage

Yifan Wang<sup>a</sup>, Marvin Volkmer<sup>a</sup>, Dörthe Franzisca Hagedorn<sup>a</sup>, Christiane Reinert<sup>a</sup>, and Niklas von der Assen<sup>a,\*</sup>

<sup>a</sup> Institute of Technical Thermodynamics, RWTH Aachen University, 52062 Aachen, Germany

\* Corresponding Author: [niklas.vonderassen@itt.rwth-aachen.de](mailto:niklas.vonderassen@itt.rwth-aachen.de).

## ABSTRACT

The synthesis of energy systems necessitates simultaneous optimization of both design and operation across all components within the energy system. In real-world applications, this synthesis poses a mixed-integer nonlinear programming (MINLP) problem, considering nonlinear behaviours such as investment cost curves and part-load performance. The complexity increases further when seasonal energy storage is involved, as it requires temporal coupling of the full time series. Although numerous solution approaches exist to solve the synthesis problems simplified by linearization, methods for solving a full-scale problem are currently missing. In this work, we introduce a rigorous method, RiNSES<sup>4</sup>, to manage the nonlinear aspects of energy system synthesis, particularly focusing on long-term time-coupling constraints. RiNSES<sup>4</sup> calculates the upper and lower bounds of the initial synthesis problem in two separate branches. The proposed method yields feasible solutions through upper bounds, while evaluating the solution quality via lower bounds. The solution quality is iteratively enhanced by increasing the resolution for calculating upper bounds and tightening the relaxations for computing lower bounds. Both branches work simultaneously and independently, with their outcomes compared after each iteration within each branch. The iterations continue until a predefined optimality gap is reached. We apply RiNSES<sup>4</sup> to design a photovoltaic and battery energy system, considering the seasonality of both energy supply and demand sides. In comparison with a state-of-the-art commercial solver, RiNSES<sup>4</sup> enables to solve the MINLP synthesis problem with great temporal detail and shows high potential.

**Keywords:** Mixed-integer nonlinear programming, time series aggregation, linearization, decomposition, relaxation

## 1. INTRODUCTION

Mathematical modeling and optimization can aid in identifying the optimal design and operation of energy systems, spanning from industrial to international scale. The synthesis problem of an energy system necessitates simultaneous optimization of both design and operation, across all components within the energy system [1]. At the design level, the types and sizes of energy system components are determined. At the operation level, decisions are made regarding the on/off status and load allocations for each time step. In general, the synthesis poses a mixed-integer nonlinear programming (MINLP) problem, taking into account nonlinear behaviors such as investment cost curves at the design level and part-load

performance at the operation level [2]. However, solving an MINLP problem is generally challenging due to its intrinsic complexity. The complexity increases further when incorporating seasonal energy supply and storage, as it requires extensive temporal data input as well as temporal coupling of the full time series [3].

In the literature, energy system modelers adopt various approaches to address synthesis problems, aiming to achieve computationally tractable results. Kotzur et al. [4] applied time series aggregation methods to efficiently reduce the size of synthesis problems. Gabrielli et al. [5] and Kotzur et al. [6] further proposed alternative modeling approaches to reduce the complexity of synthesis problems from the temporal coupling aspects. However, such approaches only yield solutions for a simplified

version of the synthesis problem.

To tackle the synthesis problem employing the full time series, Baumgärtner et al. [7] developed the RiSES<sup>4</sup> method. The method is designed to solve the linear synthesis of energy systems with seasonal storage, ensuring a solution with known quality. RiSES<sup>4</sup> integrates time series aggregation [8] with superposition seasonal storage modeling [6] in the synthesis problem, and subsequently solve an operational optimization problem with full time series directly through commercial solvers. RiSES<sup>4</sup> employs a rigorous method [9] for measuring the quality of the resulting solutions. In cases where the resulting operational optimization problems remain computationally challenging, a decomposition-based method, DeLoop, proposed by Baumgärtner et al. [10], could be potentially incorporated to more effectively address the long-term operational optimization of energy systems. RiSES<sup>4</sup> has been applied to design an industrial energy system using a mixed-integer linear programming (MILP) formulation, as well as a national energy system using a linear programming (LP) formulation, demonstrating promising performance [7].

Nevertheless, the application of RiSES<sup>4</sup> is limited to linear synthesis problems. The nonlinear nature of an energy system's synthesis is neglected in advance. As a result, the solution obtained for the linearized problem might be infeasible for the initial nonlinear synthesis problem.

### 1.1. Contribution of this work

In this work, we propose the RiNSES<sup>4</sup> method, an extension of the RiSES<sup>4</sup> method, specifically designed to address the nonlinear aspects of energy system synthesis.

Similar to the RiSES<sup>4</sup> method, RiNSES<sup>4</sup> independently computes the upper and lower bounds of a synthesis problem, but with an MINLP formulation. For calculating the upper bounds, building upon RiSES<sup>4</sup>, we use a linearized problem with aggregated time series to find design candidates. These design candidates are then fixed, and the initial MINLP problem is solved as an operational optimization problem to obtain a feasible solution, thereby establishing an upper bound for the initial problem. For computing the lower bounds, we relax various constraints within the synthesis problem to accelerate computation. We iteratively improve the solution quality by increasing the resolution for calculating upper bounds and tightening the relaxation for computing lower bounds. To assess the performance of the proposed method, we apply RiNSES<sup>4</sup> to design an energy system including photovoltaic panels and a battery for seasonal energy supply and storage.

The structure of this paper is as follows: Section 2 explicates the problem statement and the RiNSES<sup>4</sup> method. In Section 3, the proposed method is applied to a case study. Section 4 concludes the work.

## 2. METHOD

### 2.1. General MINLP formulation

$$\begin{aligned}
 & \min_{\dot{E}_k^N, E_k^N, \dot{E}_{e,t}^{\text{buy}}, \dot{E}_{e,t}^{\text{sell}}, \dot{E}_{k,e,t}^{\text{in}}, \dot{E}_{k,e,t}^{\text{out}}, E_{k,e,t}, \mathbf{y}, \mathbf{z}} TAC = \frac{1}{APVF} CAPEX + OPEX & (1a) \\
 \text{with } & CAPEX = \sum_{k \in \mathcal{K}} I_k = \sum_{k \in \mathcal{K} \setminus \mathcal{K}^{\text{stor}}} I_k^{\text{ref}} \left( \frac{\dot{E}_k^N}{\dot{E}_k^{\text{N,ref}}} \right)^{M_k} & (1b) \\
 & \quad + \sum_{k \in \mathcal{K}^{\text{stor}}} I_k^{\text{ref}} \left( \frac{E_k^N}{E_k^{\text{N,ref}}} \right)^{M_k} \\
 & OPEX = \sum_{k \in \mathcal{K}} c_k^m I_k + \sum_{t \in \mathcal{T}} \Delta t_t \sum_{e \in \mathcal{E}^{\text{ext}}} (c_{e,t}^{\text{buy}} \dot{E}_{e,t}^{\text{buy}} - c_{e,t}^{\text{sell}} \dot{E}_{e,t}^{\text{sell}}) & (1c) \\
 \text{s.t. } & \sum_{k \in \mathcal{K}} (\dot{E}_{k,e,t}^{\text{out}} - \dot{E}_{k,e,t}^{\text{in}}) + \dot{E}_{e,t}^{\text{buy}} - \dot{E}_{e,t}^{\text{sell}} = \dot{E}_{e,t}^{\text{D}} & (1d) \\
 & \quad \forall e \in \mathcal{E}, \forall t \in \mathcal{T} \\
 & E_{k,e,t} (1 - \eta_{k,e}^{\text{self}} \Delta t_t) + \Delta t_t \left( \eta_{k,e}^{\text{in}} \dot{E}_{k,e,t}^{\text{in}} - \frac{\dot{E}_{k,e,t}^{\text{out}}}{\eta_{k,e}^{\text{out}}} \right) = E_{k,e,t+1} & (1e) \\
 & \quad \forall k \in \mathcal{K}^{\text{sto}}, \forall t \in \mathcal{T} \\
 & E_{k,e,t=1} = E_{k,e,t=|\mathcal{T}|+1} & (1f) \\
 & \quad \forall k \in \mathcal{K}^{\text{sto}} \\
 & \mathbf{g}(\dot{E}_k^N, E_k^N, \mathbf{y}, \mathbf{z}) \leq 0 & (1g) \\
 & \quad \forall k \in \mathcal{K}, \forall e \in \mathcal{E} \\
 & \mathbf{h}(\dot{E}_k^N, E_k^N, \dot{E}_{e,t}^{\text{buy}}, \dot{E}_{e,t}^{\text{sell}}, \dot{E}_{k,e,t}^{\text{in}}, \dot{E}_{k,e,t}^{\text{out}}, E_{k,e,t}, \mathbf{y}, \mathbf{z}) \leq 0 & (1h) \\
 & \quad \forall k \in \mathcal{K}, \forall e \in \mathcal{E}, \forall t \in \mathcal{T} \\
 & \dot{E}_k^N, E_k^N, \dot{E}_{e,t}^{\text{buy}}, \dot{E}_{e,t}^{\text{sell}}, \dot{E}_{k,e,t}^{\text{in}}, \dot{E}_{k,e,t}^{\text{out}}, E_{k,e,t} \in \mathbb{R}^+ & (1i) \\
 & \quad \forall k \in \mathcal{K}, \forall e \in \mathcal{E}, \forall t \in \mathcal{T} \\
 & \mathbf{y} \in R^{N_y}, \mathbf{z} \in \{0,1\}^{N_z} & (1j) \\
 & \quad \forall k \in \mathcal{K}, \forall e \in \mathcal{E}, \forall t \in \mathcal{T}
 \end{aligned}$$

Problem 1

In Problem 1, we state the generic synthesis problem of an energy system using an MINLP formulation. To start with, we assign all components inside the energy system ( $k \in \mathcal{K}$ ) to two categories: storage components ( $k \in \mathcal{K}^{\text{sto}}$ ) and conversion components ( $k \in \mathcal{K} \setminus \mathcal{K}^{\text{stor}}$ ). Storage components transport products ( $e \in \mathcal{E}$ ) from one time step to other time steps ( $t \in \mathcal{T}$ ), while conversion components convert one product to any other product(s). Additionally, the category “exogenous inputs” indicates that the energy system is connected with exogenous energy supply ( $\tilde{e} \in \mathcal{E}^{\text{ext}} \subseteq \mathcal{E}$ ) and demand ( $\dot{E}_{e,t}^{\text{D}}$ ).

We minimize the total annualized costs  $TAC$  that consists of the capital and operational expenditures,  $CAPEX$  and  $OPEX$  (Equation (1a)). The annualized present value factor is defined as  $APVF = \frac{(1+i)^n - 1}{(1+i)^n \cdot i}$  with the interest rate  $i$  and the number of periods  $n$ .  $CAPEX$  is the sum of the investment costs of all energy system components, which follow the capacity power law (Equation (1b)). The investment costs  $I_k$  of the component  $k$  are defined as the reference cost  $I_k^{\text{ref}}$  multiplied by the ratio of the nominal size  $\dot{E}_k^{\text{N}}$  or  $E_k^{\text{N}}$  and a reference value  $\dot{E}_k^{\text{N,ref}}$  or  $E_k^{\text{N,ref}}$  to the power of  $M_k \leq 1$ . Please note that the indicators for conversion and storage components differ in their respective units.  $OPEX$  are defined as the sum of maintenance costs and external energy costs of each time step, where  $c_k^m$ ,  $c_{e,t}^{\text{buy}}$  and  $c_{e,t}^{\text{sell}}$  indicate a maintenance factor, and external energy buying and selling prices, respectively (Equation (1c)).  $\Delta t_t$  represents the length of time step  $t$ .

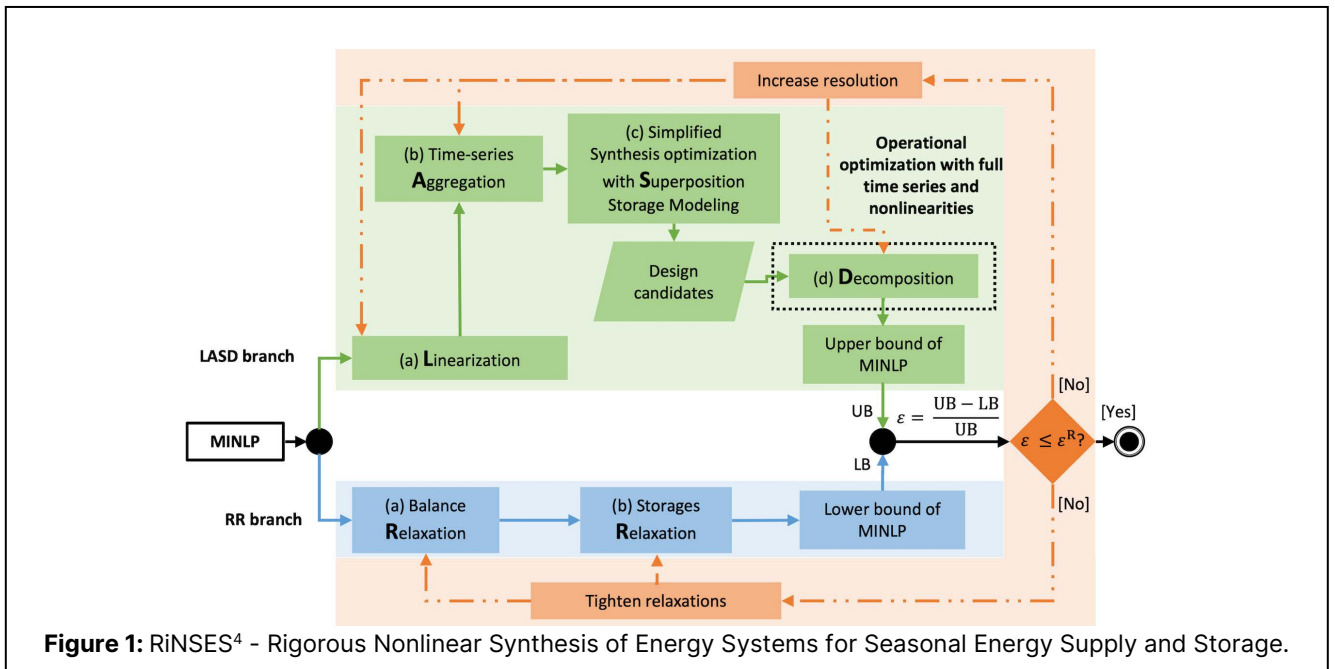
Given the input time series, the designed energy system needs to fulfill the energy demands for each time step through energy conversion or by purchasing energy from external energy supplies (Equation (1d)). Here,  $\dot{E}_{k,e,t}^{\text{in}}$

and  $\dot{E}_{k,e,t}^{\text{out}}$  denote the input and output flows, respectively, of product  $e$  at time step  $t$  for component  $k$ . The state of charge  $E_{k,e,t+1}$  of a storage component  $k$  at the time step  $t + 1$  depends on its' state of charge  $E_{k,e,t}$  at time step  $t$  and the product input and output flow at the time step  $t$ , considering the efficiencies of self-discharge  $\eta_{k,e}^{\text{self}}$ , charging  $\eta_{k,e}^{\text{in}}$  and discharging  $\eta_{k,e}^{\text{out}}$  (Equation (1e)). The so-called cycling constraint ensures that the product is conserved (Equation (1f)). All other constraints are concisely encapsulated in Equations (1g) and (1h), including the part-load performance of each conversion component. All decision variables and their respective bounds are summarized in Equations (1i) and (1j). The surrogate vectors  $\mathbf{y}$  and  $\mathbf{z}$  represent other decision variables that are not specified here, encompassing investment decisions and on/off decisions for each energy system component.

In addressing both the design and operation levels of an energy system synthesis problem, we utilize the index  $t$  to differentiate between design variables (without the index  $t$ ) and operation variables (with the index  $t$ ). This distinction also extends to constraints: design constraints, which are independent of the number of time steps, and operation constraints, which recur at each time step. This distinction is crucial for the method we propose in the following section.

## 2.2. RiNSES<sup>4</sup> Method

The proposed RiNSES<sup>4</sup> method solves Problem 1 with the full time series. As depicted in Figure 1, RiNSES<sup>4</sup> handles the initial MINLP problem in two separate branches to compute the upper (illustrated as the green segment in Figure 1) and lower bounds (shown as the blue segment in Figure 1) of the MINLP problem,



**Figure 1:** RiNSES<sup>4</sup> - Rigorous Nonlinear Synthesis of Energy Systems for Seasonal Energy Supply and Storage.

respectively. Both branches incorporate independent iterations (represented by the red segment in Figure 1) to improve the solution quality.

The upper bounds (UBs) are feasible solutions to the MINLP Problem. To calculate an upper bound for Problem 1, we use four distinct techniques: linearization (L), time-series aggregation (A), superposition storage modeling (S) and decomposition (D). The tetrad of techniques is collectively referred to as the LASD branch. A detailed description of the LASD branch is provided in Section 2.2.1. The quality of the optimal solution derived from the LASD branch is evaluated using lower bounds (LBs). The lower bounds for Problem 1 are computed via a two-stage relaxation (R) process, named as the RR branch. A detailed discussion of the RR branch is provided in Section 2.2.2. After each iteration, RiNSEs<sup>4</sup> compares the current UB and LB to calculate the current optimality gap  $\varepsilon$ , as described in detail in Section 2.2.3. We iteratively improve the solution quality by increasing the resolution for calculating UBs and tightening the relaxations for computing LBs. The two branches work simultaneously and individually from each other. The RiNSEs<sup>4</sup> method terminates if a predefined optimality gap  $\varepsilon^R$  is reached.

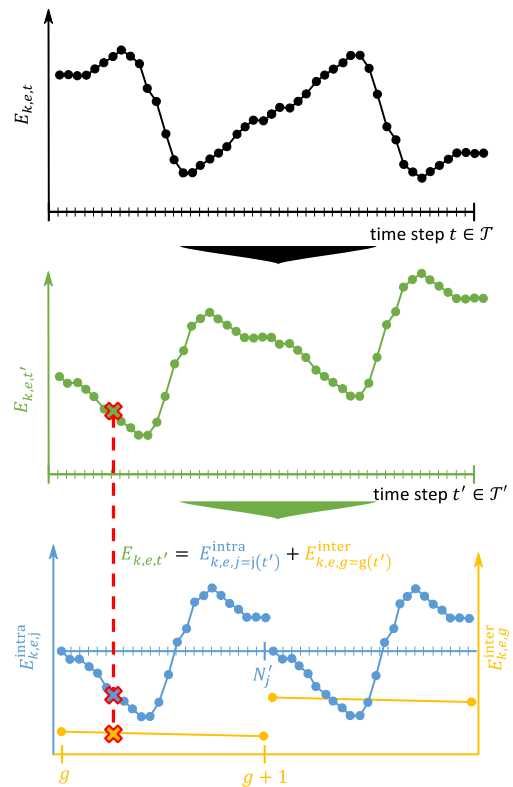
### 2.2.1. Calculating the upper bounds in LASD branch

In the LASD branch, feasible solutions (upper bounds) of the initial synthesis, Problem 1, are calculated based on four steps, as represented in the green segment in Figure 1: We first simplify the synthesis problem in steps (a) and (b) to obtain values for the design variables in step (c). Then, we validate the feasibility of the design variables in an operational optimization problem with the full time series and all nonlinearities in step (d). To efficiently solve the resulting MINLP operational problem, a decomposition approach can be applied in step (d). If there is a feasible solution to the operational optimization problem, the solution is also a feasible solution to Problem 1 and, thus, an upper bound (UB) for Problem 1.

In step (a), we first linearize nonlinearities of the initial MINLP problem to obtain an MILP formulation. The nonlinearities include investment cost curves, as shown in Equation (1b), and nonlinear part-load performance, as outlined in Equation (1h). The linearization is based on the non-separable piecewise-linear optimization approach proposed by Vielma et al. [11], which can be readily integrated using a Python package [12]. Through linearization, we simplify the complexity of the nonlinear constraints. However, the linearization results in an increase in the number of the binary decision variables in the resulting MILP problem, potentially leading to longer computing time.

Therefore, in step (b), we further simplify the synthesis problem through time series aggregation to reduce the size of *exogenous inputs*. As explained in [8] and in

[9], the initial input time series  $\mathcal{T}$ , consisting of  $N_t$  time steps, is divided into  $N_k$  periods  $\mathcal{P}$  with  $N_j$  time steps in each period, where  $N_k = N_t/N_j$ . Through time series aggregation,  $N_k$  and  $N_j$  are aggregated to  $N'_k$  typical periods  $\mathcal{P}'$  with  $N'_j$  segments in each typical period, i.e., the number of time steps is reduced to  $N'_k \times N'_j$ , which is significantly smaller than  $N_t$ . In this work, we employ the tsam Python package [13] for aggregating time-series data, taking advantage of its various available aggregation approaches. Kotzur et. al [4] analyzed the impact of various aggregation approaches on their optimization results, revealing that the choice of the aggregation approach had only minor impacts. In our study, we hence adopt the k-mean aggregation, a prevalent approach in energy system optimization, for time series aggregation.



**Figure 2:** Overview of the superposition storage modeling. The figure is adapted from Kotzur et al. [6].

Upon completing steps (a) and (b), we derive a simplified synthesis problem in step (c). Here, to consider product transport on a seasonal scale, we adopt the superposition storage modeling approach, initially introduced by Kotzur et al. [6] and also implemented in RiSES<sup>4</sup> [7]. In our work, we extend the model by incorporating typical segments. Figure 2 illustrates the basic principles of this superposition modeling approach.

In step (c), the state of charge associated with the initial input time series  $E_{k,e,t}$  ( $t \in \mathcal{T}$ ), illustrated in the top

part of Figure 2, is replaced with the state of charge for the aggregated time series  $E_{k,e,t'}$  ( $t' \in \mathcal{T}'$ ), shown in the middle part of Figure 2. The superposition modeling approach decomposes  $E_{k,e,t'}$  (in green) into two parts: the intra-period state of charge  $E_{k,e,j}^{\text{intra}}$  ( $j \in \{1, \dots, N'_j\}$ , colored in blue) and inter-period state of charge  $E_{k,e,g}^{\text{inter}}$  ( $g \in \{1, \dots, N'_g\}$ , colored in orange), depicted in the bottom part of Figure 2.  $E_{k,e,j}^{\text{intra}}$  describes the storage behavior within the typical periods, whereas  $E_{k,e,g}^{\text{inter}}$  represents the storage behavior between these periods. The relationships  $j = j(t')$  and  $g = g(t')$  denote the corresponding values of  $j$  and  $g$  associated with the time step  $t'$ .

Please note that in superposition storage modeling, the sequence of time steps is of paramount importance. Therefore, we use the symbol  $N'_g$  to denote the typical periods, rather than  $N'_k$  mentioned earlier. The key difference between  $N'_g$  and  $N'_k$  is that  $N'_g$  encompasses the sequence of occurrence of typical periods, and, thus, has the same size as the periods  $\mathcal{P}$  previously described. For an in-depth explanation of superposition storage modeling approach, please refer to references [6, 7].

As a consequence, for the *exogenous inputs*, the initial full time series  $\mathcal{T}$  with  $N_t$  time steps is reduced to the aggregated time series  $\mathcal{T}'$  with  $N'_g \times N'_j$  time steps. In the following, we refer to the linearized and aggregated synthesis problem with superposition storage modeling as Problem 2. In Problem 2, Equations (2a) to (2d) are applied as operation constraints for all storage components, replacing the constraints (1e) and (1f).

$$\frac{E_{k,e,j+1}^{\text{intra}} - E_{k,e,j}^{\text{intra}}}{\Delta t_j} = -\eta_{k,e}^{\text{self}} E_{k,e,j}^{\text{intra}} \quad (2a)$$

$$+ \Delta t_j \left( \eta_{k,e}^{\text{in}} \hat{E}_{k,e,j}^{\text{in}} - \frac{E_{k,e,j}^{\text{out}}}{\eta_{k,e}^{\text{out}}} \right)$$

$$E_{k,e,g+1}^{\text{inter}} = E_{k,e,g}^{\text{inter}} (1 - \eta_{k,e}^{\text{self}} \Delta t_t)^{N'_j} + E_{k,e,N'_j+1}^{\text{intra}} \quad (2b)$$

$$E_{k,e,1}^{\text{inter}} = E_{k,e,N'_g+1}^{\text{inter}} \quad (2c)$$

$$E_{k,e,t'} = E_{k,e,j=j(t')}^{\text{intra}} + E_{k,e,g=g(t')}^{\text{inter}} \quad (2d)$$

Problem 2, formulated as an MILP problem with a small-scale input time series  $\mathcal{T}'$ , facilitates an efficient solution process in step (c). Based on the results of Problem 2, we fix the design variables of the original MINLP synthesis problem to the solution of Problem 2, receiving an MINLP operational optimization problem. The resulting operational optimization problem with the full time series is a large-scale MINLP problem (Problem 3).

In step (d), we solve the resulting Problem 3. Depending on its size, we first try to solve Problem 3 using a commercial MINLP solver. If Problem 3 cannot be solved within a predetermined computing time frame, we employ the DeLoop method, developed by Baumgärtner et al. [10]. DeLoop handles time-coupled long-term operational optimization problems via decomposition and parallel computing, and it systematically reduces the number

of decomposed subproblems in an iterative manner. In the worst-case scenario, the original MINLP operational problem is tackled.

If Problem 2 provides feasible design decisions for Problem 3, where the full time series and nonlinearities at the operation level is addressed, we yield an upper bound (UB) to the initial synthesis problem by combining the design decisions of Problem 2 and the operation decisions of Problem 3. However, we cannot guarantee that the design decisions identified in Problem 2 allows feasible operation in Problem 3. If an infeasibility occurs, we iteratively repeat steps (b)-(d) with enhanced time series resolutions, until a feasible solution is found, or the original size of the input time series is applied in step (b). If no feasible solution could be found during aggregation, the resulting superposition storage modeling in step (c) are equivalent to Equations 1(e) and 1(f) in the original synthesis problem. If all attempted steps still result in infeasibility, the final move is to increase the number of breakpoints during linearization. As the number of breakpoints increases, so does the accuracy of the linearization. In the case of infinitely many breakpoints, the linearized problem closely approximates the original MINLP synthesis problem. Should this scenario arise, the LASD branch addresses the original MINLP synthesis problem.

With increasing time series resolution and accuracy of linearization, the whole LASD branch converges to the original MINLP synthesis problem in the worst-case scenario. In practice, we observed that the iterative approach consistently identified a feasible solution before necessitating the use of the full-scale time series in Problem 2. Thus, this work has not yet explored the aspect of infinite number of breakpoints. During implementation, we define a maximum number of breakpoints, for example, four. When the LASD branch reaches this maximum, it forcibly transitions to tackling the original MINLP synthesis problem.

### 2.2.2. Computing the lower bounds in RR branch

In the RR branch, we underestimate Problem 1 using two-stage relaxations and obtain a relaxed synthesis problem. Solving the relaxed synthesis problem provides a lower bound, which serves as the lower bound for Problem 1, as visually represented by the blue segment in Figure 1.

In stage (a), we employ the same time series aggregation methods as those in the LASD branch, which leads to typical periods  $P'$  with aggregated segments  $S$  within each typical period. In each segment, we identify the maximal and minimal external energy demands  $\hat{E}_{e,t}^{\text{D,max}}$  and  $\hat{E}_{e,t}^{\text{D,min}}$  for each energy form ( $\forall e \in \mathcal{E}$ ). Subsequently, we relax the constraints by replacing Equation (1d) with Equations (3a) and (3b). Please note that, unlike the LASD branch, we maintain the initial size and sequence of the input time series  $\mathcal{T}$  in this stage. The number of



decision variables of the synthesis problem therefore remains unchanged, only the solution space is larger.

$$\sum_{k \in \mathcal{K}} (\dot{E}_{k,e,t}^{\text{out}} - \dot{E}_{k,e,t}^{\text{in}}) + \dot{E}_{e,t}^{\text{buy}} - \dot{E}_{e,t}^{\text{sell}} \leq \dot{E}_{e,t}^{\text{D,max}} \quad (3a)$$

$$\sum_{k \in \mathcal{K}} (\dot{E}_{k,e,t}^{\text{out}} - \dot{E}_{k,e,t}^{\text{in}}) + \dot{E}_{e,t}^{\text{buy}} - \dot{E}_{e,t}^{\text{sell}} \geq \dot{E}_{e,t}^{\text{D,min}} \quad (3b)$$

In stage (b), we further underestimate Problem 1 by relaxing the constraints associated with the storage components. We replace Equations (1e) and (1f) with Equations (4a) and (4b) for each time step ( $\forall t \in \mathcal{T}$ ), which effectively decouples the constraints between two adjacent time steps. We refer to the resulting large-scale MINLP problem as Problem 4, which we solve directly using an MINLP solver. Upon solving Problem 4, its lower bound is a lower bound to the initial synthesis problem (Problem 1).

$$E_{k,e,t} (1 - \eta_{k,e}^{\text{self}} \Delta t_t) + \Delta t_t \left( \eta_{k,e}^{\text{in}} \dot{E}_{k,e,t}^{\text{in}} - \frac{\dot{E}_{k,e,t}^{\text{out}}}{\eta_{k,e}^{\text{out}}} \right) \geq 0 \quad (4a)$$

$$E_{k,e,t} (1 - \eta_{k,e}^{\text{self}} \Delta t_t) + \Delta t_t \left( \eta_{k,e}^{\text{in}} \dot{E}_{k,e,t}^{\text{in}} - \frac{\dot{E}_{k,e,t}^{\text{out}}}{\eta_{k,e}^{\text{out}}} \right) \leq E_k^N \quad (4b)$$

In the RR branch, we expand the solution space by relaxing different constraints of Problem 1, potentially resulting in an increased computing time for Problem 4. Consequently, we employ computing time as an additional termination criterion, and adopt the lower bound of Problem 4 as the lower bound for Problem 1. If Problem 4 is infeasible, the RR branch proceeds to the next iteration.

For new iterations, we gradually revert the relaxations: starting with step (b) for each storage component, followed by step (a) for each energy balance equation. This sequential tightening of underestimation continues until, in the worst-case scenario, the original synthesis problem is fully addressed.

### 2.2.3. Optimality gap and Iteration

Finally, as marked by the red segment in Figure 1, we compare the best resulting lower and upper bounds obtained using Equation (5) and verify whether the pre-defined optimality gap  $\varepsilon^R$  is satisfied:

$$\varepsilon = \frac{TAC^{\text{UB}} - TAC^{\text{LB}}}{TAC^{\text{UB}}} \leq \varepsilon^R \quad (5)$$

RiNSES<sup>4</sup> identifies global optimal solutions for the initial MINLP synthesis problem (Problem 1),  $TAC^{\text{UB}}$ , with a known quality  $TAC^{\text{LB}}$ . It assesses the progress by comparing the current optimality gap  $\varepsilon$  against the required optimality gap  $\varepsilon^R$ . RiNSES<sup>4</sup> solves Problem 1 in the LASD

branch and the RR branch with different levels of simplification. If the achieved optimality gap  $\varepsilon$  fails to meet the required optimality gap  $\varepsilon^R$ , new iterations are triggered in both branches.

In the LASD branch, we enhance the time step resolution for time series aggregation in step (b) and, potentially, the number of breakpoints for linearization in step (a). Should Problem 2 or Problem 3 exhibit infeasibility, the computing process proceeds directly to the next iteration. In the RR branch, we undo the relaxations gradually, first for storage components in step (b), then for energy balance equations in step (a). Should Problem 4 encounter infeasibility, the computing process proceeds directly to the next iteration.

The iterations stop as soon as the required optimality gap  $\varepsilon^R$  is reached. The current best upper bound is a feasible solution to the initial nonlinear synthesis problem, Problem 1, with known quality. Since both the LASD and RR branches converge to the original MINLP problem in the worst-case scenario, the RiNSES<sup>4</sup> method guarantees the convergence to the initial MINLP synthesis problem (Problem 1).

## 3. CASE STUDY

### 3.1. Set-up

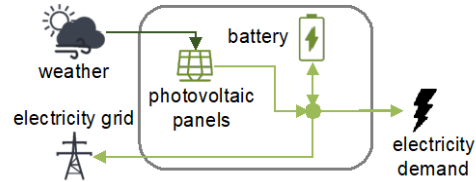


Figure 3: Structure of a PV-BAT energy system.

To evaluate the effectiveness of the proposed method, we apply RiNSES<sup>4</sup> to design a small energy system for seasonal energy supply and storage. We compare the method to an MINLP commercial solver. Thus, we choose a system small enough to be solvable by the MINLP commercial solver as well. As depicted in Figure 3, the energy system includes a conversion component (photovoltaic panels, PV), a storage component (battery, BAT), and three exogenous data inputs (weather, electricity selling price and demand). The renewable electricity generated by photovoltaic panels can be directly used

**Table 1:** Model parameters of the case study for the energy system components at the design level. The energy system components include photovoltaic panels (PVs) and a battery (BAT).

$k \in \mathcal{K}$	$I_k^{\text{ref}}$	$\dot{E}_k^{\text{N,ref}}$	$M_k$	$c_k^m$	$\dot{E}_k^{\text{N}}$
PVs	4264.3	1 $kW_{el}$	0.9592	0.01	[100, 20000] in $kW_{el}$
BAT	$I_k^{\text{ref}}$	$\dot{E}_k^{\text{N,ref}}$	0.8382	0.025	$\dot{E}_k^{\text{N}}$
	2116.1	1 $kWh_{el}$			[40, 600000] in $kWh_{el}$

to fulfill the electricity demand, stored in a battery, or fed into the electricity grid.

In the case study, we assume an interest rate  $i = 8\%$  and an investment period  $n$  of 10 years to calculate the annualized present value factor. The model parameters and the nominal size bounds at the design level are from the literature [14] and summarized in Table 1. For details on the model at the operation level and input time series, please refer to our previous work [15]. Please note that this work adopts a 4-hour resolution, representing the average values derived from every four time steps of the input time series used in our previous research [15]. Additionally, electricity selling prices here are based on the electricity buying prices cited in literature [15].

We set the predefined optimality gap  $\varepsilon^R$  to 5%. All resulting MILP problems are solved by the commercial solver Gurobi [16] with an optimality gap of 1%, whereas all MINLP problems are handled using the commercial solver BARON 22.11.3 [17] with an optimality gap of 0.2%. We model the energy system on the COMANDO platform [12] using Python 3.8. All computational tasks are carried out on Intel® Xeon® W-2155 processors with 3.3GHz and 128 GB RAM.

### 3.2. Results and discussion

The initial synthesis problem of the energy system contains 26,284 constraints and 17,524 decision variables. We address this synthesis problem with two solution approaches:

- Using the state-of-the-art MINLP global solver, BARON 22.11.3 [17], and
- Implementing our proposed method RiNSES<sup>4</sup>.

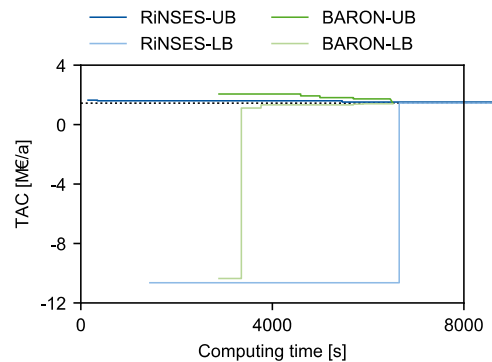
Figure 4 shows the results, illustrating the computing times for the upper and lower bounds of both, RiNSES<sup>4</sup> (blue lines) and BARON (green lines). The black dotted line represents the global optimal solution to the MINLP problem. Additionally, as explained in Section 2, the RiNSES<sup>4</sup> method independently calculates the upper and lower bounds, resulting in varying computing times for each.

In our case study, RiNSES<sup>4</sup> finds the first feasible solution within 150 seconds, which is nearly 20 times faster than BARON (2844 seconds). In addition, the initial feasible solution of RiNSES<sup>4</sup> is of high quality, closely approaching the optimal solution, in contrast to that of BARON. When comparing the upper bounds with the optimal solution, the relative error of the first feasible solution of RiNSES<sup>4</sup> is already within the predefined optimality gap  $\varepsilon^R$  5%.

In terms of the lower bounds, RiNSES<sup>4</sup> calculates the initial lower bound in 1444 seconds, which is twice as fast as BARON (2844 seconds). However, the initial lower bounds of both methods are significantly distant from the optimal solution. While BARON improves its lower bound in 2885 seconds, RiNSES<sup>4</sup> requires considerably more

time, achieving a better lower bound in 6649 seconds.

The solver BARON achieves the predefined solution quality in 6518s, with an optimality gap of 2.59%, whereas RiNSES<sup>4</sup> needs a slightly longer computing time of 6649s to reach an optimality gap of 4.97%. This is due to the slower improvement of lower bounds in the RR branch. All in all, RiNSES<sup>4</sup> can find good feasible solutions to MINLP energy system synthesis problems faster than the state-of-the-art solver BARON, even if the proof of optimality takes longer for the regarded case study.



**Figure 4:** Lower and upper bounds of RiNSES<sup>4</sup> and BARON as function of the computing time for designing a PV-BAT energy system.

### 4. CONCLUSION

The synthesis of energy systems for seasonal energy supply and storage initially results in large-scale MINLP problems, which are computationally challenging and often simplified and reformulated as (MI)LP problems to enhance computational tractability.

This work introduces the RiNSES<sup>4</sup> method, specifically designed to address nonlinearity in energy system synthesis. RiNSES<sup>4</sup> features two separate branches to under- and overestimate the initial MINLP problem simultaneously and independently. The method provides feasible solutions by upper bounds, which involve linearization, aggregation, superposition storage modeling and decomposition. For computing lower bounds, which are crucial for assessing solution quality, two-stage relaxations are utilized.

RiNSES<sup>4</sup> is applied to design a photovoltaic and battery energy system, with the results evaluated in computational studies. In comparison to the commercial solver BARON, the proposed RiNSES<sup>4</sup> method finds the initial optimal solution very quickly, albeit with a higher optimality gap.

The RiNSES<sup>4</sup> method is generally applicable to two-stage, time-dependent synthesis problems with coupling decision variables and constraints in complex energy systems. However, the method's approach for computing lower bounds, which relies on two-stage relaxations,

requires further enhancement. One potential enhancement could involve integrating nonlinear relaxation into the RR branch. This integration could help to effectively underestimate the nonlinear constraints and reduce computing time by solving MILP formulations within the RR branch. Overall, the foundational design and adaptability of RiNSES<sup>4</sup> render it a promising tool for advancing the field of nonlinear synthesis of energy systems for seasonal energy supply and storage.

## DIGITAL SUPPLEMENTARY MATERIAL

The digital supplementary material of this work can be found in a Git repository: <https://git-ce.rwth-aachen.de/ltp/ptg-es4>.

## ACKNOWLEDGEMENTS

This study is funded by the ‘Europäischer Fonds für regionale Entwicklung (EFRE)’ (EFRE-0801844). The support is gratefully acknowledged.

## REFERENCES

- Lin F, Leyffer S, Munson T. A two-level approach to large mixed-integer programs with application to cogeneration in energy-efficient buildings. *Computational Optimization and Applications* 65: 1–46 (2016).
- Goderbauer S, Bahl B, Voll P, Lübbecke ME, Bardow A, Koster AMCA. An adaptive discretization MINLP algorithm for optimal synthesis of decentralized energy supply systems. *Computers & Chemical Engineering* 95: 38–48 (2016).
- Kotzur L, Nolting L, Hoffmann M, Groß T, Smolenko A, Priesmann J, Büsing H, Beer R, Kullmann F, Singh B, Praktijnjo A, Stolten D, Robinius M. A modeler's guide to handle complexity in energy systems optimization. *Advances in Applied Energy* 4: 100063 (2021).
- Kotzur L, Markewitz P, Robinius M, Stolten D. Impact of different time series aggregation methods on optimal energy system design. *Computers & Chemical Engineering* 117: 474–487 (2018).
- Gabrielli P, Gazzani M, Martelli E, Mazzotti M. Optimal design of multi-energy systems with seasonal storage. *Applied Energy* 219: 408–424 (2018).
- Kotzur L, Markewitz P, Robinius M, Stolten D. Time series aggregation for energy system design: Modeling seasonal storage. *Applied Energy* 213: 123–135 (2018).
- Baumgärtner NJ, Temme F, Bahl B, Hennen MR, Hollermann DE, Bardow A. RiSES4: Rigorous Synthesis of Energy Supply Systems with Seasonal Storage by relaxation and time-series aggregation to typical periods. *Proceedings of the International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems (ECOS 2019)* 263–274 (2019).
- Bahl B, Söhler T, Hennen M, Bardow A. Typical periods for two-stage synthesis by time-series aggregation with bounded error in objective function. *Frontiers in Energy Research* 5: 35 (2018).
- Baumgärtner N, Bahl B, Hennen M, Bardow A. RiSES3: Rigorous Synthesis of Energy Supply and Storage Systems via time-series relaxation and aggregation. *Computers & Chemical Engineering* 127: 127–139 (2019).
- Baumgärtner N, Shu D, Bahl B, Hennen M, Hollermann DE, Bardow A. DeLoop: Decomposition-based Long-term operational optimization of energy systems with time-coupling constraints. *Energy* 198: 117272 (2020).
- Vielma JP, Ahmed S, Nemhauser G. Mixed-Integer Models for Nonseparable Piecewise-Linear Optimization: Unifying Framework and Extensions. *Operations Research* 58: 303–315 (2010).
- Langiu M, Shu DY, Baader FJ, Hering D, Bau U, Xhonneux A, Müller D, Bardow A, Mitsos A, Dahmen M. COMANDO: A Next-Generation Open-Source Framework for Energy Systems Optimization. *Computers & Chemical Engineering* 152: 107366 (2021).
- Hoffmann M, Kotzur L, Stolten D. The Pareto-optimal temporal aggregation of energy system models. *Applied Energy* 315: 119029 (2022).
- Baumgärtner N, Delorme R, Hennen M, Bardow A. Design of low-carbon utility systems: Exploiting time-dependent grid emissions for climate-friendly demand-side management. *Applied Energy* 247: 755–765 (2019).
- Wang Y, Bornemann L, Reinert C, von der Aßen N. A Power-to-Gas energy system: modeling and operational optimization for seasonal energy supply and storage. In: *33rd European Symposium on Computer Aided Process Engineering*. Ed: Elsevier (2023).
- Gurobi Optimization LLC. *Gurobi Optimizer Reference Manual*. (2022).
- The Optimization Firm LLC. <https://minlp.com/baron-solver>.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Simulation and Comparative Analysis of Conventional Steam-Methane Reforming Models for Reactor Electrification

Yufei Zhao<sup>a</sup>, Chengtian Cui<sup>a</sup>, and Cornelius. M. Masuku<sup>a\*</sup>

<sup>a</sup> Purdue University, Davidson School of Chemical Engineering, West Lafayette, Indiana, USA

\* Corresponding Author: [cmasuku@purdue.edu](mailto:cmasuku@purdue.edu).

## ABSTRACT

This study delves into the development and examination of various mathematical models for conventional steam-methane reforming (SMR) reactors, establishing a foundational basis for an electrified SMR reactor design. Distinct mathematical models with different scales and dimensions are derived. A basic 1D-fluid, 0D-catalyst (1D-0D) pseudo-homogeneous model is validated with plant data, and progressively advanced to a 2D-0D model considering radial transfer, then further extended to a rigorous 2D-1D model considering transfer phenomena between catalyst particle and fluid. Simulation cases are conducted under uniform design parameters, heat source and operation conditions. Comparative analyses focus on several key performance aspects, including temperature, reaction rate distribution, and outlet characteristics such as temperature, pressure, flow rate, composition and CH<sub>4</sub> conversion. The models effectively describe the industrial SMR reactor behavior. Influences of scale and dimension of mathematical model on reactor performance are highlighted. The rigorous 2D-1D model is identified as the most suitable model for adapting to electrified reactor configurations due to its precise capture of transfer phenomena and detailed illustration of both fluid and catalyst behaviors.

**Keywords:** Steam Methane Reforming, Reactor Design, Multi-Scale Modeling, Decarbonization, Hydrogen Production.

## 1 INTRODUCTION

Hydrogen (H<sub>2</sub>) is a crucial industrial chemical with a reported demand of approximately 90 million tonnes in the year 2020 and a projected increase in demand ranging from 28% to 45% by 2030 [1]. H<sub>2</sub> plays a vital role in various sectors including oil refining, and ammonia production among others [2]. About 62% of H<sub>2</sub> is produced from natural gas especially through steam-methane reforming (SMR) [1]. However, conventional SMR has disadvantages including: consumption of fossil fuels, emissions of carbon dioxide, heat transfer limitation of tube-furnace design. Electrification through induction heating as a substitute of heat source presents a noteworthy alternative [3]. To numerically assess an electrically-heated SMR (E-SMR), an electromagnetic model for induction heating is required. Concurrently, a comprehensive SMR model that can correlate the heat generation at

catalyst sites with the overall behavior of the reactor is essential.

There has been abundant research about modeling of SMR reactor. Latham *et al.* [4] developed a mathematical model of the reformer by segments which were discretized axially. Constant effectiveness factors were implemented with intrinsic kinetic model for reaction rate calculation. The model can predict temperature profiles for the outer-tube wall, inner-tube wall, furnace gas and process gas. Kuncharam, B. V. R *et al.* [5] employed a multi-scale modeling approach, integrating a steady state two-dimensional model for fluid-phase with a one-dimensional model for the pellet-phase. The effectiveness factor assumption was avoided. Their findings assert that the application of the Ergun equation to calculate the pressure drop in SMR reformer yields satisfactory results. Furthermore, the incorporation of molar change is necessary. Tacchino *et al.* [6] presented a validated

steady state multi-scale model of an industrial SMR reactor comprised of the furnace and reformer. A significant variability in effectiveness factor was observed, indicating the necessity of including catalyst particle simulation. The model proved its robustness by showing no sensitive reliance on adjustable parameter, the flame length.

Current advances in SMR modeling reveal a notable discrepancy in the approaches across different scales and dimensions. Models have been developed based on distinct parameters, operating conditions, and empirical expressions. Key factors such as fluid mechanics, heat input, temperature and reaction rate distributions, showing significant influence on the overall behavior including outlet condition, CH<sub>4</sub> conversion and efficiency of the reformer, have dependencies on scale and dimensionality. There lacks a comprehensive comparative analysis rooted in a uniform basis. In order to lay a foundation for E-SMR model development, and to quantify the key factors and the influence of varied-scale and varied-dimension on reactor behavior, numerical simulation of different models and analysis are required [7].

In this paper, for fluid-catalyst dimensions, 1D-0D, 2D-0D, 2D-1D mathematical models of conventional SMR are developed based on the same design parameters and operation conditions. Temperature and reaction rate distributions corresponding to different scenarios are illustrated. SMR reactor properties are recognized. Outlet behaviors, including temperature, pressure, flow rate, composition and CH<sub>4</sub> conversion are observed. Sensitivity analysis is conducted for the radial conductivity. Differences of the modeling schemes are recognized and analyzed.

## 2 MODEL DEVELOPMENT

**Table 1:** Characteristics of reactor tube and catalyst particles.

Parameters	Value	Unit
Length of tube	12.5	m
Radius of tube	0.063	m
Diameter of catalyst	0.0035	m
Conductivity of catalyst	33	W/m/K
Density of catalyst	3690	kg/m <sup>3</sup>
Heat capacity of catalyst	880	J/kg/K

Industrial SMR reactors contain parallel tubes arranged in rows and columns inside the furnace, each packed with catalyst. This paper sets up a mathematical model of an industrial-scale single tube to observe the overall behavior of the reactor. Ni/MgAl<sub>2</sub>O<sub>4</sub> catalyzes the reaction. Characteristics of reactor tube and catalyst particles are listed in Table.1.

The following assumptions have been made for establishing the models:

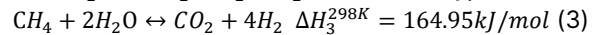
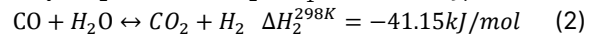
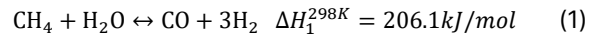
## 2.1 Assumptions

1. Operational conditions and performance data can be linearly extrapolated from a single tube to the overall reactor;

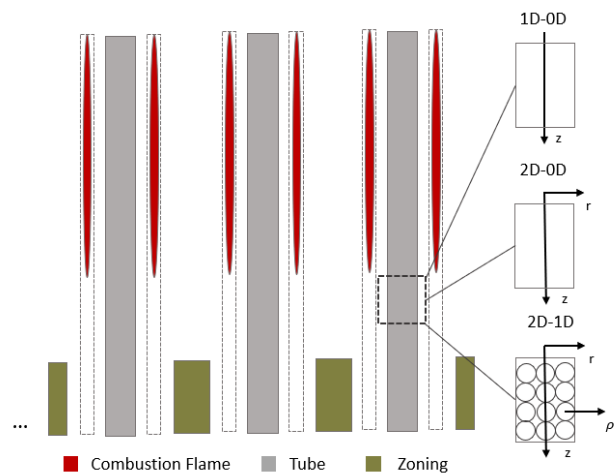
2. The fluid phase follows ideal gas behaviors;

3. Spherical catalyst particle and cylindrical tube have constant porosity respectively;

4. 3 main reactions occur:



5. Inner tube wall temperature is a polynomial function of axial space position fitted from industrial data.



**Figure 1.** Abbreviated diagram of industrial SMR reactor (reproduced from Ref. [4]) and illustration of three models regarding scale and dimension.

Three models covering different scales and dimensions are shown in Figure 1, along with industrial SMR reactor diagram. Tube expressions are derived in cylindrical coordinate, while catalyst particle expressions in spherical coordinate. Considering geometric symmetry, the most rigorous scenario is a multi-scale model considering multi-dimension including axial and radial domains for the fluid phase in tube, and radial domain for catalyst particle (2D-1D). Additional assumptions may be included owing to the distinct characteristics in each model. In the 2D-0D model, reactor tube is assumed to be pseudo-homogeneous simplifying the heterogeneity. For 1D-0D model, plug flow behavior is additionally assumed.



**Table 2:** Nomenclature.

Symbol	Unit	Name
C	mol/m <sup>3</sup>	concentration
u	m/s	Interstitial velocity
U	m/s	Superficial velocity
D	m <sup>2</sup> /s	diffusivity
y		Composition
f	Hz	Frequency of alternating magnetic field
a <sub>p</sub>	m <sup>2</sup> /m <sup>3</sup>	surface area per unit volume of catalyst
d <sub>p</sub>	m	diameter of catalyst
R <sub>tube</sub>	m	Radius of reactor tube
k <sub>m</sub>	m/s	solid-fluid mass transfer coefficient
k <sub>h</sub>	W/m <sup>2</sup> /K	solid-fluid heat transfer coefficient
ρ <sub>f</sub>	kg/m <sup>3</sup>	density of fluid
ρ <sub>p</sub>	kg/m <sup>3</sup>	density of catalyst particle
c <sub>p</sub>	J/kg/K	heat capacity
T	K	temperature
λ	W/m/K	conductivity
R	mol/kgcat/s	reaction rate
L	m	length of reactor tube
ΔH	J/mol	reaction enthalpy
ε		porosity
v		stoichiometry of reactions
h	W/m <sup>2</sup> /K	Tube-fluid heat transfer coefficient
A <sub>hys</sub>	J/kg	Hysteresis area
Pr	μc <sub>p</sub> /λ	Prandtl number
Pe	ud/D <sub>i</sub>	Peclet number
Re	ρud/μ	Reynolds number
Sh	k <sub>m</sub> d/D <sub>i</sub>	Sherwood number
Sc	μ/D <sub>i</sub> ρ	Schmidt number
<b>Subscript</b>		
i,j		component or reaction index
z		fluid axial domain
r		fluid radial domain
ρ		catalyst radial domain
s		catalyst surface
b		catalyst bed
f		Fluid phase
p		Catalyst phase
<b>Superscript</b>		
e		effective coefficient
eo		quiescent bed effective coefficient

## 2.2 Mathematical Expressions

The mathematical models encompass mass, heat

and momentum conservations under steady state for fluid mechanics description, and kinetic model for reaction description. Mass diffusion and convection, and heat conduction and convection, are considered for transfers. Pressure drop expression is derived from momentum conservation with friction factor calculated by the Ergun equation.

For the purpose of elucidating the expressions, the rigorous 2D-1D model is shown below as a representative example since it is the most detailed.

### 2.2.3 Fluid Phase

#### 2.2.3.1 Mass Conservation

$$\frac{\partial C_i}{\partial z} u_z + C_i \frac{\partial u_z}{\partial z} = \frac{1}{r} D_{i,r}^e \frac{\partial C_i}{\partial r} + \frac{\partial D_{i,r}^e}{\partial r} \frac{\partial C_i}{\partial r} + D_{i,r}^e \frac{\partial^2 C_i}{\partial r^2} + \frac{\partial D_{i,z}^e}{\partial z} \frac{\partial C_i}{\partial z} + D_{i,z}^e \frac{\partial^2 C_i}{\partial z^2} + \frac{(1-\varepsilon_b)}{\varepsilon_b} a_p k_{i,m} (C_{i,s} - C_i) \quad (1)$$

Boundary conditions:

$$C_i|_{z=0} = c_{i0}; \quad \frac{\partial C_i}{\partial z}|_{z=L} = 0; \\ \frac{\partial C_i}{\partial r}|_{r=0} = 0; \quad \frac{\partial C_i}{\partial r}|_{r=R_{tube}} = 0 \quad (2)$$

#### 2.2.3.2 Heat Conservation

$$\varepsilon_b \rho_f u_z c_p \frac{\partial T}{\partial z} = \frac{1}{r} \lambda_r^e \frac{\partial T}{\partial r} + \frac{\partial \lambda_r^e}{\partial r} \frac{\partial T}{\partial r} + \lambda_r^e \frac{\partial^2 T}{\partial r^2} + \frac{\partial \lambda_z^e}{\partial z} \frac{\partial T}{\partial z} + \lambda_z^e \frac{\partial^2 T}{\partial z^2} + (1 - \varepsilon_b) a_p k_h (T_s - T) \quad (3)$$

Boundary conditions:

$$-\lambda_r^e \frac{\partial T}{\partial r}|_{r=R_{tube}} = h(T_w - T); \quad \frac{\partial T}{\partial r}|_{r=0} = 0; \\ T|_{z=0} = T_0; \quad \frac{\partial T}{\partial z}|_{z=L} = 0 \quad (4)$$

### 2.2.4 Catalyst Phase

#### 2.2.4.1 Mass Conservation

$$0 = \frac{2}{\rho} D_{i,p,\rho} \frac{\partial C_{i,p}}{\partial \rho} + \frac{\partial D_{i,p,\rho}}{\partial \rho} \frac{\partial C_{i,p}}{\partial \rho} + D_{i,p,\rho} \frac{\partial^2 C_{i,p}}{\partial \rho^2} + \frac{1}{\varepsilon_p} \rho_p \sum_{j=1}^3 \nu_{ij} R_j \quad (5)$$

Boundary conditions:

$$\frac{\partial C_{i,p}}{\partial \rho}|_{\rho=0} = 0; \quad -\varepsilon_p D_{i,p,\rho} \frac{\partial C_{i,p}}{\partial \rho}|_{\rho=R_p} = k_{i,m} (C_{i,s} - C_i) \quad (6)$$

#### 2.2.4.2 Heat Conservation

$$[(1 - \varepsilon_p) \rho_p c_{p,p} + \varepsilon_p \rho_f c_{p,f}] \frac{\partial T_p}{\partial t} = \frac{2}{\rho} \lambda_{p,\rho}^e \frac{\partial T_p}{\partial \rho} + \frac{\partial \lambda_{p,\rho}^e}{\partial \rho} \frac{\partial T_p}{\partial \rho} + \lambda_{p,\rho}^e \frac{\partial^2 T_p}{\partial \rho^2} - (1 - \varepsilon_p) \rho_p \sum_{j=1}^3 R_j \Delta H_j \quad (7)$$

Boundary conditions:

$$\frac{\partial T_p}{\partial \rho}|_{\rho=0} = 0; \quad -\lambda_{p,\rho}^e \frac{\partial T_p}{\partial \rho}|_{\rho=R_p} = k_h (T_s - T) \quad (8)$$

### 2.2.5 Coefficients

### 2.2.5.1 Diffusivity

Axial and radial effective diffusivity are expressed as following:

$$\frac{1}{Pe_z} = \frac{0.72}{Re_p Sc} + \frac{0.52}{\left(1 + \frac{9.0}{Re_p Sc}\right)} \quad (9)$$

$$\frac{1}{Pe_r} = \frac{0.34}{(Re_o Sc)^{0.80}} + \frac{0.08}{\left(1 + \frac{10.8}{Re_o Sc}\right)} \quad (10)$$

### 2.2.5.2 Conductivity

$$\frac{\Delta_f^e}{\lambda_f} = \frac{\lambda^{eo}}{\lambda_f} + 0.1PrRe \quad (11)$$

$$\frac{\lambda_z^e}{\lambda_f} = \frac{\lambda^{eo}}{\lambda_f} + 0.5PrRe \quad (12)$$

### 2.2.5.3 Solid-fluid transfer coefficient

$$Sh = \frac{k_{i,m} d_p}{D_{i,m}} = 2 + 1.1Sc^{1/3} Re_o^{0.6} \quad (13)$$

$$Nu = \frac{k_h d_p}{\lambda_m} = 2 + 1.1Pr^{1/3} Re_o^{0.6} \quad (14)$$

### 2.2.5.4 Tube-fluid heat transfer coefficient

$$H = \alpha \frac{\lambda}{d_p} \left( 2.58 Re_p^{1/3} Pr^{1/3} + 0.094 Re^{0.8} Pr^{0.4} \right) \quad (15)$$

### 2.2.6 Kinetics

An intrinsic Langmuir-Hinshelwood-Hougen-Watson kinetic model developed by Xu and Froment [8] is applied.

## 3 RESULTS

The equation-based mathematical models composed of PDEs are coded in FORTRAN and solved in Aspen Custom Modeler V12.1. Physical properties are calculated by calling Aspen properties bank. Finite difference methods are used for discretization. Newton Method solves the system of nonlinear equations.

### 3.1 1D-0D Model Validation

This model describes homogeneous reactor tube with only axial domain transfer. The simulation has 9,887 equations after decomposition. Operating conditions of the three cases implemented are listed in Table 3.

The model validation focuses on temperature profile along the tube, CH<sub>4</sub> conversion of methane and outlet pressure. Root mean standard error (RMSE) and Pearson correlation Coefficient (PCC) defined in Equation 16 are calculated to measure deviation and correspondence of change tendency of profile. The smaller RMSE is, and the closer CF is to 1, the better consistency the simulated profile shows to industrial case. Relative error (RE) is calculated for evaluating CH<sub>4</sub> conversion and pressure. Figure 2 shows comparative results between temperature profile derived from our model and those documented in

literature under Case 3 operation conditions as an example. Modeled composition profiles are plotted in Figure 3, along with industrial outlet values. Table 4 shows the evaluation coefficients quantifying the comparison of temperature profile. The developed 1D-0D model has lowest PCC of 0.9979, highest RE of 3.61%, highest RMSE of 14.28 compared with 800-1150K range, showing a consistency with literature reported. The outlet behaviors are listed in Table 5. The results are aligned with plant data.

$$CF = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-\bar{y})^2}} \quad (16)$$

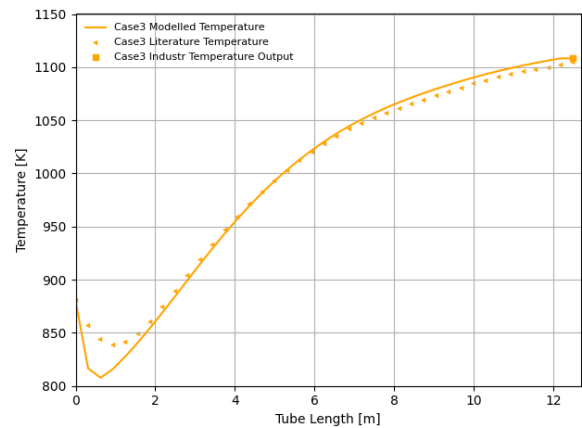
Heat is input into a discretized plug flow segment for 1D-0D model. The profile is plotted in Figure 4 along with reaction rate. The maximum heat input is reached at about L=3-5m, which is the length of a combustion flame. The majority of reactions happen near the entrance of the tube length area, leading to a temperature drop to minimum at about L=0.625m.

**Table 3:** Inlet operation conditions of 3 industrial cases [4].

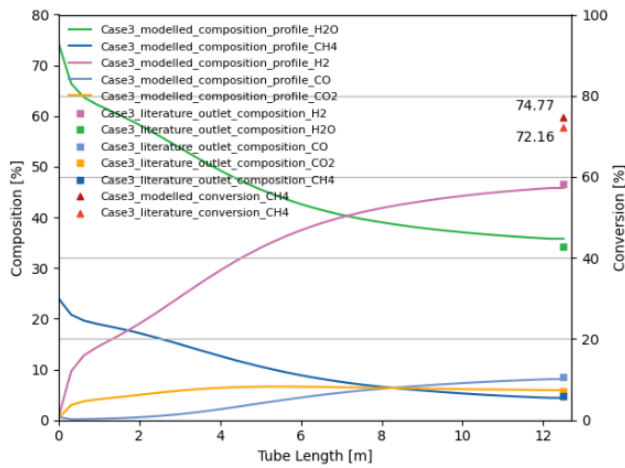
	Case1	Case2	Case3
T [K]	884.55	887.05	879.85
P [bar]	30.060	29.440	28.085
Velocity [m/s]	1.279	1.147	0.925
y <sub>CH4</sub>	0.2421	0.2487	0.2401
y <sub>H2O</sub>	0.7461	0.7377	0.7437
y <sub>CO</sub>	0.0036	0.0042	0.0060
y <sub>CO2</sub>	0.0047	0.0053	0.0042
y <sub>H2</sub>	0.0036	0.0042	0.0060

**Table 4:** Evaluation coefficients of 3 cases compared with literature documented [4].

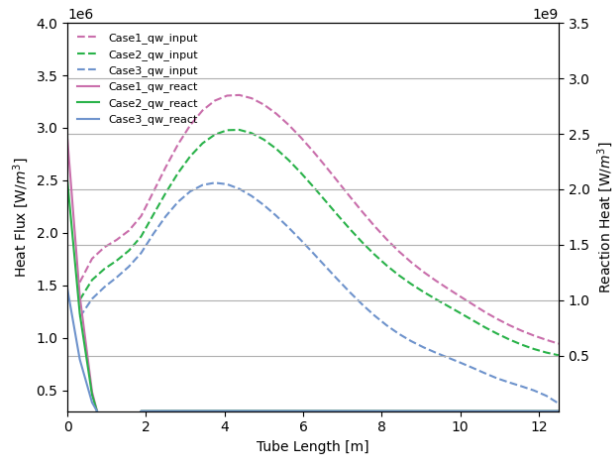
	Case1	Case2	Case3
RMSE	14.28	12.58	10.46
PCC	0.9979	0.9981	0.9979



**Figure 2.** Simulated (solid), literature (dashed) temperature profile along the tube of Case 3 with RMSE of 10.4.



**Figure 3.** Simulated industrial composition profile along the tube, and CH<sub>4</sub> conversion of case 3.



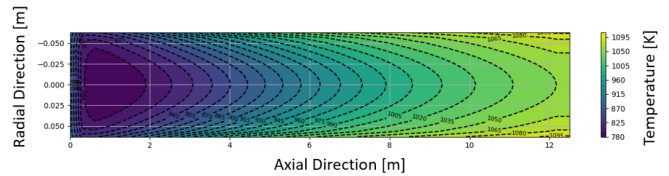
**Figure 4.** Simulated reaction heat (solid) and input heat (dashed) profile along the tube.

### 3.2 2D-0D Model

Radial transfer is introduced in 2D-0D homogenous model based on 1D-0D model. The simulation has 43,035

equations after decomposition. Case 3 is implemented for comparison.

Figure 5 shows simulated temperature distribution. Axially lowest temperature appeared at  $L=0.3125-0.625m$ , which is at similar position compared with 1D-0D. Radial temperature difference is observed, with a highest of 126.31K at  $L=4.0625m$ , which is slightly delayed compared with the maximum heat input at  $L=3.75m$ . Radial temperature difference follows the same trend as heat input as shown in Figure 6 (top). Radial temperature difference indicating heat loss is inevitable considering the relatively low radial heat transfer efficiency. However, the average temperature along axial direction still shares a similar trend compared with 1D-0D. 2D-0D case has a lower average temperature of 49.4K, leading to a CH<sub>4</sub> conversion decrease from 75% to 64%.



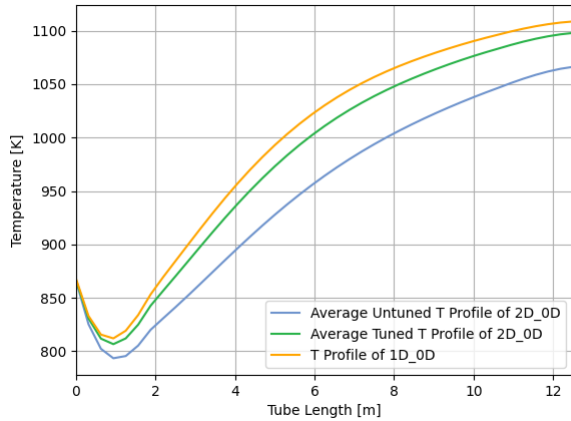
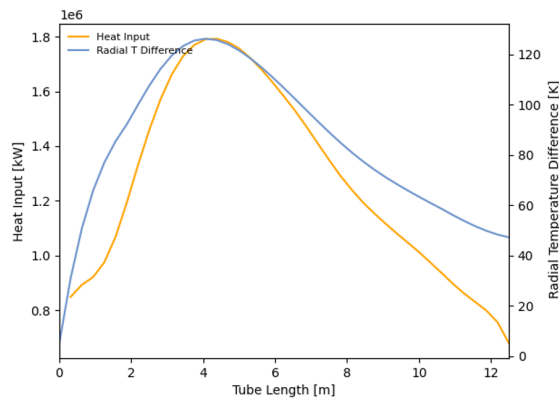
**Figure 5.** Simulated temperature distribution in 2D-0D case.

Figure 7 illustrates the distribution of the reaction rate, using Reaction 3 as a representative case. 99.5% of the overall reaction occurs with the length  $L=0.9375m$ . For the rest of the tube length, 69%-83% of the reaction occurs near the reactor tube wall where temperature is highest radially, indicating the reactor being heat transfer limited.

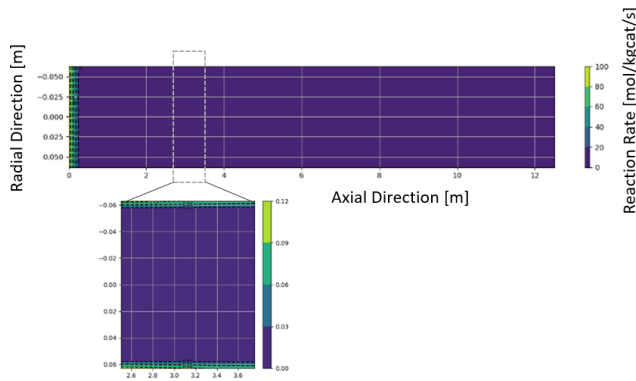
The different behaviors between 2D-0D and 1D-0D are resulted from with or without plug-flow assumption. Radial transfer coefficient is assumed to be infinitely large to reach uniform distribution. This tendency is illustrated in Figure 8 that the larger the radial conductivity is, the closer the average outlet temperature is to 1D-0D model.

**Table 5:** Modelled outlet behaviors compared with industrial data [4].

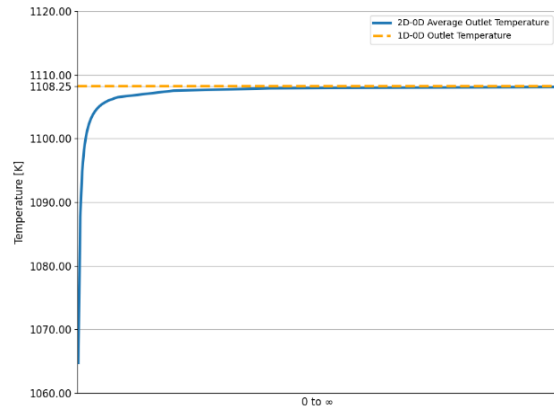
	Case 1	Plant 1	Case 2	Plant 2	Case 3	Plant 3
Temperature [K]	1107.51	1105.55	1108.77	1107.15	1108.25	1108.95
Pressure [bar]	28.461	28.189	28.040	27.990	27.234	27.155
Flow Rate [mol/s]	8.935	8.970	7.797	8.082	6.136	6.024
CH <sub>4</sub> Conversion [%]	73.79	71.30	73.11	70.70	74.77	72.16
y <sub>CH4</sub>	0.0461	0.0505	0.0484	0.0526	0.0482	0.0440
y <sub>H2O</sub>	0.3614	0.3507	0.3517	0.3417	0.3430	0.3580
y <sub>H2</sub>	0.4544	0.4583	0.4593	0.4631	0.4655	0.4580
y <sub>CO</sub>	0.0790	0.0820	0.0820	0.0845	0.0850	0.0808
y <sub>CO2</sub>	0.0591	0.0582	0.0587	0.0575	0.0574	0.0592



**Figure 6.** (top) Heat input profile (left) and radial temperature difference (right) profile along the reactor tube; (bottom) 1D-0D temperature profile, untuned and tuned 2D-0D average temperature profile.



**Figure 7.** Reaction rate distribution in reactor tube with as a function of a specific length.



**Figure 8.** 2D-0D average outlet temperature versus radial conductivity.

To ensure that the 2D model has better accordance with industrial data regarding the  $\text{CH}_4$  conversion, the amount of heat input is recalculated by parameter tuning of the tube-fluid heat transfer coefficient  $\alpha$ . The tuned-conversion reaches less than  $5e-4\%$  RE compared with industrial data. Heat input after tuning has an average of  $1.7e5 \text{ W/m}^3$  more in amount compared with untuned case. Outlet average temperature has 1.03% RE, while pressure has 0.15% RE. The tuned temperature distribution has an average temperature of 14.3K lower than 1D-0D due to the radial temperature distribution, but 35.1K higher than untuned 2D-0D shown in Figure 6 (bottom). The outlet behaviors are listed in Table 6. The tuned case approaches the plant data more than untuned case.

**Table 6.** Comparison of tuned, untuned 2D-0D cases and industrial case.

Outlet	Tuned	Untuned	Plant
Temperature [K]	1096.86	1064.77	1108.95
Pressure [bar]	27.19	27.25	27.23
Flow Rate [mol/s]	6.133	6.015	6.024
$\text{CH}_4$ Conversion [%]	72.16	64.32	72.16
$y_{\text{CH}_4}$	0.0496	0.0560	0.0482
$y_{\text{H}_2\text{O}}$	0.3661	0.3756	0.3430
$y_{\text{H}_2}$	0.4480	0.4364	0.4655
$y_{\text{CO}_2}$	0.0605	0.0619	0.0573
$y_{\text{CO}}$	0.0757	0.0702	0.0850

### 3.3 2D-1D Model

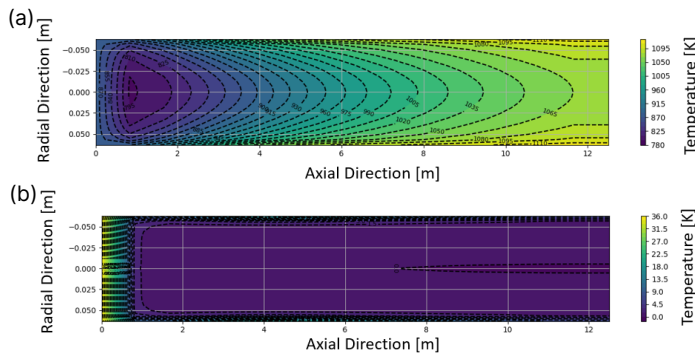
Solid catalyst particle phase is taken into consideration for 2D-1D heterogenous model. External transfer between catalyst surface and fluid phase, and internal transfer in catalyst particle are introduced. The simulation has 205,389 equations after decomposition. Case 3 is implemented for comparison.

Temperature difference exists between the fluid, the catalyst surface, and catalyst center, leading to

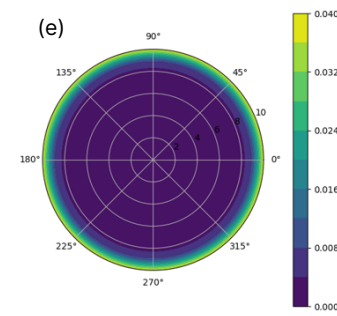
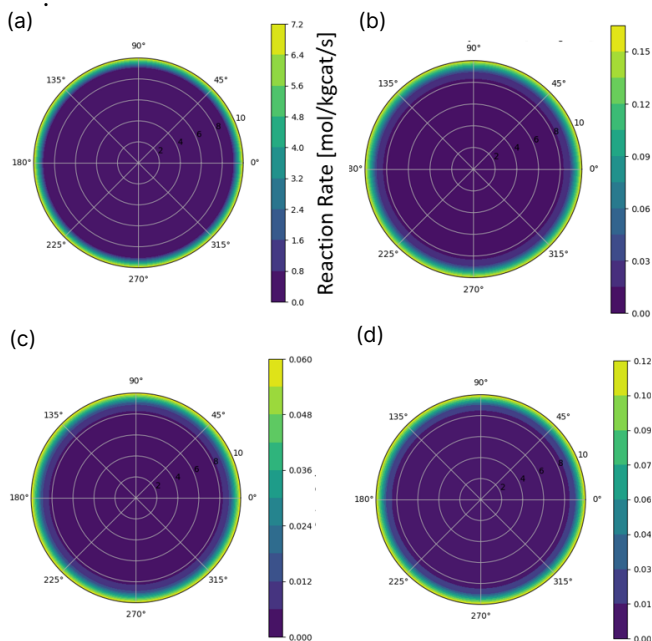
additional heat loss. Figure 9 (a) shows the temperature distribution of fluid. Temperature differences from catalyst center to catalyst surface is less than 0.5 K. Temperature difference from catalyst surface to fluid distribution is shown in Figure 9 (b), the average is 3.69 K. The maximum difference axially is at entrance, while radially is at tube wall, where the majority of reactions occur, which is consistent with 2D-0D observation. A drop in CH<sub>4</sub> conversion of 24% is observed under the same condition as 2D-0D.

Reaction rate distributions of catalyst particle at specific positions are shown in Figure 10. The majority of reactions occur near the surface, which indicates that the catalytic reactions are diffusion-limited.

The outlet behaviors are shown in Table 7. The 2D-1D outlet behaviors show good consistency with 2D-0D case, except temperature and CH<sub>4</sub> conversion.



**Figure 9.** (a) Temperature distribution of 2D-1D fluid phase; (b) Temperature difference between fluid and catalyst surface distribution of 2D-1D.



**Figure 10.** Reaction rate distribution of catalyst particle at position (a) L/L<sub>tube</sub>=0; (b) L/L<sub>tube</sub>=0.20; (c) L/L<sub>tube</sub>=0.47; (d) L/L<sub>tube</sub>=0.73; (e) L/L<sub>tube</sub>=1.

**Table 7.** Comparison of 2D-1D with tuned 2D-0D cases.

Outlet	2D-1D	2D-0D
Temperature [K]	1080.09	1096.86
Pressure [bar]	27.19	27.19
Flow Rate [mol/s]	6.083	6.133
CH <sub>4</sub> Conversion [%]	47.99	72.16
Y <sub>CH4</sub>	0.0498	0.0496
Y <sub>H2O</sub>	0.3665	0.3661
Y <sub>H2</sub>	0.4481	0.4480
Y <sub>CO2</sub>	0.0606	0.0605
Y <sub>CO</sub>	0.0750	0.0757

### 3.4 E-SMR Model

The development of conventional SMR model lays foundation for heat electrification that E-SMR model is developed by incorporation of induction heat model. Hysteresis heat is regarded as dominant induction heat source, which is estimated by the following expression:

$$P_{hys} = A_{hys} f \rho \quad (17)$$

where hysteresis area  $A_{hys}$  is dependent on material electromagnetic property, magnetic field strength and temperature.

In conventional SMR reactors, the heat flux source is introduced as a boundary condition for the fluid phase boundary at tube layer. While for induction-heated E-SMR reactor which utilizes dual-function catalysts, heat is directly generated at the catalyst sites. The heat flux is introduced internally from the catalyst phase in the way that the  $P_{hys}$  term is included as a source term in the energy conservation equation of the catalyst/fluid phase.

## 4 CONCLUSIONS

Mathematical models covering varied scales and dimensions of SMR Reactor tube have been developed. Comparative analyses have been carried out focusing on



temperature, reaction rate distribution and outlet behaviors, including temperature, pressure, flow rate, composition and CH<sub>4</sub> conversion.

Validation with industrial data shows that the 1D-0D numerical simulation can precisely describe the SMR reactor behavior. When developed on the same basis, the 2D-0D shows a radial temperature distribution. When resistance of radial transfer is set to infinitely small, the behavior of 2D-0D is approaching 1D-0D, which is a validation of the developed 2D-0D model. Higher reaction rates are mainly distributed near entrance and near tube wall positions, indicating a heat transfer limited property. 2D-0D shows a lower CH<sub>4</sub> conversion than 1D-0D considering the heat resistance behavior in the way of radial temperature distribution. A heat transfer coefficient corresponding to industrial plant CH<sub>4</sub> conversion can be obtained by parameter fitting. The fitted 2D-0D case shows a high correlation with plant data. 2D-1D model was further developed with the results showing additional heat resistance in the way of temperature difference between the fluid, catalyst surface, and catalyst center. A further decrease of CH<sub>4</sub> conversion compared with 2D-0D is observed. On catalyst sites, the majority of reactions occur at position close to the catalyst surface, indicating the diffusion-limited property.

Simplification of model in terms of scale and dimension neglects some of the heat resistances. Due to the fixed tube wall assumption for all three cases, and same initial conditions, the neglected heat resistance manifests in the way of decreased CH<sub>4</sub> conversion. However, distinct models are suitable for different tasks. When an estimation of effectiveness factor is unavailable and catalyst behaviors are required, rigorous 2D-1D is more suitable.

## ACKNOWLEDGEMENTS

The authors would like to thank the Center for Innovative and Strategic Transformation of Alkane Resources (CISTAR) and the National Science Foundation under Cooperative Agreement No. EEC-1647722 for funding.

## REFERENCES

1. International Energy Agency. Global Hydrogen Review 2022. (2022)
2. Department of Energy.  
<https://www.energy.gov/eere/fuelcells/h2scale>
3. Wismann ST, et al. Electrified methane reforming: A compact approach to greener industrial hydrogen production. *Science* 364(6442): 756-759 (2019).
4. Latham DA. Mathematical Modelling of an Industrial Steam-Methane Reformer. PhD Thesis, Queen's University, Kingston, ON (2008).
5. Kuncharam BVR, Dixon AG. Multi-scale two-dimensional packed bed reactor model for industrial steam methane reforming. *Fuel Process Technol* 200 :106314 (2020).
6. Tacchino V, et al. Multi-scale model of a top-fired steam methane reforming reactor and validation with industrial experimental data. *Chem Eng J* 428: 131492 (2022).
7. Masuku CM, et al. Process decarbonization through electrification. *Curr Opin Chem Eng* 44: 101011 (2024).
8. Xu J, Froment GF. Methane steam reforming: II. diffusional limitations and reactor simulation. *AIChE J* 35: 97-103 (1989).

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Part 1

## Peer-Reviewed Articles

Section 5: Design and Sustainability



# Machine Learning Methods for the Forecasting of Environmental Impacts in Early-stage Process Design

Emmanuel A. Aboagye<sup>a</sup>, Austin L. Lehr<sup>a</sup>, Ethan Shumaker<sup>a</sup>, Jared Longo<sup>a</sup>, John Pazik<sup>a</sup>, Robert P. Hesketh<sup>a</sup> and Kirti M. Yenkie<sup>a\*</sup>

<sup>a</sup> Rowan University, Department of Chemical Engineering, Glassboro, NJ, USA

\* Corresponding Author: [yenkie@rowan.edu](mailto:yenkie@rowan.edu).

## ABSTRACT

Initial design stages are inherently complex and often lack comprehensive information, posing challenges in evaluating sustainability metrics. Machine Learning (ML) emerges as a valuable solution to address these challenges. ML algorithms, particularly effective in predicting environmental impacts of new chemicals with limited data, enable more informed decisions in sustainable design. This study focuses on employing ML for predicting the environmental impacts related to human health, ecosystem quality, climate change, and resource utilization to aid in early-stage environmental impact assessment of chemical processes. The effectiveness of the ML algorithm, eXtreme Gradient Boosting (XGBoost) tested using a dataset of 350 points, divided into training, testing, and validation sets. The study also includes a practical application of the model in a cradle-to-cradle LCA of N-Methylpyrrolidone (NMP), demonstrating its utility in sustainable chemical process design. This approach signifies a significant advancement in the early stages of process design, highlighting the potential of ML in enhancing environmental sustainability in the chemical industry.

**Keywords:** Machine Learning, Life Cycle Analysis, Process Design, Modelling, Process Synthesis

## INTRODUCTION

Amidst growing climate change concerns and heightened environmental awareness, industries are increasingly scrutinized for their environmental impact [1]. This scenario underscores the importance of environmental impact assessment at early-stage process synthesis [2], [3], where operational processes are initially formulated and assessed. Decisions made at this stage have significant implications for the environmental footprint of the entire operation. In this context, Machine Learning (ML) [4], [5] presents a transformative solution. By integrating ML in the early stages, industries can efficiently utilize its potential for rapid, accurate, and comprehensive assessments of sustainability.

ML offers a significant advantage in systems with non-obvious relationships in that, once trained, ML models can be used to predict such environmental metrics, facilitating prompt design modifications and improvements. Its adaptability also allows for easy integration with optimization strategies, aiding industries in

developing processes during early-stages that balance environmental and economic factors. Financially, ML is invaluable; addressing sustainability issues early on helps industries avoid costly later-stage modifications, leading to substantial cost savings. In essence, incorporating ML into early-stage process synthesis represents a forward-thinking move for industries striving towards sustainability.

Thermodynamic properties such as enthalpy, entropy, Gibbs free energy [6], [7], among others provide crucial insights into the energy requirements of a process, operational efficiency, and overall feasibility. These attributes significantly influence the energy consumption of the process, impacting essential sustainability metrics such as GWP and total carbon footprint. On the other hand, molecular characteristics [8], [9], including molecular weight, bond energies, and functional groups, among others offer valuable information about the inherent qualities of chemical substances such as reactivity [10], potential toxicity [11], and environmental impacts [12], [13]. Often times, data on both thermodynamic and molecular

properties are available during the initial stages of process design. Therefore, by developing a ML model that incorporates thermodynamic and molecular descriptors as input and sustainability metrics as outputs, it is possible to predict sustainability metrics for both new and existing chemicals that lack established sustainability data. This approach enables a more holistic and informed evaluation of sustainability at the early stages of chemical process synthesis.

Previous research has shown that ML can be utilized to effectively enhance energy efficiency [14] and forecast corporate carbon footprints [15], among others [16], [17]. Building on this, the current study applies ML, specifically, eXtreme Gradient Boosting (XGBoost) [18], [19], to predict endpoint impact metrics for chemicals, particularly new molecules. This approach can lead to the development of safer and more sustainable chemical alternatives and circular process designs. Subsequently, the developed ML model is used together with other methods to predict the entire cradle-to-cradle environmental impact of NMP.

## METHODOLOGY

In this section, we discuss the data gathering process, preprocessing, ML model building, and evaluation metrics.

### Data Acquisition

A comprehensive dataset of 350 common solvents, including alcohols, esters, hydrocarbons, and ethers, was compiled for this study. The dataset is divided into two parts: the feature set and the label set. The feature set consists of thermodynamic and molecular descriptor data, encompassing the chemical properties used for model training. The label set, on the other hand, represents the data that the model aims to predict. For each chemical, 15 thermodynamic properties are gathered, such as critical temperature, pressure, volume, heat capacity, boiling point, and standard Gibbs-free energy. The data collection process begins with extracting the SMILES string [20] and chemical formula for each solvent.

These SMILES strings are used to extract the corresponding thermodynamic properties. This is achieved through two Python libraries: "chemicals" [21] and "thermo", which host extensive databases of pure and calculated chemical properties. The local databank in these libraries includes over 20,000 chemicals and their properties, compiled from sources like the National Institute of Standards and Technology (NIST), Design Institute for Physical Properties (DIPPR), PubChem, CRC Handbook, Perry's Chemical Engineers' Handbook, and various scientific papers and publications.

The molecular descriptor properties for each

chemical were obtained using RDKit [22] (version 2023.3.3), an open-source Python library renowned in cheminformatics. For this study, RDKit was employed to acquire 200 molecular descriptors for each chemical, covering various properties. These include molecular weight, carbon count, maximum partial charge, functional group, number of heterogeneous atoms, number of radical atoms, and the number of aliphatic rings, among others.

For the label data, SimaPro® [23] (version 9.4.0.2) is used to gather the cradle-to-gate metrics for each chemical. The metrics include human health impact (HHI), ecosystem quality impact (EQI), global warming potential (GWP), and resource utilization impact (RUI). These four endpoint metrics are chosen due to decision-making relevance, ease of communication, and depth of analysis.

### Data Preprocessing

The initial step is to address the issue of missing data in the label set. While removing rows with missing data is a typical solution, ML models benefit from larger datasets. Therefore, the k-Nearest Neighbors (kNN) method, a well-established technique in data imputation, was employed for this analysis. Upon completion, the feature dataset was scaled to be in a range of 0 and 5.

Given the large number of features available in the dataset, it was necessary to identify and select those features that contribute most significantly to the model. This process of feature selection not only reduces computational time but also eliminates redundant features, thereby enhancing model accuracy. A balanced approach was adopted, choosing a total of 10 features with an equal number (5) from both the thermodynamic and molecular feature sets. This equal representation ensures that each feature set contributes fairly to the model. The streamlined selection of just 10 features also adds practical value for users, simplifying the prediction process for specific chemicals by requiring only a limited set of properties. Additionally, the model was tailored to each of the four metrics it predicts, with a unique set of features for each metric. This customization ensures that only the most relevant and impactful features are used for predictions in each specific case, optimizing the effectiveness and precision of the model. The Sequential Backward Feature Selection (SBFS) methodology, with linear regression and Mean-Squared-Error (MSE) criterion, was used to achieve this aim.

### Model Training and Hyperparameter Tuning

Once the feature set for each label is finalized, the next step is to build the ML model for the prediction. Two models were developed, XGBoost and Artificial Neural Network (ANN), but in this paper, we present the XGBoost model.

XGBoost, an advanced ensemble ML model, is an

efficient implementation of the gradient boosting framework, particularly suitable for optimizing large-scale ML problems. It functions by iteratively building and refining models, each new model correcting the inaccuracies of its predecessors. This refinement is guided by the gradient descent method, which addresses the weaknesses in the existing ensemble by adding new decision trees, continuing until a predetermined error limit is reached or a specified number of trees is included.

In this study, the data was divided into training, validation, and testing sets. This split was not fixed but was determined based on the label being predicted. To enhance the performance of the model, key hyperparameters of XGBoost (version 1.7.6) were optimized. This optimization involved selecting and tuning four to six hyperparameters that most significantly affect the model. This process was facilitated by the "hyperopt" [24] (version 0.2.7) library, which employs a Bayesian optimization framework [25]. The hyperparameters adjusted include the maximum depth of a tree, learning rate, number of trees, minimum child weight in a node, subsample fraction for growing trees, and the fraction of features chosen for tree development. The optimal hyperparameters were determined using the validation set, with an objective function designed to minimize the MSE between the actual and predicted values after training on the training set. The test set, crucially, was reserved exclusively for evaluating the generalizability and overall performance of the model.

For model evaluation, two key metrics were used: R-squared ( $R^2$ ) value and the Root-Mean-Squared-Error (RMSE). The  $R^2$ , also known as the coefficient of determination, indicates the proportion of the variance in the dependent variable that the independent variables in the model can explain. On the other hand, the RMSE measures the average magnitude of the errors between the predicted and actual outcomes, providing a direct assessment of the accuracy of the model. This metric gives an absolute measure of the fit of the model, quantifying the average deviation in the predictions.

## RESULTS AND DISCUSSION

In this section we discuss the results from the feature selection, the developed XGboost model and an NMP case study.

### Feature Selection Result

Table 1 outlines the chosen features for evaluating various sustainability metrics, following the application of SBFS. The selected features underscore the importance of both thermodynamic properties and molecular descriptors in providing a comprehensive assessment of sustainability. Notably, critical temperature and heat capacity emerge as common thermodynamic features

across all metrics, highlighting their universal applicability in sustainability evaluations. These properties are fundamental in understanding the energy dynamics and efficiency of chemical processes. Furthermore, the inclusion of XLogP and boiling point in three out of the four metrics signifies their relevance in assessing different environmental impacts. Molecular descriptors, particularly HallKierAlpha, have been selected for their ability to represent the three-dimensionality of molecules, a factor crucial in understanding the environmental behavior and impact of chemicals. HallKierAlpha, selected for three of the four metrics, specifically captures shape representation and molecular branching, aspects essential for evaluating the environmental compatibility of chemical substances.

**Table 1:** Selected features for each endpoint metric

Metric	Thermodynamic Feature	Molecular Descriptor Feature
HHI	heat of vaporization, heat capacity, XLogP, acentric factor, critical temperature	Chi0, HallKierAlpha, SMR_VSA7, VSA_EState6, NumValenceElectrons
EQI	heat capacity, standard formation enthalpy (gas), boiling Point, critical temperature, critical volume	Chi2v, BertzCT, HallKierAlpha, qed, fr_halogen
GWP	heat capacity, boiling point, XLogP, critical temperature, critical molar volume	BertzCT, ExactMolWt, HallKierAlpha, PEOE_VSA6, NOCount
RUI	heat capacity, boiling point, XLogP, critical pressure, critical temperature	ExactMolWt, MaxAbsPartialCharge, MaxPartialCharge, NumRotatableBonds, SMR_VSA2

### Model Result

Figure 1 presents a parity plot of the predictive accuracy of the XGBoost model across the various environmental metrics. This parity plot offers a comprehensive view of the model's performance, highlighting its strengths and areas for further improvement.

Starting with the Human Health Impact (HHI) metric, as depicted in Figure 1 (a), the model demonstrates remarkable accuracy. The test set notably outperforms both the training and validation sets, achieving an  $R^2$  of 0.997. This high score indicates a strong correlation



between the predicted and actual values, signifying the efficacy of the model in predicting HHI metric. Additionally, the RMSE values across the train-validation-test sets are closely aligned, further underscoring the reliability of the predictions. The range of HHI predictions, spanning 0.63 – 12 ( $\times 10^{-6}$ ) DALY/kg<sub>chem</sub> with a 95% confidence interval, reflects the model's comprehensive coverage of potential human health impacts.

The Ecosystem Quality Impact (EQI) metric, illustrated in Figure 1 (b), presents a slightly different picture. Although the RMSE values remain within acceptable limits, suggesting general reliability, the model exhibits a significant discrepancy in the  $R^2$  value for train-validation sets but performs well on the test set. This variance indicates a need for refinement in the model to achieve a more dependable  $R^2$  value for EQI predictions. The predictions for EQI range from 0.022 – 3.0 PDF.m<sup>2</sup>.yr/ kg<sub>chem</sub>.

In the case of the Global Warming Potential (GWP) metric, the model shows good predictive performance, though it is not without its challenges. The GWP model displays a tendency to generalize to a good degree on the validation set but performs less efficiently on the testing set based on the  $R^2$  however, the RMSE for both validation and training set are similar in magnitude and order. The predicted values for GWP, ranging from 0.81 to 9.0 kgCO<sub>2</sub>-eq/kg<sub>chem</sub> within a 95% confidence interval, demonstrate capability of the model in this domain.

Lastly, the Resource Utilization Impact (RUI) metric, which can be interpreted as the Cumulative Energy Demand (CED) for chemical production, performs less in terms of prediction on the test set. The predictions for RUI, ranging from 4.5 to 15 ( $\times 10^1$ ) MJ-primary/kg<sub>chem</sub> with a 95% confidence interval, indicate a high degree of accuracy and reliability, showcasing the model's strengths in this area. In this study, the 95% confidence intervals were derived through a bootstrapping technique. This resampling approach enables us to assess the variability in the predictions made by our models. Specifically, for each model (HHI, EQI, GWP, RUI), we generated 1,000 bootstrap samples. This was achieved by randomly selecting 50% of the dataset with replacement in each iteration. Subsequently, we utilized the models to make predictions for each bootstrap sample. The construction of the 95% confidence intervals involved determining the 2.5th and 97.5th percentiles from the distribution of these bootstrap predictions. It is important to note that these intervals reflect the variability associated with the model predictions themselves, rather than the uncertainty in the hyperparameters of the XGBoost model or the variability inherent in the dataset. Table 2 also shows the comparison of the 95% confidence intervals for the actual data and the developed XGBoost model indicating a good agreement with the original data.

**Table 2:** Comparison of confidence interval for actual data and XGBoost model

Metric	Actual Data	XGBoost Model
HHI ( $\times 10^{-6}$ )	0.55 – 12	0.63 – 12
EQI	0.022 – 3.1	0.022 – 3.0
GWP	0.68 – 9.8	0.81 – 9.0
RUI ( $\times 10$ )	3.5 – 17	4.5 – 15

### Case Study: Cradle-to-cradle Life Cycle Assessment (LCA) of N-Methylpyrrolidone (NMP)

(NMP) is a polar aprotic solvent, notable for its high boiling point, and is widely used in the chemical industry, particularly in the production of polymers. Its role in polymer manufacturing is significant, but it also raises environmental concerns. The issue with NMP lies in its non-consumptive use in synthesis and processing, leading to its release as waste, a common occurrence in the fine and specialty chemical industries. The environmental and health risks associated with the disposal of NMP are well-recognized. However, the lack of suitable and safer alternatives to NMP and similar dipolar aprotic solvents has resulted in its continued widespread use in specialty chemical applications. Given these circumstances, the importance of solvent recovery after its usage becomes paramount. Recovering NMP not only mitigates the environmental and health risks but also addresses waste management concerns in the chemical industry, emphasizing the need for sustainable and responsible solvent usage practices.

#### Process Description

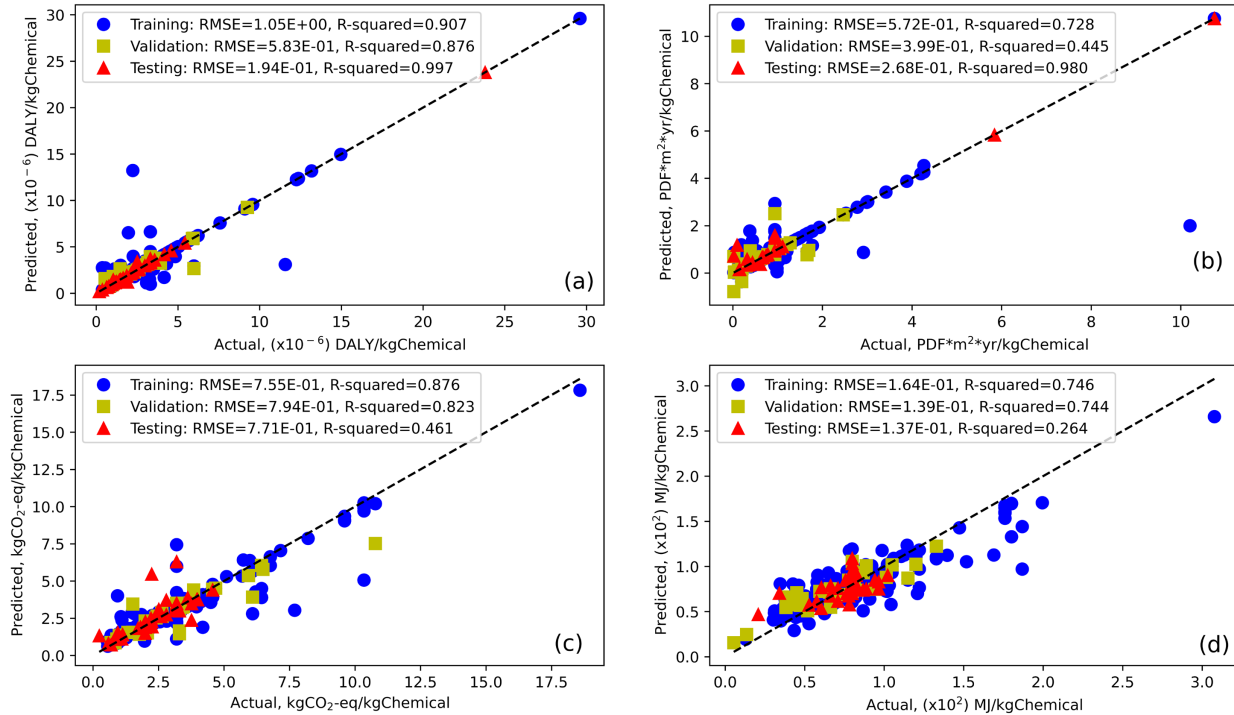
We consider a specific case where fresh solvents (n-methyl-2-pyrrolidone (NMP)) and reagents (trifluoroacetic acid (TFA), hydroxyethyl methacrylate (HEMA), hydrochloric acid (HCl)) are initially sent to a reactor together with monomers (oxydianiline (ODA), pyromellitic dianhydride (PMDA)) for making a resin (polyimide (PI)) precursor. The role of the NMP is to dissolve the ODA while the reagents are added to the reaction medium to improve the photosensitivity of the PI produced. Once the reaction is complete, the product stream flows to the washing stage where ultrapure water is used to wash the produced resin resulting in three main streams: 1) the resin precursor stream which is sent to a filter press for further processing, 2) a hazardous waste stream containing NMP, and 3) a wastewater stream. We look at three main stages in the impact assessment: 1) the cradle-to-gate (production phase) impact which entails the feed stream containing the solvents, reagents and monomers, 2) gate-to-gate (usage phase) impact which entails the energy and water usage from the reaction and washing stages, and 3) gate-to-cradle (end-of-life phase) which entails the recovery of NMP from the hazardous waste

stream. Here, the XGboost model is used to predict the cradle-to-gate impact metrics, which represent the production phase. Additionally, we use the predicted values in the gate-to-cradle phase which represent the end-of-life phase of the waste solvent. Table 3 highlights the specifications for the case study.

Equations (1) – (3) gives the environmental impacts of each phase of the Life Cycle Assessment (LCA).

$$LCA_{i,production} = \sum_j^n LCA_{i,j,production} \quad (1)$$

$$LCA_{i,use-phase} = \sum_j^m LCA_{i,k,use-phase} + LCA_{i,water,use-phase} \quad (2)$$



**Figure 1:** Parity plot for each metric from the XGBoost model. (a) HHI, (b) EQI, (c) GWP, (d) RUI

**Table 3:** NMP case study specification

Component	Inlet mass flowrate to Reactor (kg/y)	Ultrapure water for washing (kg/y)	Hazardous waste solvent stream composition (%wt)
NMP	183416	-	17
ODA	24055	-	-
PMDA	26202	-	-
HEMA	5448	-	0.5
TFA	5448	-	0.5
HCl	5448	-	0.5
H <sub>2</sub> O	-	4114148	81.5

**Table 4:** Impact metric prediction for each chemical for the production phase from XGboost model

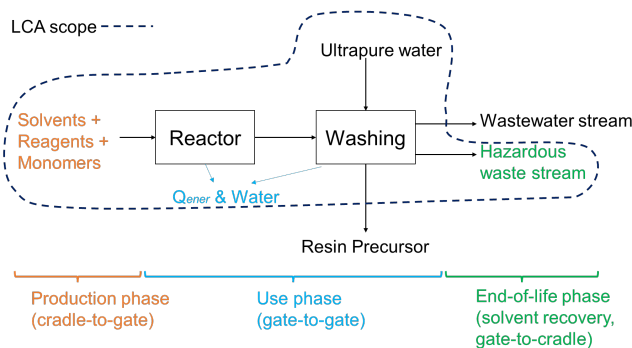
Component	HHI (x10 <sup>-6</sup> , DALY/kg <sub>Chem</sub> )	EQI (PDF.m <sup>2</sup> .yr/kg <sub>Chem</sub> )	GWP (kgCO <sub>2</sub> eq/kg <sub>Chem</sub> )	RUI (x10, MJ/kg <sub>Chem</sub> )
NMP	7.6	1.9	2.7	11.3
ODA	1.7	0.6	2.5	14.1
PMDA	3.9	0.9	2.0	16.5
HEMA	1.8	1.0	2.5	7.3
TFA	1.2	0.9	2.7	5.4
HCl	0.7	0.2	0.7	0.2

$$LCA_{i,EoL} = (1 - R_{rec,NMP})LCA_{i,NMP,EoL} + \sum_{j=1}^n LCA_{i,k,EoL} + \sum_j^m LCA_{i,k,EoL} \quad (3)$$

Here,  $LCA_{i,j,production}$  is the environment impact metric  $i$  for the production of chemical  $j$ , and  $n$  is the total number of chemicals.  $LCA_{i,k,use-phase}$  is the environmental metric  $i$  for the energy demand of technology  $k$ ,  $m$  is the total number of technologies in the process,  $LCA_{i,water,use-phase}$  is the impact metric for the total amount of water used in the process.  $R_{rec,NMP}$  is the amount of NMP recovered for reuse,  $LCA_{i,NMP,EoL}$  is the environmental impact metric for NMP,  $LCA_{i,j,EoL}$  is the environmental impact of the remaining chemicals not being recovered, and is the environmental impact due to the energy demand of the technologies (pervaporation in this case) for the solvent recovery process. The total cradle-to-cradle impact assessment per kg of NMP is given by Equation (4).

$$LCA_{i,cradle-to-cradle} = LCA_{i,production} + LCA_{i,use-phase} + LCA_{i,EoL} \quad (4)$$

Here,  $i$  is the environmental impact indicator (HHI, EQI, GWP, RUI),  $LCA_{i,production}$  is the impact metric for the production phase of the chemicals (cradle-to-gate),  $LCA_{i,use-phase}$  is the impact metric for the use-phase of the chemicals – in this case the energy demand and water usage in the reaction and washing stages (gate-to-gate) and  $LCA_{i,EoL}$  is the impact metric for the EoL phase (gate-to-cradle) for the chemicals in the hazardous waste stream. NMP is the functional unit for the analysis hence the impact metric analysis is per kg of NMP basis. Figure 2 shows the LCA scope and the associated stages of the cradle-to-cradle assessment.



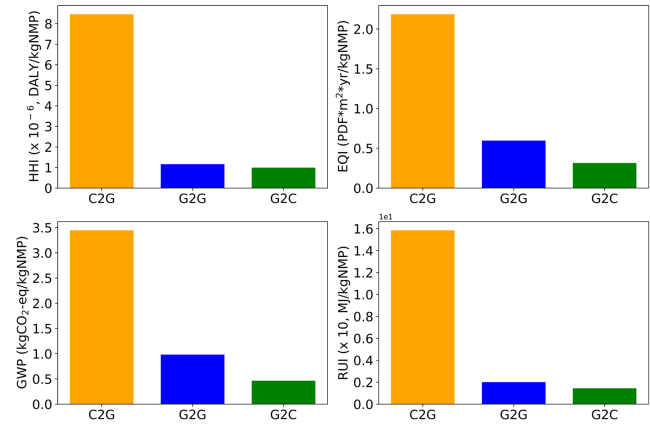
**Figure 2:** LCA scope and the various aspects of the cradle-to-cradle assessment

### Case Study Results & Discussion

Table 4 shows the predictions for each chemical for the production phase assessment from the developed XGBoost model. To give a perspective of how the predictions compare with actual data from the SimaPro® software, the HHI, EQI, GWP, and RUI are  $7.6 \times 10^{-6}$ , 1.93, 7.7,

and 168.9, respectively for NMP. It should be noted that for water, we used SimaPro® values for the analysis since we did not consider it to be a chemical.

Figure 3 shows the comprehensive impact of the NMP lifecycle.



**Figure 3:** Endpoint impacts of NMP across its lifecycle. C2G:Cradle-to-Gate; G2G:Gate-to-Gate; G2C:Gate-to-Cradle

In the HHI, the results indicate a markedly higher impact in the C2G phase, illustrating the significant health-related implications during the initial stages of NMP production. This could be attributed to the extraction and processing of raw materials, underscoring the need for stringent health and safety measures during these processes. The EQI shows similar trend with the highest impact also observed in the C2G phase. However, the G2G is higher in this case signifying a higher impact contribution from the use-phase. For GWP, the difference between the C2G and G2G is 71.5%, while that between the G2G and G2C is 52.7%. Similar trend is observed with the RUI metric.

### CONCLUSION

In this work, an XGBoost model is developed to predict four endpoint impact metrics of chemicals based on thermodynamic properties and molecular descriptors. The developed model was subsequently used in a case study where a cradle-to-cradle life cycle assessment is performed with NMP as the functional unit. The model is used to predict the production phase of the various chemicals used, and subsequently used in solvent recovery which is the considered route for the end-of-life phase. The use-phase is analysed using the utilities from the reaction and washing processes. The model results indicate that the human health impact has the best accuracy. While the remaining three metrics had significant improvements on the validation set, the models could be improved to enhance the predictions on the test set. Regarding the life cycle assessment, it is observed that the

cradle-to-gate stage has the significant impact on the lifecycle followed by gate-to-gate, and finally, gate-to-cradle for all four metrics. Additionally, this case study shows that ML model predictions can be used to substitute unknowns data for cradle-to-gate, and even gate-to-cradle life cycle assessment. Furthermore, the developed model can be incorporated during the early-design stage to provide initial estimates of impact metrics for better decision-making.

## ACKNOWLEDGEMENTS

The authors would like to acknowledge the United State's Environmental Protection Agency Pollution Prevention Grant Program funded by the Bipartisan Infrastructure Law (Grant# 4U – 96236522).

## REFERENCES

1. H. Cabezas and U. Diwekar, *Sustainability: Multi-Disciplinary Perspectives*, 1st ed. Bentham Science Publishers, 2012.
2. A. Argoti, A. Orjuela, and P. C. Narváez, "Challenges and opportunities in assessing sustainability during chemical process design," *Current Opinion in Chemical Engineering*, vol. 26, pp. 96–103, Dec. 2019, doi: 10.1016/j.coche.2019.09.003.
3. P. Karka, S. Papadokonstantakis, and A. Kokossis, "Environmental impact assessment of biomass process chains at early design stages using decision trees," *Int J Life Cycle Assess*, p. 26, 2019.
4. P. Karka, S. Papadokonstantakis, and A. Kokossis, "Digitizing sustainable process development: From ex-post to ex-ante LCA using machine-learning to evaluate bio-based process technologies ahead of detailed design," *Chemical Engineering Science*, vol. 250, p. 117339, Mar. 2022, doi: 10.1016/j.ces.2021.117339.
5. P. Karka, S. Papadokonstantakis, and A. Kokossis, "Predictive LCA - a systems approach to integrate LCA decisions ahead of design," in *Computer Aided Chemical Engineering*, vol. 46, Elsevier, 2019, pp. 97–102. doi: 10.1016/B978-0-12-818634-3.50017-5.
6. K. D. Dahm and D. P. Visco, "Fundamentals of Chemical Engineering Thermodynamics," 2015.
7. S. I. Sandler, *Chemical, biochemical and engineering thermodynamics*, Fifth edition. Hoboken, NJ: Wiley, 2017.
8. R. Parthasarathi and A. Dhawan, "Chapter 5 - In Silico Approaches for Predictive Toxicology," in *In Vitro Toxicology*, A. Dhawan and S. Kwon, Eds., Academic Press, 2018, pp. 91–109. doi: 10.1016/B978-0-12-804667-8.00005-5.
9. S. Hongmao, "Chapter 6 - Quantitative Structure-Property Relationships Models for Lipophilicity and Aqueous Solubility," in *A Practical Guide to Rational Drug Design*, S. Hongmao, Ed., Woodhead Publishing, 2016, pp. 193–223. doi: 10.1016/B978-0-08-100098-4.00006-5.
10. T. Stuyver, F. De Proft, P. Geerlings, and S. Shaik, "How Do Local Reactivity Descriptors Shape the Potential Energy Surface Associated with Chemical Reactions? The Valence Bond Delocalization Perspective," *J. Am. Chem. Soc.*, vol. 142, no. 22, pp. 10102–10113, Jun. 2020, doi: 10.1021/jacs.0c02390.
11. R. Srivastava, "Theoretical Studies on the Molecular Properties, Toxicity, and Biological Efficacy of 21 New Chemical Entities," *ACS Omega*, vol. 6, no. 38, pp. 24891–24901, Sep. 2021, doi: 10.1021/acsomega.1c03736.
12. E. A. Aboagye *et al.*, "Systematic Design of Solvent Recovery Pathways: Integrating Economics and Environmental Metrics," *ACS Sustainable Chem. Eng.*, vol. 10, no. 33, pp. 10879–10887, Aug. 2022, doi: 10.1021/acssuschemeng.2c02497.
13. O. Jolliet *et al.*, "IMPACT 2002+: A new life cycle impact assessment methodology," *Int J LCA*, vol. 8, no. 6, pp. 324–330, Nov. 2003, doi: 10.1007/BF02978505.
14. D. A. C. Narciso and F. G. Martins, "Application of machine learning tools for energy efficiency in industry: A review," *Energy Reports*, vol. 6, pp. 1181–1199, Nov. 2020, doi: 10.1016/j.egyr.2020.04.035.
15. Q. Nguyen, I. Diaz-Rainey, and D. Kurupparachchi, "Predicting corporate carbon footprints for climate finance risk analyses: A machine learning approach," *Energy Economics*, vol. 95, p. 105129, Mar. 2021, doi: 10.1016/j.eneco.2021.105129.
16. S. Boobier, D. R. J. Hose, A. J. Blacker, and B. N. Nguyen, "Machine learning with physicochemical relationships: solubility prediction in organic solvents and water," *Nat Commun*, vol. 11, no. 1, p. 5753, Dec. 2020, doi: 10.1038/s41467-020-19594-z.
17. A. Carranza-Abaid, H. F. Svendsen, and J. P. Jakobsen, "Surrogate modelling of VLE: Integrating machine learning with thermodynamic constraints," *Chemical Engineering Science: X*, vol. 8, p. 100080, Nov. 2020, doi: 10.1016/j.cesx.2020.100080.
18. T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco California USA: ACM, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.
19. M. Chen, Q. Liu, S. Chen, Y. Liu, C.-H. Zhang, and

- R. Liu, "XGBoost-Based Algorithm Interpretation and Application on Post-Fault Transient Stability Status Prediction of Power System," *IEEE Access*, vol. 7, pp. 13149–13158, 2019, doi: 10.1109/ACCESS.2019.2893448.
20. "CIRpy — CIRpy 1.0.2 documentation." Accessed: Aug. 25, 2023. [Online]. Available: <https://cirpy.readthedocs.io/en/latest/>
  21. "chemicals: Chemical properties component of Chemical Engineering Design Library (ChEDL) — Chemicals 1.1.4 documentation." Accessed: Aug. 25, 2023. [Online]. Available: <https://chemicals.readthedocs.io/>
  22. "RDKit." Accessed: Aug. 25, 2023. [Online]. Available: <https://www.rdkit.org/>
  23. M. Goedkoop, M. Oele, J. Leijting, T. Ponsioen, and E. Meijer, "Introduction to LCA with SimaPro." *PRE Sustainability*, 2016.
  24. J. Bergstra, D. Yamins, and D. Cox, "Hyperopt: A Python Library for Optimizing the Hyperparameters of Machine Learning Algorithms," presented at the Python in Science Conference, Austin, Texas, 2013, pp. 13–19. doi: 10.25080/Majora-8b375195-003.
  25. P. I. Frazier, "Bayesian Optimization," in *Recent Advances in Optimization and Modeling of Contemporary Problems*, in INFORMS TutORials in Operations Research. , INFORMS, 2018, pp. 255–278. doi: 10.1287/educ.2018.0188.

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Design and Optimization of Circular Economy Networks: A Case Study of Polyethylene Terephthalate (PET)

Abdulhakeem Ahmed<sup>a</sup>, Ana I. Torres<sup>a\*</sup>

<sup>a</sup> Carnegie Mellon University, Department of Chemical Engineering, Pittsburgh, Pennsylvania, USA

\* Corresponding Author: [aitorres@cmu.edu](mailto:aitorres@cmu.edu)

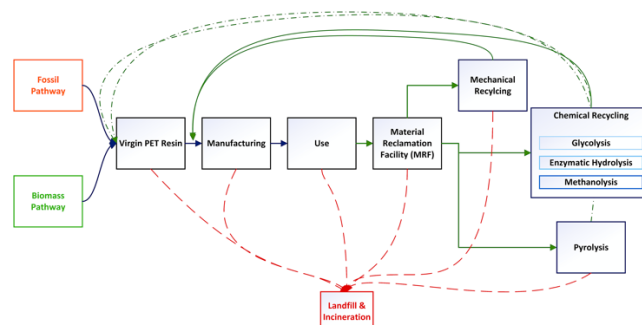
## ABSTRACT

Circular systems design is an emerging approach for promoting sustainable development. Despite its perceived advantages, the characterization of circular systems remains loosely defined and ambiguous. This work proposes a network optimization framework that evaluates three objective functions related to economic and environmental domains and employs a Pareto analysis to illuminate the trade-offs between objectives. The US polyethylene terephthalate (PET) value chain is selected as a case study and represented via a superstructure containing various recycling pathways. The superstructure optimization problems are modeled as a mixed integer linear program (MILP) and linear programs (LPs), implemented in Pyomo, and solved with CPLEX for a one-year assessment horizon. Solutions to the circular economy models are then compared to the corresponding solutions of linear economy models. Preliminary results show that the optimal circular network is advantageous over the optimal linear network for all objectives subject to the current market supply of raw materials and the total cost of production. However, when considering the present chemical processing infrastructure of the US economy and unrestricted biomass feedstock availability, a linear economy is favorable as an outcome of low operating cost and carbon sequestration.

**Keywords:** Circular Economy, Supply Chain Optimization, Sustainability, Plastic Recycling

## INTRODUCTION

With the growth of modern societies, waste management and finite resource depletion have become problematic. A primary facilitator of this phenomenon is the extensive employment of linear systems where materials are extracted, used, and discarded. Due to heightened concerns regarding finite resource depletion and the environmental effects of mismanaged waste, there is growing interest in adopting circular economies. A circular economy (CE) strives to eliminate waste and pollution, circulate products and materials at their highest level, and regenerate nature [1]. Applications of a CE have been present, to some extent, at the micro-level of some organizations. However, what remains elusive is effective representation at the macro-level and well-defined metrics and methodologies for achieving and quantifying circularity.



**Figure 1.** High-level superstructure representation.

Each year, plastics constitute roughly 400 million tonnes of generated waste, of which approximately 80% enters landfills or is emitted into the environment [2]. Of all produced plastics, polyethylene terephthalate (PET) is one of the most widely consumed, with applications of PET ranging from consumer packaging to films and textiles (polyester). In the US, historical demands for polyester and bottle-grade PET comprise over 85% of all PET

products [3]. However, recycling rates for bottles and general textiles are approximately 30% and 15%, respectively [4,5]. For these reasons, we consider the PET economy as an application case study for assessment.

A holistic representation of the PET value chain (product life cycle) and potential circular processing pathways is depicted in Figure 1. The analysis is posed as a superstructure optimization problem. The superstructure includes fossil and biomass feedstocks for PET precursors, production routes from these precursors to virgin PET (vPET), manufacturing and use phases for PET-based products, and end-of-life (EoL) processing technologies. EoL processes include traditional and novel processing of the following grouping: landfilling, incineration, mechanical recycling (MR), chemical recycling (CR), and thermochemical recycling.

The application of various recycling processes allows for material retention within the value chain. However, it remains unclear how favorable these processes are over a dominantly linear PET economy. To enhance this understanding, we formulate a mixed integer linear programming problem (MILP) where we assess an economic, environmental, and material utilization objective. The optimization problems are solved to determine the optimal selection of processing technologies and material flows that satisfy demand and are then compared to the optimal solutions of a linear economy reference model. The linear reference model is obtained by imposing flow restriction constraints, which void recycle flows. A Pareto analysis is then performed using the epsilon-constrained method defined in [6] to identify the trade-offs between the objectives.

## METHODS

### Modeling Framework

For modeling purposes, a State Task Network (STN) formulation where processes (tasks) consume and produce materials (states) is followed. Two types of nodes are defined, one constituting key chemical components and the other technologies that convert components. Technology nodes ( $j$ ) accept flows from component nodes, performing composition transformations and acting as the influx to other component nodes. Component nodes ( $i$ ) combine fluxes from different technology nodes and are also connected to the external market, allowing purchases and sales of material. The characterization of the nodes is depicted in Figure 2. A technology matrix ( $A$ ) is defined to contain process conversion information representing transformations across technology nodes. The technology matrix is informed by industrial data and process simulations from the literature.

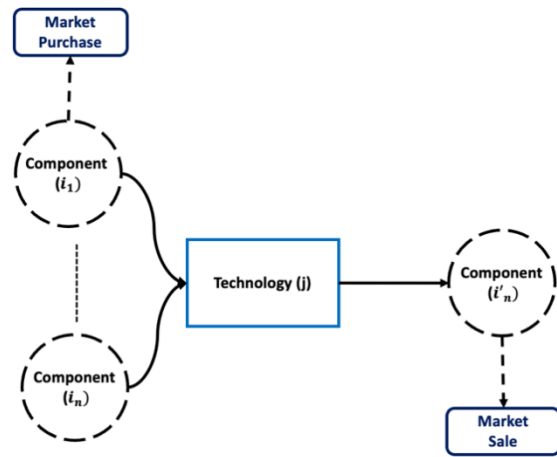


Figure 2. Network node characterization.

The generalized model equations for the given framework are defined by equations (1-7). Eqs. (1) and (2) define the generation and consumption of component  $i$  in technology  $j$ . Where  $\gamma_{ij}$ , a conversion factor, is positive for generative transformations and negative for consumption. Parameter  $\gamma_{ij}$  constitutes the members of matrix ( $A$ ) and is defined by industrial data aggregated by Franklin Associates [7], the US EIA [8], the US EPA [9], and Aspen simulations documented in the literature [10-17]. Variable  $l_j$  represents the total flow through a technology node. Sets  $J^G$  and  $J^C$  designate producing and consuming technology nodes, respectively.

$$g_{ij} = \gamma_{ij} l_j \quad \forall i \in I, \forall j \in J^G \quad (1)$$

$$c_{ij} = -\gamma_{ij} l_j \quad \forall i \in I, \forall j \in J^C \quad (2)$$

$$\sum_{j \in J} \gamma_{ij} l_j + p_i = s_i \quad \forall i \in I \quad (3)$$

Eq. (3) represents the general material balance for all components in the set  $I$ , stating that the purchase plus the total generation and consumption of  $i$  across the set of technologies ( $J$ ) is equal to the market sales or excess of  $i$ . Furthermore, component purchases ( $p_i$ ) are bounded by an upper limit ( $p_i^L$ ), which in this analysis is the five-year historical average supply for all raw materials. Eq. (5) states that the demand for end products must be satisfied for all  $I$  belonging to a subset  $I^D$ . Table 1 represents the annual rate of PET consumption (megaton) in the US in 2018 [3,18].

$$p_i \leq p_i^L \quad \forall i \in I \quad (4)$$

$$\sum_{j \in J^G} g_{ij} \geq d_i \quad \forall i \in I^D \quad (5)$$

$$g_{ij'} \leq \eta_i \sum_{j \in J^G} g_{ij} \quad \forall i \in I^G \quad (6)$$

$$c_{ij'} \leq \eta_i \sum_{j \in J^C} c_{ij} \quad \forall i \in I^C \quad (7)$$

Lastly, Eqs. (6-7) generalizes the process constraints imposed on network flows to satisfy select technical requirements. These equations state that the

production or consumption of component  $i$  by  $j$  is limited by the net generation or consumption of  $i$  multiplied by a scalar  $\eta_i$ . The parameter  $\eta_i$  represents flow ratios and corresponds to current process capabilities detailed in the literature.  $I^G$  and  $I^C$ , are subsets of  $I$ . And  $j'$  is a referenced technology node corresponding to  $\eta_i$ . For example, one such constraint is the allowable composition of mixed virgin PET resin and recycled PET (rPET) resin produced by mechanical recycling that satisfies intrinsic viscosity requirements for processing into new bottles. Here,  $\eta_i$  is 0.35 [19] and  $j'$  is the mechanical recycling of post-use bottles. Additional technical constraints include an upper bound on the fraction of end-of-life textiles (EoL) to be downcycled to satisfy a portion of demand. Secondly, there is an upper bound on the amount of non-bottle/non-polyester EoL materials that can be recycled, which discounts products that generally cannot be recycled, such as food containers.

**Table 1:** Case study demand specification.

Product Type	Demand (Mt/year)
Bottles	2.84
Polyester	7.84
Films/sheets/others	1.17

## Design Objectives

Three design objectives are assessed to explore economic and environmental domains. These include total annualized cost of production, process greenhouse gas emissions (GHGs), total virgin raw material utilization.

### Total Annualized Cost

$$TAC = \sum_{j \in J^E} (opex_j l_j + ACCR \sum_{t \in T} capex_{jt}) \quad (8)$$

The economic dimension is evaluated with Eq. (8) which states that the total annualized cost of production is equal to the sum of operational costs and annualized capital expenses (ACC) for the technologies. Operating cost ( $opex_j$ ) equals the sum of feed, utility, and labor minus the sale of by-products for each  $j$ . We assume a ten-year amortization period and an interest rate of 15%, which correlates to an annual capital charge ratio (ACCR) of 0.199 [20]. Investment costs ( $capex_{jt}$ ) are indexed by plant ( $t$ ) belonging to the set of plants ( $T$ ). Set  $J^E$  is a subset of economically evaluated technologies.

$$\frac{capex_j^{ref}}{(capacity_j^{ref})^{0.6}} \left[ a^{0.6} + \frac{b^{0.6} - a^{0.6}}{b - a} (capacity_{jt} - az_{jt}) \right] \quad (9)$$

$$capacity_{jt} \leq Mz_{jt} \quad \forall j \in J, \forall t \in T \quad (10)$$

$$l_j \leq \sum_{t \in T} capacity_{jt} \quad \forall j \in J \quad (11)$$

To account for economies of scale, the six-tenths rule is applied to estimate  $capex_{jt}$ . This rule leads to a non-linear non-convex equality. Thus, a secant

linearization [21] as presented in Eq. (9) is applied, where  $a$  and  $b$  are lower and upper bounds of the approximating function.  $capex_j^{ref}$  and  $capacity_j^{ref}$  are reference parameters taken from the literature. And the capacity of plant  $t$  for process  $j$  is represented by  $capacity_{jt}$ . Furthermore, the binary variable  $z_{jt}$  indicates whether a plant  $t$  is selected for the process  $j$ . Eq. (10) ensures that capacity is zero if plant  $t$  of process  $j$  is not selected, with  $M$  equal to the maximum plant capacity. Lastly, Eq. (11) bounds net flow through  $j$  by the total capacity of  $j$ .

### GHG Emissions

$$GHG = \sum_{j \in J^E} \phi_j g_{ij} \quad \forall i \in I^R \quad (12)$$

The environmental objective presented by Eq. (12) represents raw material extraction, processing, and manufacturing emissions. It states that the net greenhouse gas (GHG) emissions are equal to the total generation of components  $i$  in the set of reference products  $I^R$  multiplied by the emission factor  $\phi_j$  which measures carbon equivalents ( $CO_2e$ ). Set  $I^R$  is defined such that emission parameters are normalized per reference product belonging to each evaluated technology node. The considered GHGs are carbon dioxide, methane, and nitrous oxide. Parameter  $\phi_j$  is determined with process data available in the literature.

### Virgin Raw Material Utilization

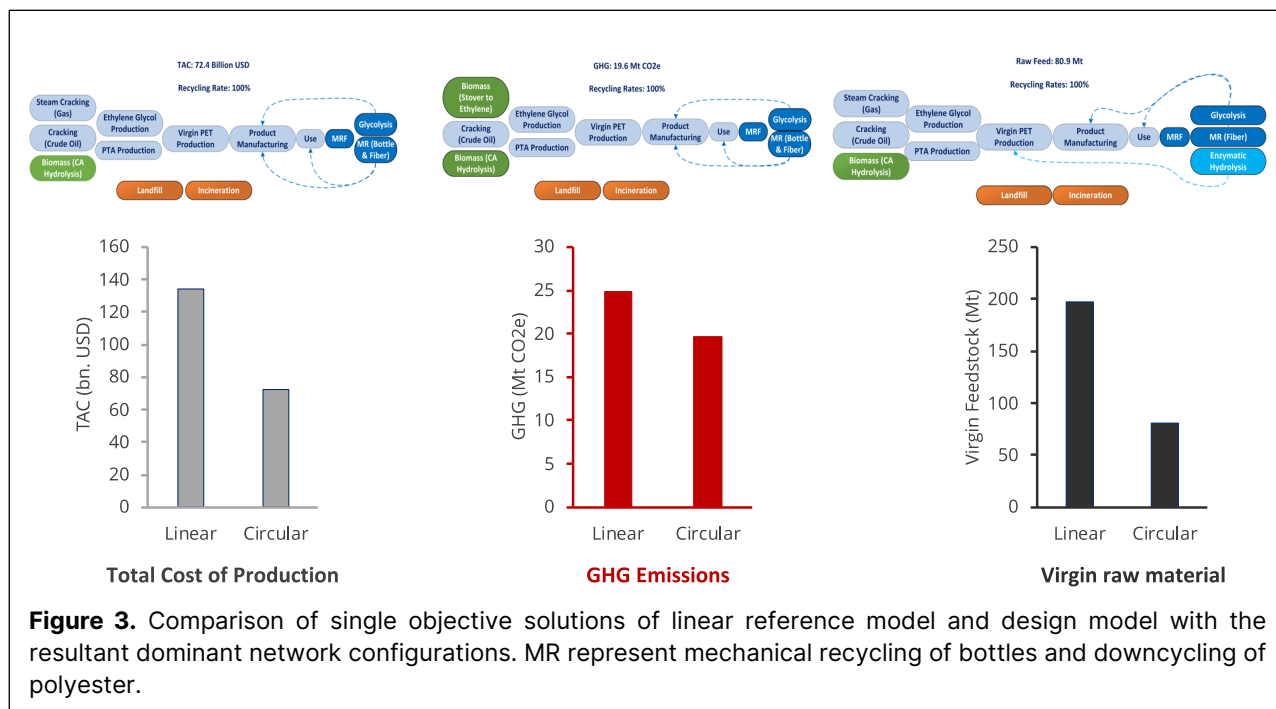
$$VF = \sum_{i \in I^F} p_i \quad (13)$$

The final design goal considers raw material use and measures network material efficiency. This objective serves as a proxy for quantifying circularity. The objective function is defined by Eq. (13) and represents the sum of market purchases of virgin feedstock, where  $I^F$  describes feedstock components.

### Optimization Problem

The analysis involves solving the optimization problem defined in Eq. (14). Here,  $Z$  represents each objective function, where Eq. (14) is a MILP for the annualized cost objective but an LP for both environmental and material use objectives as Eq. (9-11) pertain only to the economic problem.

$$\begin{aligned} & \min Z \\ & s. t. Eq. (1 - 7), Eq(9 - 11) \\ & l_j, s_i, p_i, capex_{jt}, capacity_{jt} \geq 0, z_{jt} \in \{0,1\} \end{aligned} \quad (14)$$



## RESULTS

### Single-Objective Solutions

The optimization problem is solved using CPLEX version 22.1.1 and Pyomo, a Python optimization modeling language [22]. The results and comparisons for the single objective problems are presented in Figure 3. Considering the economic objective, the optimal network configuration of the design model is circular, employing both mechanical recycling and glycolysis. Final products from recycling include rPET that is sent directly to the manufacturing stage to create final PET end products. Comparing this optimal network to the solution of the corresponding linear reference model shows that a circular network offers a 46% cost reduction over the linear counterpart.

Likewise, for the GHG emission objective, the solution of the design model leads to a circular network achieved with mechanical recycling and glycolysis. However, contrary to the network configuration of the economic objective, more biomass feedstock is used over fossil feedstock. Additionally, comparing the solution of the GHG objective to the corresponding linear reference gives a 21% reduction in supply chain process emissions. This reduction is lower relative to that achieved by the economic objective.

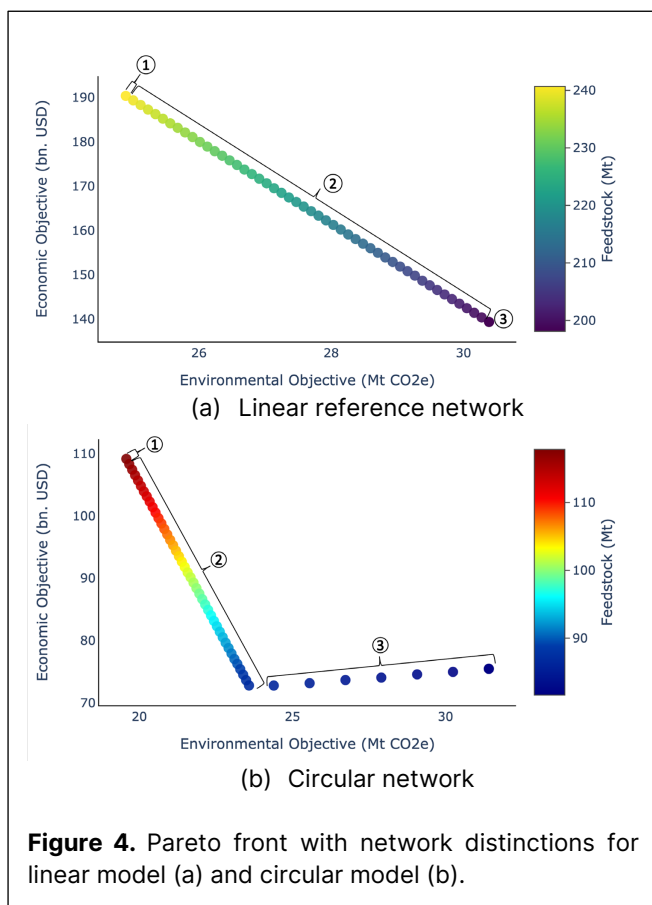
Lastly, the solution of the circular design model for the virgin raw material objective is a circular network with a 59% reduction in raw material consumption. However, compared to the previous design solutions, this network consists of a much higher application of chemical recycling via glycolysis and enzymatic hydrolysis and the rest

of the downcycling of polyester. In addition to rPET production, monomers are produced and sent upstream for polymerization into PET.

In addition to the base cases, additional considerations are evaluated for the economic and environmental objectives. The first assigns zero Capex to developed US processes to account for existing infrastructure. The resultant networks for both linear and circular models converge to an equivalent linear network. The selected network utilizes a purely fossil-based pathway employing steam cracking of natural gas and cracking of crude oil. The second consideration is unrestricted access to all feedstock, where unlimited is defined by altering market bounds to allow one gigatonne for all feed components. Like the first consideration, solving the reference and circular model leads to an equivalent linear network. However, contrary to the previous network, the obtained network follows a purely biomass-based pathway via biomass-to-ethylene conversion and acid hydrolysis of biomass to produce paraxylene.

### Pareto Assessment

The Pareto fronts for the design and linear models are shown in Figure 4. The linear model exhibits a linear trade-off between all objectives, where reducing emissions leads to higher-cost networks that consume more virgin feedstock. This trend occurs as increasing biomass feed consumption leads to reduced emissions. Yet, because process conversions of biomass processes are relatively low compared to fossil processes, higher raw material input is required to satisfy demand. These observations are further validated via the change in configuration of the selected optimal networks numbered in Figure 4.



**Figure 4.** Pareto front with network distinctions for linear model (a) and circular model (b).

In Figure 4(a), network one corresponds to the solution of the GHG emission objective presented in Figure 3 without selections for recycling. Likewise, network three is consistent with the cost objective network in Figure 3, without recycling. Network two corresponds to an intermediary of network one with steam cracking of natural gas without recycling, supporting the observable transition from biomass pathways to better conversion fossil pathways.

For the circular model, the Pareto front is piecewise linear. As shown in Figure 4(a), the leftmost trend line abides by the reasoning discussed for the linear model. However, after some threshold, attaining increased material circularity (lesser virgin feedstock consumption) leads to more emissions and higher annualized costs. The rationale for this behavior is that further reductions in material consumption demand a reasonably material-efficient network. Such a network is achievable by selecting more chemical recycling processes, contributing to increased capital investment and GHG emissions. Like the linear model, these trends are supported by the network configurations spanning the Pareto front. Looking at Figure 4(b), network three corresponds to the solution of the virgin raw material objective presented in Figure 3, plus the mechanical recycling of bottles. Network one corresponds to the GHG emission objective in Figure 3. And network two, an intermediary of network one with steam

cracking of natural gas.

## CONCLUSIONS

This work formulates a superstructure optimization problem for the US polyethylene terephthalate value chain to determine an optimal network selection consisting of processing technologies and material flow pathways satisfying each design goal for a one-year assessment period. The three assessed design goals are total annualized cost, GHG process emissions, and virgin feedstock consumption, which served as a proxy for measuring material circularity. Finally, a multi-objective Pareto assessment was performed to illuminate existing trade-offs between competing objectives for the linear and circular design models.

The results of the single-objective analysis indicate that circular PET supply chains offer lower GHG process emissions, water consumption, and virgin feedstock usage. Additionally, under current market conditions, a circular PET network is a better investment than a linear network when constructed from a zero-infrastructure ground state. Next, the results of the Pareto assessment portray a linear trade-off between objectives for the linear model, with a correlation between total annualized cost and virgin feedstock consumption, which vary inversely with GHG process emissions. The same observation is initially present for the circular model. However, attaining better material circularity increases the carbon intensity and cost for the selected networks, chiefly due to the additional processing technologies required to realize greater network efficiency.

## ACKNOWLEDGEMENTS

This work is funded by Carnegie Mellon University (A.I. Torres start-up package), the Steinbrenner Institute for Environmental Education and Research Doctoral Fellowship, and the GEM Fellowship.

## REFERENCES

1. Ellen MacArthur Foundation. Circular economy introduction- what is a circular economy. Webpage, 2023. <https://ellenmacarthurfoundation.org/topics/circular-economy-introduction/overview>.
2. Environment, U.N. Drowning in plastics – marine litter and plastic waste vital graphics. Report, (2021, October 21).
3. Raymond L Smith, Sudhakar Takkellapati, and Rachele C Riegerix. Recycling of plastics in the united states: plastic material flows and polyethylene terephthalate (pet) recycling processes. *ACS sustainable chemistry & engineering*, 10(6):2084–2096, 2022.
4. NAPCOR. Pet recycling report (2018). Report, (2020, November 30). <https://napcor.com/>



- reports-resources/.
5. US EPA. Advancing sustainable materials management: 2018 tables and figures. Report, (2020, December). [https://www.epa.gov/sites/default/files/2021-01/document/2018\\_tables\\_and\\_figures\\_dec\\_2020\\_fnl\\_508.pdf](https://www.epa.gov/sites/default/files/2021-01/document/2018_tables_and_figures_dec_2020_fnl_508.pdf).
  6. George Mavrotas. Effective implementation of the  $\epsilon$ -constraint method in multi-objective mathematical programming problems. *Applied mathematics and computation*, 213(2):455–465, 2009.
  7. Franklin Associates. Cradle-to-resin life cycle analysis of polyethylene terephthalate resin. Report, 2020. [https://circularsolutionsadvisors.com/wp-content/uploads/2022/09/PET\\_NAPCOR\\_Study.pdf](https://circularsolutionsadvisors.com/wp-content/uploads/2022/09/PET_NAPCOR_Study.pdf).
  8. EIA. Oil and petroleum products explained: Refining crude oil. Article, 2023, June 12. <https://www.eia.gov/energyexplained/oil-and-petroleum-products/refining-crude-oil-inputs-and-outputs.php>. Ap 42, fifth edition, volume i, chapter 6: Organic chemical process industry. Website, 2018, November 19. <https://www3.epa.gov/ttnchie1/ap42/ch06/index.html>.
  9. Abhay Athaley, Praneeth Annam, Basudeb Saha, and Marianthi Ierapetritou. Techno-economic and life cycle analysis of different types of hydrolysis process for the production of p-xylene. *Computers & Chemical Engineering*, 121:685–695, 2019.
  10. Patnarin Benyathiar, Pankaj Kumar, Gregory Carpenter, John Brace, and Dharmendra K Mishra. Polyethylene terephthalate (pet) bottle-to-bottle recycling for the beverage industry: A review. *Polymers*, 14(12):2366, 2022.
  11. Md Emdadul Haque, Namit Tripathi, Srinivas Palanki, Qiang Xu, and Krishna DP Nigam. Plant-wide modeling and economic analysis of monoethylene glycol production. *Processes*, 10(9):1755, 2022.
  12. Jiaze Ma, Philip A Tominac, Horacio A Aguirre-Villegas, Olumide O Olafasakin, Mark M Wright, Craig H Benson, George W Huber, and Victor M Zavala. Economic evaluation of infrastructures for thermochemical upcycling of post-consumer plastic waste. *Green Chemistry*, 25(3):1032–1044, 2023.
  13. Andrea P Ortiz-Espinoza, Mohamed MB Noureldin, Mahmoud M El-Halwagi, and Arturo Jimenez-Gutiérrez. Design, simulation and techno-economic analysis of two processes for the conversion of shale gas to ethylene. *Computers & Chemical Engineering*, 107:237–246, 2017.
  14. Avantika Singh, Nicholas A Rorrer, Scott R Nicholson, Erika Erickson, Jason S DesVeaux, Andre FT Avelino, Patrick Lamers, Arpit Bhatt, Yimin Zhang, Greg Avery, et al. Techno-economic, life-cycle, and socioeconomic impact analysis of enzymatic recycling of poly (ethylene terephthalate). *Joule*, 5(9):2479–2503, 2021.
  15. Minbo Yang, Xueyu Tian, and Fengqi You. Manufacturing ethylene from wet shale gas and biomass: comparative technoeconomic analysis and environmental life cycle assessment. *Industrial & Engineering Chemistry Research*, 57(17):5980–5998, 2018.
  16. Taylor Uekert, Avantika Singh, Jason S DesVeaux, Tapajyoti Ghosh, Arpit Bhatt, Geetanjali Yadav, Shaik Afzal, Julien Walzberg, Katrina M Knauer, Scott R Nicholson, et al. Technical, economic, and environmental comparison of closed-loop recycling technologies for common plastics. *ACS Sustainable Chemistry & Engineering*, 11(3):965–978, 2023.
  17. NAPCOR. Pet recycling report (2018). Report, (2020, November 30). <https://napcor.com/reports-resources/>.
  18. Utkarsh S Chaudhari, Yingqian Lin, Vicki S Thompson, Robert M Handler, Joshua M Pearce, Gerard Caneba, Prapti Muhuri, David Watkins, and David R Shonnard. Systems analysis approach to polyethylene terephthalate and olefin plastics supply chains in the circular economy: a review of data sets and models. *ACS Sustainable Chemistry & Engineering*, 9(22):7403–7421, 2021.
  19. Gavin Towler and Ray Sinnott. *Chemical engineering design: principles, practice and economics of plant and process design*. Butterworth-Heinemann, 2021.
  20. Ignacio E Grossmann. *Advanced optimization for process systems engineering*. Cambridge University Press, 2021.
- Bynum, Michael L., Gabriel A. Hackebeil, William E. Hart, Carl D. Laird, Bethany L. Nicholson, John D. Sirola, Jean-Paul Watson, and David L. Woodruff. *Pyomo - Optimization Modeling in Python*. Third Edition Vol. 67. Springer, 2021.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Economic Optimization and Impact of Utility Costs on the Optimal Design of Piperazine-Based Carbon Capture

Ilayda Akkor<sup>a</sup>, Shachit S. Iyer<sup>b</sup>, John Dowdle<sup>b</sup>, Le Wang<sup>b</sup>, and Chrysanthos Gounaris<sup>a</sup>

<sup>a</sup> Carnegie Mellon University, Department of Chemical Engineering, Pittsburgh, PA, USA

<sup>b</sup> The Dow Chemical Company, Lake Jackson, Texas, USA

\* Corresponding Author: [iakkor@andrew.cmu.edu](mailto:iakkor@andrew.cmu.edu).

## ABSTRACT

Recent advances in process design for solvent-based, post-combustion capture (PCC) processes, such as the Piperazine/Advanced Flash Stripper (PZ/AFS) process, have led to a reduction in the energy required for capture. Even though PCC processes are progressively improving in Technology Readiness Levels (TRL), with a few commercial installations, incorporating carbon capture adds cost to any operation. Hence, cost reduction will be instrumental for proliferation. The aim of this work is to improve process economics through optimization and to identify the parameters in our economic model that have the greatest impact on total cost to build and operate these systems. To that end, we investigated changes to the optimal solution and the corresponding cost of capture considering changes in the price of utilities and solvent. We found that changes in solvent price had the most effect on the cost of capture. However, re-optimizing the designs in the event of price changes did not lead to significant improvements in the case of piperazine, cooling water and electricity, whereas re-optimizing for changes in steam prices lead to yearly saving of 3.8%. These findings show that the design choices obtained at the nominal optimal solution are insensitive to utility price changes except for the case of steam and that there is a need for altered designs for locations where the steam prices are different.

**Keywords:** post-combustion carbon capture, rate-based model, optimization, nonlinear programming, sensitivity analysis.

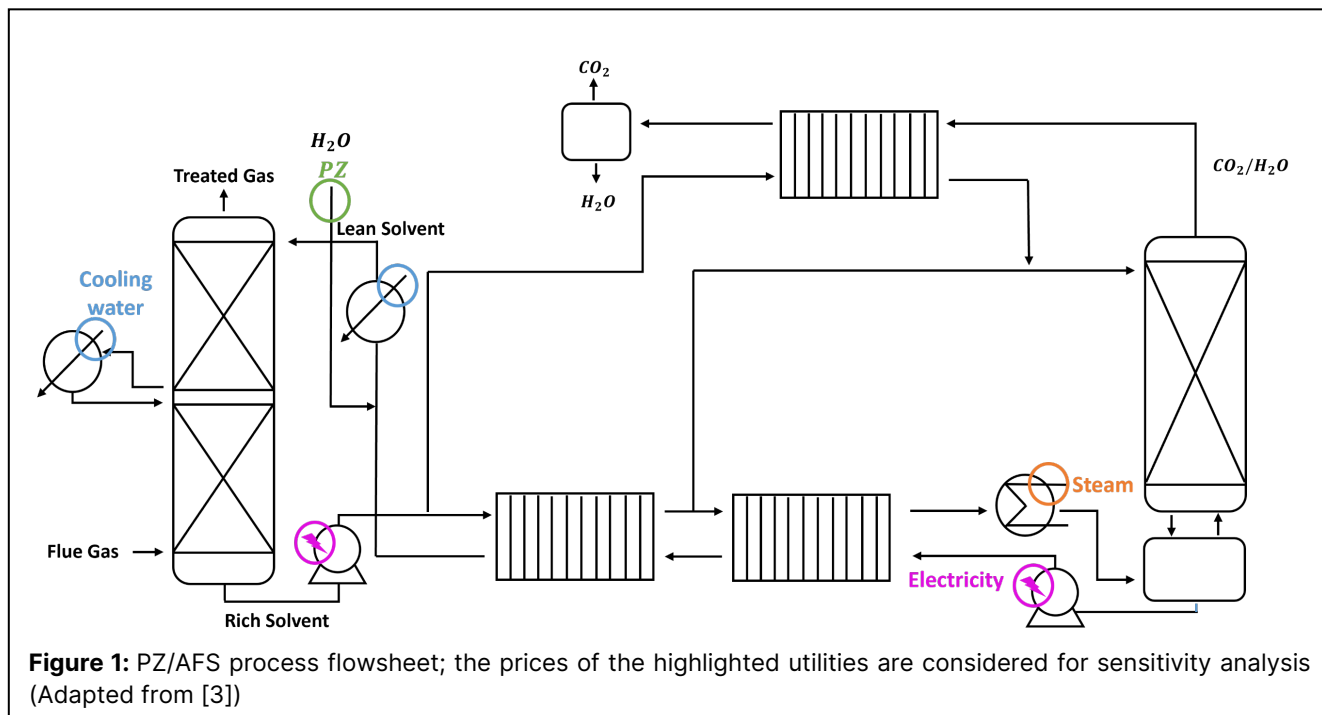
## INTRODUCTION

### Solvent-Based Carbon Capture Process Modeling

Solvent-based post-combustion carbon capture (PCC) processes are the most promising way to mitigate the increasing CO<sub>2</sub> emissions. There has been significant focus on trying to reduce the energy and cost required for capture, one prominent example being the Piperazine/Advanced Flash Stripper (PZ/AFS) process [1], shown in Fig 1. This novel process uses PZ as the solvent and the AFS process design, as shown. Flue gas containing CO<sub>2</sub> is contacted with the CO<sub>2</sub>-lean solvent in a packed absorber column. An intercooler in the middle of the column, as well as a cooler at the top, are used to drive down the solvent temperature to improve absorption. After leaving the absorber, the CO<sub>2</sub>-rich solvent passes through a heat exchanger network that has two

split streams; one of which is pumped directly to the top of a stripper column, while the other is pumped to a steam heater as shown in Fig 1. After the solvent is further heated in a steam heater, it gets flashed inside of a flash tank where the resulting vapor then enters the stripper column. After stripping, a vapor phase containing a CO<sub>2</sub> - water mixture exits the column from the top. It is then cooled down in a heat exchanger and finally separated out in a condenser drum. The lean solvent recovered in the flash tank is pumped back to the absorber column with the addition of make-up PZ and water.

The economics of PCC processes can be improved by optimization, which requires a mathematical model of the process. Hence, a detailed model of the PZ/AFS process was built using an equation-oriented approach in the IDAES framework, an open source, Python-based modeling medium [2]. In addition to the absorber and stripper columns, our model also includes the flash tank, all heat exchangers, and a recycle stream with solvent



make-up. For the columns, a rate-based modeling approach was used (as opposed to equilibrium-based), and a finite-difference scheme was used for discretization of the system of differential algebraic equations (DAE). This led to a highly nonlinear programming (NLP) model with over 8600 variables and constraints. The solution strategy includes a tailored, multi-level initialization cascade for reliable model convergence [3]. Once the model results were validated with pilot plant data [4], economic optimization was performed, which resulted in identifying the design and operating conditions that led to a considerable reduction in the capital and operational costs compared to the baseline, at the pilot scale [3]. To test the robustness of the model and to aid the commercialization of this process, optimization was performed for varying plant capacities (processing up to 1500 times more flue gas than the pilot scale) and for feed gases with different  $\text{CO}_2$  concentrations to represent different flue gas sources that may arise in practice [3].

### Parameter Uncertainty and Variability

In chemical process models, there are many parameters that are uncertain or subject to variability which, if disregarded by the model, lead to solutions that may not be truly optimal or even feasible in real life. Hence, it is crucial to account for these during the design phase. Uncertainty in a parameter stems from a lack of knowledge of its exact value, whereas variability is concerned with a parameter that may take on a different value, for example at a different point in time or at another location [5]. In the context of process models, these parameters can be categorized into three groups: (1) *epistemic parameters*, which are inherent in any process model, and which stem

from the approximating nature of correlations used to evaluate properties (kinetic, thermodynamic, transport) and the overall assumptions made therein; (2) *operational parameters*, which pertain to the upstream conditions and may be subject to temporal variations, e.g., composition, flowrate and/or temperature of the flue gas; and (3) *economic parameters*, such as the applicable prices for utilities and the solvent.

To identify which of the above are most impactful and to quantify the extra costs necessary for designs to insure against them, a series of sensitivity analyses were conducted. More specifically, the aim of this work was to leverage the equation-oriented model that we had previously built to gather insight on the impact of the economic parameters on the optimal process designs. In our previous study it was seen that, at the commercial scale, along with the purchase cost of the absorber column, the operating expenses for cooling water, steam, electricity, and solvent renewal were the dominating costs [3]. Therefore, this study further investigated the sensitivity of the plant design and its total cost to the prices of said utilities and to the piperazine solvent.

### METHODOLOGY

For this study, we chose to perform sensitivity analyses on the commercial scale, natural gas combined cycle (NGCC) flue gas processing plant. The feed gas flowrate was 30,000 mol/s, and the parallel train configuration was used as it was done in the Mustang Station plant, presented in the work of Clossmann et al. [6]. The optimization problem was structured to minimize the total annualized cost (TAC), subject to some performance

constraints, such as a 96% capture target and flooding constraints for the columns [3]. Thirteen degrees of freedom were the packed length and the inner diameter of the absorber and stripper columns, the areas of the three heat exchangers, the two coolers and the steam heater, two bypass ratios, and solvent circulation rate. The first ten were the decision variables related to design, which must be committed before building the plant, whereas the bypass ratios and the solvent amount were operational decisions that can be adjusted.

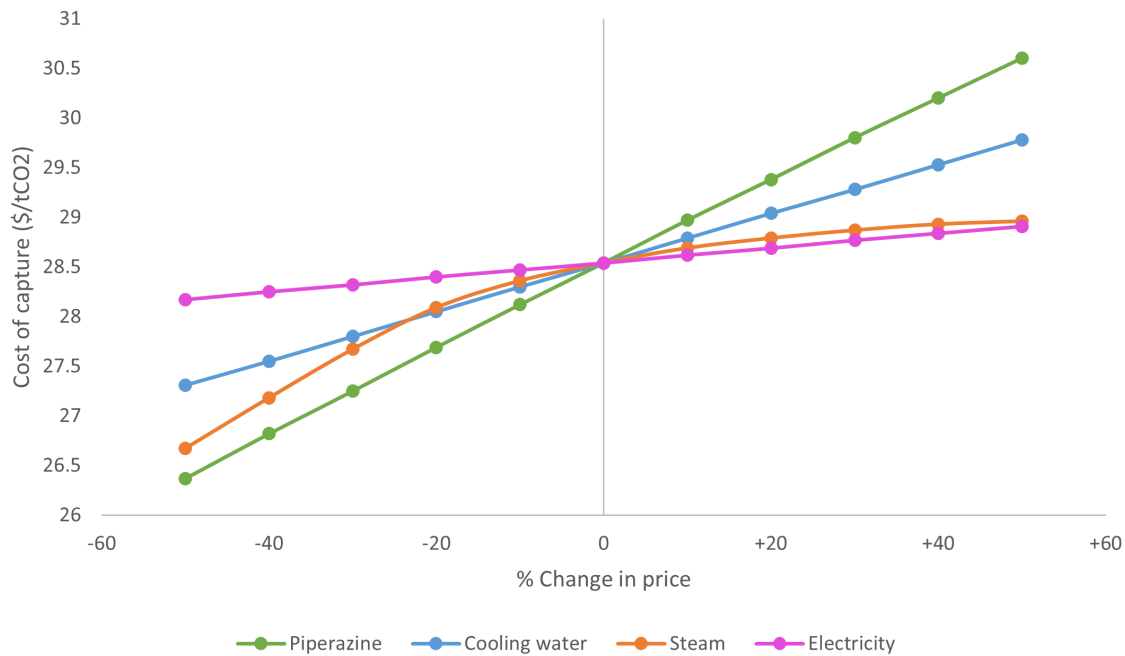
The parameters that we focused on were the prices of the utilities (e.g., steam, cooling water and electricity) and piperazine, which were nominally assumed to be at the values given in Table 1. Cooling water was used for the two coolers, as shown in Fig 1. It must be noted that the optimal solution for the nominal case for the commercial-scale natural gas combined cycle (NGCC) flue gas processing plant does not use the intercooler. Hence, in the nominal solution, cooling water is only needed for the lean cooler at the top of the absorber column. However, the optimal solution for the cases that consider price variation, the inclusion of intercooler is allowed as a degree of freedom. Steam is used to heat up the rich solvent in a heater before it enters the flash tank. Electricity is needed for pumping the lean and rich solvents between columns. Our economic model considers operational expenses for these utilities as a fixed per unit price. To compensate for any losses, make-up of water and piperazine is needed to close the recycle loop. However, in a real-world plant operation, the lost piperazine is mostly

recovered in a water wash section above the absorber column. Therefore, cost of make-up is not additionally considered in our economic model. Instead, the cost of the solvent initially used in loading up the system is included in the capital costs at the design stage. Thereafter, it is assumed that the entirety of the solvent will be renewed every three years to make up for losses and degradation. This cost is included in the operational costs.

**Table 1:** Fixed prices assumed for utilities and piperazine.

Utility	Price
Piperazine	3750 \$/ton
Steam	14.5 \$/ton
Cooling water	0.0329 \$/ton
Electricity	0.06 \$/kWh

Optimal designs and corresponding operating conditions are obtained for cases for which the prices of selected parameters are changed stepwise up to a range of  $\pm 50$ . CONOPT was used as the solver for solving this nonlinear optimization problem [7]. By reoptimizing the system for each case with respect to all degrees of freedom (both design and operational), we can see how designs change and how the overall cost changes with the utility prices. The insights gained from this study are particularly useful when designing plants for geographies or markets with different applicable utility prices. Additionally, the parameter that has the greatest impact on total



**Figure 2:** Sensitivity of the optimal cost of capture to changes in utility and solvent prices

cost can be identified along with design optimizations that can mitigate the impact of its price change.

## SENSITIVITY OF OPTIMAL DESIGN TO ECONOMIC PARAMETERS

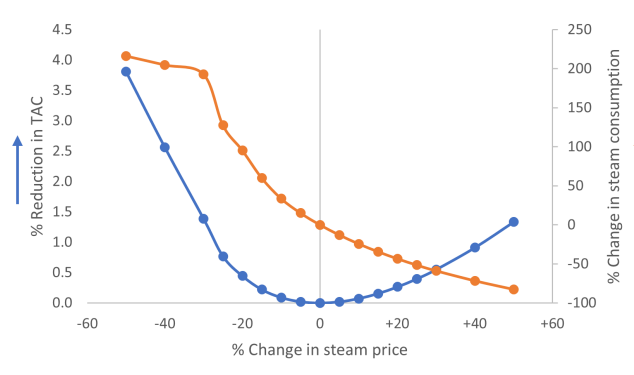
The corresponding changes in the optimal cost of capture with changing utility and solvent prices is shown in Fig 2. It is observed that the change in the price of piperazine has the most significant effect on the cost. This is due to the way solvent renewal is accounted for in our economic model, where it is a part of both capital and operational expenditures. Moreover, while there is flexibility in the usage amounts of utilities, a certain base amount of solvent needs to be used to achieve a given capture target no matter what. It is observed that electricity has the least impact on the cost of capture. This could be because our model does not consider a compressor train, so electricity is only used for running the pumps. The column pressures are not allowed to vary, therefore the only decision variable that would affect the electricity consumption is the amount of solvent circulated. While all cost curves are mostly linear, the curve representing steam becomes non-linear at lower prices, with the cost of capture becoming more sensitive to steam for price decreases beyond 20%.

To further investigate the sensitivities to price changes, the optimal solutions were compared to results from the baseline case, where price changes were implemented on the nominal solutions but without any optimization. For steam, reoptimizing the process design for price changes can lead to savings up to 3.8% (around \$1.6M yearly) compared to the baseline solution. The most significant difference is observed in the case of steam as compared to other parameters. Reoptimizing for different prices of piperazine leads to savings up to 0.2% (around \$100k yearly). However, for cooling water and electricity, reoptimizing did not lead to significant changes. The new solutions did not deviate from the original solution significantly. The nominal optimal solution is therefore robust for the changes in the price of electricity and cooling water. However, it must be noted that these are local optimal solutions, which can be dependent on initialization. Since each problem is initialized from the nominal solution, a better solution further away from the nominal case can not be ruled out.

### Sensitivity to Steam

The effect of changes in steam price on the optimal solution along with the percent reduction in the optimal Total Annualized Cost (TAC) were further investigated. Fig 3 shows the change in the steam consumption with price changes. When the price is increased by 50%, very little steam is used, with a steam heater area that is almost  $1/6^{\text{th}}$  of its original size. When the price is decreased

by 50%, the change in steam consumption is much more pronounced. This also reflects in the reduction of TAC as seen in Fig 3. For price changes between -50% and -30%, we can see a significant reduction in TAC. The steam consumption also changes sharply after the -30% price change point. This is also evident in Fig 2, where the curve representing steam shows nonlinear behavior.

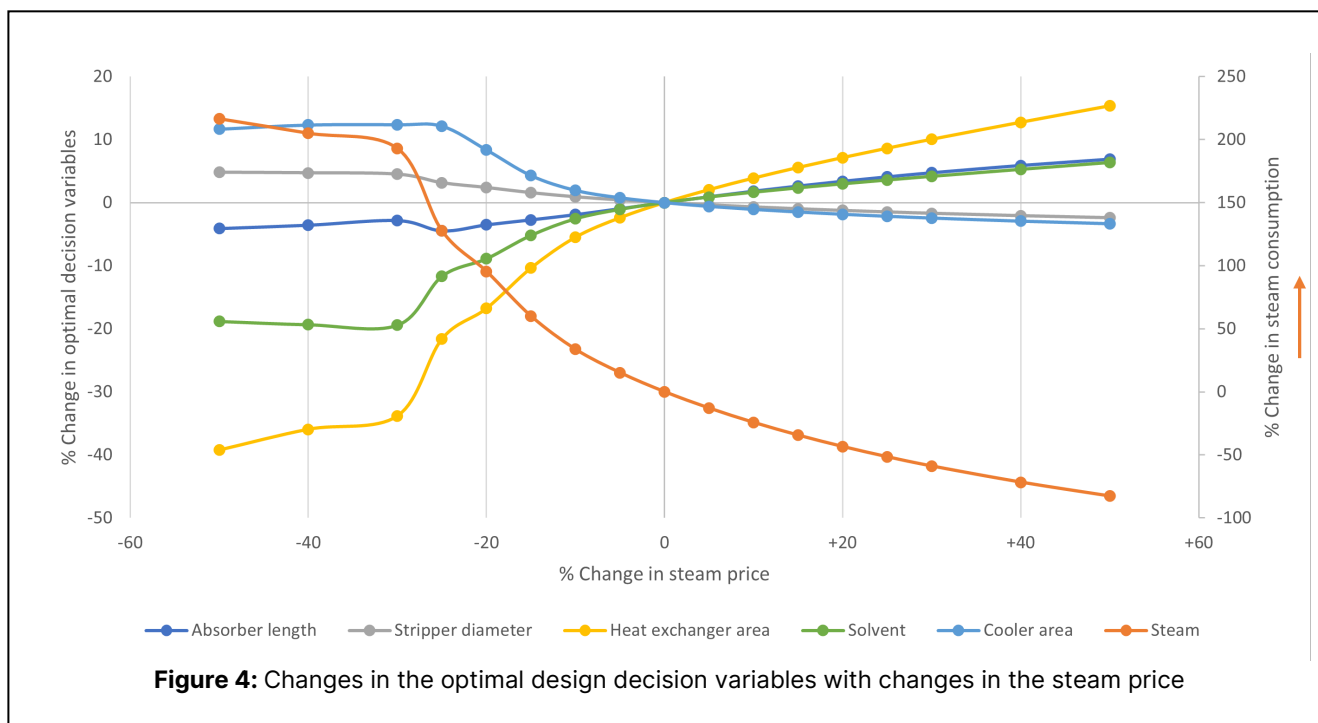


**Figure 3:** Percentage reduction in TAC by redesigning for different steam prices

The change in the values of the other decision variables given the steam price changes are shown in Fig 4. Once again, we observe the jump to a better solution at the -30% price change point. The cooler area and the stripper diameter have a similar trend with the steam consumption where they decrease with increasing steam prices. Using more steam increases the temperature of the lean solvent being pumped back to the absorber. Therefore, a bigger cooler and a higher cooling water consumption is needed when more steam is used as the system would now need more cooling. Usage of more steam also decreases the lean loading since more  $\text{CO}_2$  is flashed from the rich solvent at a higher temperature. Moreover, since more water and  $\text{CO}_2$  are flashed with increased steam usage, the gas flowrate within the stripper column increases, leading to a need for larger diameter. However, its change from the nominal solution is around 10%.

The packed absorber length, solvent flowrate and the total heat exchanger area show an opposite trend with steam consumption. Since more steam and cooling water are used, there is less of a need for the heat exchangers. Thus, at low steam prices, there is a significant (reaching up to 40%) reduction in the total heat exchanger area needed. Cooling improves the absorber performance since absorption is exothermic. As discussed above, using more steam leads to more cooling water consumption therefore, we can use a shorter column when steam prices are low. However, the change in absorber packing height is less significant. Similarly, since increased cooling water improves absorption and more steam improves stripping, less solvent may be used





to achieve the same capture target. Indeed, we can see that 20% less solvent is used for the cheapest steam price case.

## CONCLUSIONS

We leveraged the optimization capabilities of the equation-oriented model of a novel solvent-based PCC process that we had previously developed to find cost-optimal solutions considering variation in economic parameters. This was done through a series of parameter sweeps on the price of cooling water, steam, electricity, and the solvent. The prices were perturbed in steps up to a range of  $\pm 50\%$ . Piperazine had the most significant effect on the change of cost of capture while electricity had the least effect. Upon comparing the nominal solutions obtained for each price change from the baseline with those solutions obtained on re-optimizing for each price change, it was observed that only changes in steam price had a significant impact. Thus, we infer that the optimal solution for the nominal case is robust to changes in the price of cooling water, electricity and piperazine. In contrast, for steam, an optimal value with significant differences from that of the nominal solution was obtained for different prices. This change was significant for cheaper steam prices, where yearly savings of 3.8% were obtained when reoptimizing for 50% cheaper steam. For this case, a bigger cooler and a larger stripper diameter was chosen, while savings were achieved through using a shorter absorber column, lesser solvent amount, and smaller heat exchangers. So, if this plant was to be

implemented at different locations where different steam prices apply, it is worth changing the design accordingly, whereas for the other utilities this may not be a concern.

There are many more parameters in this model that could be considered uncertain. When transitioning into an optimization under uncertainty framework, it is important to pre-screen the parameters with the highest impact through sensitivity analyses as shown in this work. In the future, more parameters relating to epistemic and operational uncertainty should be investigated and the sensitivity of the total cost to said parameters should be ranked in a systematic way. Another future direction could be to assess the flexibility of a design by optimizing solely over the decision variables related to operation.

## REFERENCES

1. GT Rochelle, Y Wu, E Chen, K Akinpelumi, KB Fischer, T Gao, CT Liu, JL Selinger. Pilot plant demonstration of piperazine with the advanced flash stripper. *Int J Greenh Gas Control* 84:72-81 (2019)
2. A Lee, JH Ghose, JC Eslick, CD Laird, JD Sirola, MA Zamarripa, D Gunter, JH Shinn, AW Dowling, D Bhattacharyya, LT Biegler, AP Burgard, DC Miller. The IDAES process modeling framework and model library – Flexibility for process simulation and optimization. *J Adv Manuf Process* 3.3 e10095 (2021)
3. I Akkor, SS Iyer, J Dowdle, L Wang, CE Gounaris. Mathematical modeling and economic optimization

of a piperazine-based post-combustion carbon capture process. (To appear, 2024)

4. T Gao, JL Selinger, GT Rochelle. Demonstration of 99% CO<sub>2</sub> removal from coal flue gas by amine scrubbing. *Int J Greenh Gas Control* 83:236-244 (2019)
5. ES Rubin. Understanding the pitfalls of CCS cost estimates. *Int J Greenh Gas Control* 10:181-190 (2012)
6. F Closmann, G Rochelle, T Gao, AS Babu, M Abreu, B Drewry. Feed for piperazine with the advanced stripper™ on NGCC at Denver City, Texas. *Proceedings of the 15th Greenhouse Gas Control Technologies Conference* 15-18 (2021)
7. A Drud. Conopt: A GRG code for large sparse dynamic nonlinear optimization problems. *Math Program* 31:153-191 (1985)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# The design and operational space of syngas production via integrated direct air capture with gaseous CO<sub>2</sub> electrolysis

Hussain M. Almajed<sup>a,b</sup>, Omar J. Guerra<sup>c</sup>, Ana Somoza-Tornos<sup>d</sup>, Wilson A. Smith<sup>a,b,c</sup>, and Bri-Mathias Hodge<sup>b,c,e,f\*</sup>

<sup>a</sup> University of Colorado Boulder, Department of Chemical & Biological Engineering, Boulder, CO, USA

<sup>b</sup> University of Colorado Boulder, Renewable and Sustainable Energy Institute, Boulder, CO, USA

<sup>c</sup> National Renewable Energy Laboratory, Golden, CO, USA

<sup>d</sup> Delft University of Technology, Department of Chemical Engineering, Delft, NL

<sup>e</sup> University of Colorado Boulder, Department of Electrical, Computer and Energy Engineering, Boulder, CO, USA

<sup>f</sup> University of Colorado Boulder, Department of Applied Mathematics, Boulder, CO, USA

\* Corresponding Author: [BriMathias.Hodge@colorado.edu](mailto:BriMathias.Hodge@colorado.edu)

## ABSTRACT

The overarching goal of limiting the increase in global temperature to  $\leq 2.0^\circ\text{C}$  likely requires both decarbonization and defossilization efforts. Direct air capture (DAC) and CO<sub>2</sub> electrolysis stand out as promising technologies for capturing and utilizing atmospheric CO<sub>2</sub>. In this effort, we explore the details of designing and operating an integrated DAC-electrolysis process by examining some key parameters for economic feasibility. We evaluate the gross profit and net income to find the most appropriate capacity factor, average electricity price, syngas sale price, and CO<sub>2</sub> taxes. Additionally, we study an optimistic scenario of CO<sub>2</sub> electrolysis and perform a sensitivity analysis of the CO<sub>2</sub> capture price to elucidate the impact of design decisions on the economic feasibility. Our findings underscore the necessity of design improvements of the CO<sub>2</sub> electrolysis and DAC processes to achieve reasonable capacity factor and average electricity price limits. Notably, CO<sub>2</sub> taxes and tax credits in the order of \$400 per t-CO<sub>2</sub> or greater are essential for the economic viability of the optimistic DAC-electrolysis route, especially at competitive syngas sale prices. This study serves as a foundation for further work on designing appropriate power system models that integrate well with the presented air-to-syngas route.

**Keywords:** Carbon Dioxide Capture, Technoeconomic Analysis, Syngas, Aspen Plus, Modelling and Simulations.

## 1. INTRODUCTION

Suppressing the increase in global temperature to  $\leq 2.0^\circ\text{C}$  likely necessitates the combination of decarbonization and defossilization technologies, including point-source CO<sub>2</sub> capture (PSCC), carbon dioxide removal (CDR), CO<sub>2</sub> storage, and CO<sub>2</sub> utilization [1–7]. The path to carbon neutrality implies using defossilized carbon sources, such as biomass, sea/oceanwater, and air. There has been a significant effort on using biomass as a nature-based carbon source [8–10]. However, concerns have been raised about the effects of biomass-based technologies on crop prices, human rights, and their

competition with food lands [3,11–15].

Alternatively, one could leverage sea/oceanwater and air as defossilized carbon sources. Indeed, a key CDR method that has been gaining significant attention in the recent years is direct ocean capture (DOC) [16–20], which benefits from the CO<sub>2</sub> equilibrium between the atmosphere and the ocean as a capture step and focuses on the extraction of CO<sub>2</sub> from carbonized oceanwater. Although this technology can be promising, it is still nascent and further research is underway to understand its potential in the broader scope of defossilization efforts.

Another key CDR technology is direct air capture (DAC), which captures CO<sub>2</sub> from the atmosphere using a

chemical or a physical sorbent/solvent. DAC offers several advantages including smaller land requirements compared to nature-based CDR technologies, modular contactors, and almost no competition with food crop lands [7,21–23]. One of the advanced DAC designs utilizes a two-cycle process to capture CO<sub>2</sub> using a hydroxide-based solvent (generally, KOH) in the form of (bi)carbonates [24–27]. This technology is currently being developed and commercialized by a collaboration between Carbon Engineering and Oxy companies. However, one of the major challenges of hydroxide-based DAC is its high cost, with estimates ranging from \$94 to \$1,000 per t-CO<sub>2</sub>, and high energy consumption, which ranges from 5.50 to 8.81 GJ per t-CO<sub>2</sub> [21,24,28–30]. Most of this high energy cost originates from the elevated temperature needed (900° C) to regenerate the captured CO<sub>2</sub> via calcination [30]. Nevertheless, hydroxide-based DAC is thought to be the most scalable and cost-effective DAC technology today [21,22], making it an attractive option for obtaining defossilized carbon from air.

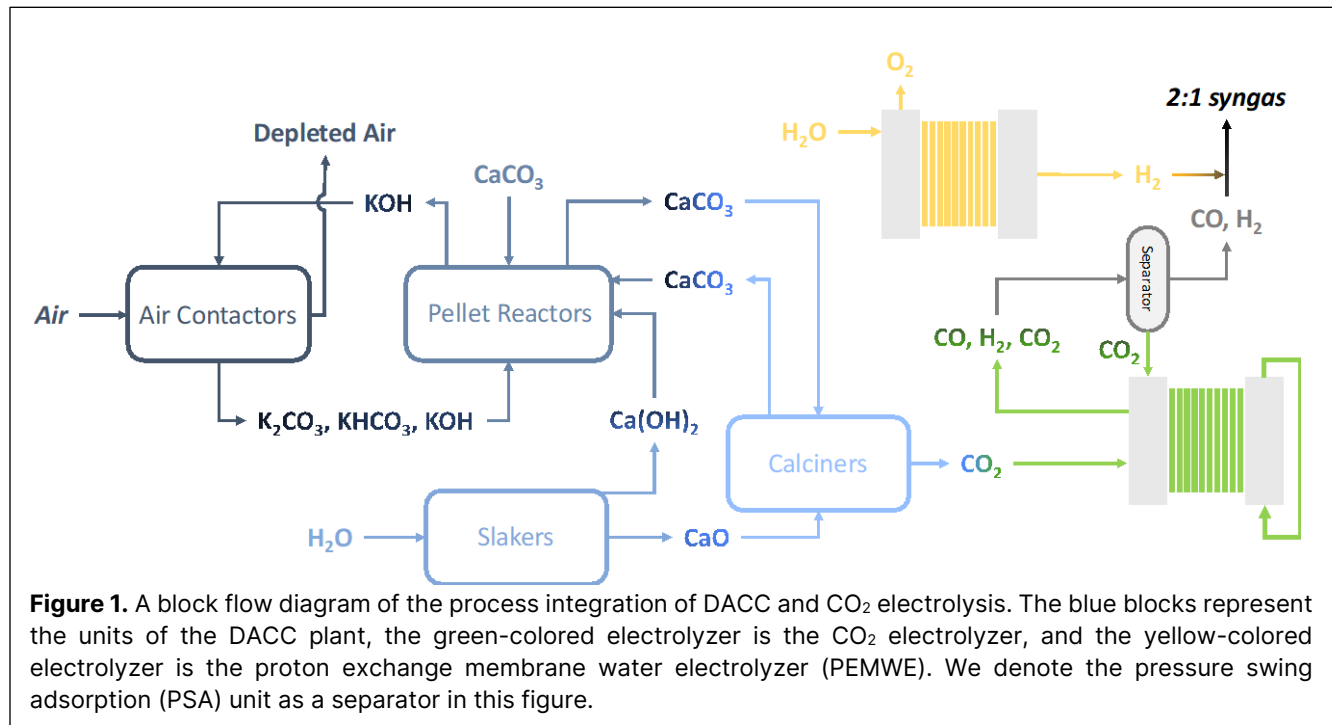
To utilize this captured CO<sub>2</sub> in a way that achieves circularity, both thermochemical and electrochemical methods can be used to convert it to more valuable products such as syngas (i.e., a mixture of H<sub>2</sub> and CO), which is a key intermediate product that can be used in the production of methanol, ethylene, jet fuel, and other high-value products [31,32]. One of the most technologically mature methods for producing syngas from a feed of CO<sub>2</sub> is reverse water gas shift (RWGS), which takes in H<sub>2</sub> as a second reactant and produces syngas at a H<sub>2</sub>:CO molar ratio of 1-2 [33,34]. The main downside of this technology is the high temperature (1,000° C) requirement that

likely necessitates the use of traditional heat sources (e.g., natural gas).

An alternative technology is low-temperature CO<sub>2</sub> electrolysis, which directly utilizes electricity to reduce CO<sub>2</sub> to CO at ambient conditions over a silver electrocatalyst in the cathode of an electrochemical cell. Although still a nascent technology, low-temperature CO<sub>2</sub> electrolysis allows the direct utilization of renewable electrons, enabling the design of a fully renewable CO<sub>2</sub> utilization method. It is worth noting that several start-up and established companies are trying to develop and upscale this technology, including Twelve [35], GIGKarasek [36], Siemens [37], Evonik [37], and Dioxide Materials [38].

The integration of DAC with low-temperature CO<sub>2</sub> electrolysis (DAC-CO<sub>2</sub>ER) can provide a promising defossilized air-to-syngas production pathway. We evaluated the techno-economics of this pathway against that of an integrated DAC-RWGS route in a previous effort [21], in which we found a potential route for DAC-CO<sub>2</sub>ER to be economically competent with more conventional routes in a future scenario. One of the main takeaways was to operate CO<sub>2</sub> electrolysis at a current density of 1,500 mA per cm<sup>2</sup> and a cell voltage of ≤ 2.00 V to enable the competition between DAC-CO<sub>2</sub>ER with DAC-RWGS and conventional methods for 2:1 syngas production.

In this work, we especially focus on the economic feasibility of the DAC-CO<sub>2</sub>ER route that will allow it to be continuously operated at a large scale. We study the influence of capacity factor, average electricity price, syngas sale price, CO<sub>2</sub> taxes, and CO<sub>2</sub> tax credits on the process economics. We first estimate the design and operational limits of a baseline CO<sub>2</sub> electrolysis system by



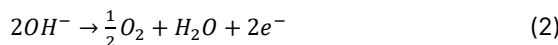
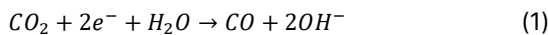
varying the capacity factor, average electricity price, and syngas sale price. We then look at an optimistic case of CO<sub>2</sub> electrolysis to understand the effect of cell voltage and current density optimization on the design and operational limits. To expand on this, we perform a sensitivity analysis of the CO<sub>2</sub> capture cost on the optimistic-scenario limits to also consider further improvements of the DAC design. We then estimate the natural gas (NG) based market syngas sale price with CO<sub>2</sub> taxes, enabling a comparison with potential future conventional syngas prices. Finally, we add CO<sub>2</sub> tax credits to the DAC-electrolysis system to search for economically feasible regions of this defossilized syngas production pathway for design and operating conditions.

## 2. METHODOLOGY

### 2.1. Process description

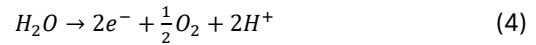
Figure 1 envisions a block flow diagram of the presented defossilized syngas production pathway. The DAC technology considered here uses KOH to capture CO<sub>2</sub> from the atmosphere in the form of K<sub>2</sub>CO<sub>3</sub> and KHCO<sub>3</sub> in an air contactor. The captured liquid solution is then sent to a set of pellet reactors along with a Ca(OH)<sub>2</sub> slurry to regenerate the liquid KOH solvent and produce CaCO<sub>3</sub> pellets. The pellets are then sent to calciners where they are dissociated into CaO and CO<sub>2</sub> at a temperature of 900° C. The CaO is then mixed with H<sub>2</sub>O in a set of slakers to re-produce a Ca(OH)<sub>2</sub> slurry, whereas the gaseous product is dehydrated to produce a purified CO<sub>2</sub> stream. This whole process is continuous, allowing for the capture of CO<sub>2</sub> in the form of potassium (bi)carbonates as well as the regeneration of liquid KOH and gaseous CO<sub>2</sub>.

The concentrated gaseous CO<sub>2</sub> can then be sent to a CO<sub>2</sub> electrolyzer where it is electrochemically reduced to CO using two electrons and two protons, Eq. (1). It is worth noting that the local pH near the catalyst surface determines the proton source, which also determines the reaction by-product (i.e., OH<sup>-</sup> for neutral-to-alkaline local pH or H<sub>2</sub>O for acidic local pH) [39]. In this work, we assume a high local pH, which corresponds to H<sub>2</sub>O being used as the proton source. On the other side, the anode electrochemical reaction is assumed to be the oxygen evolution reaction (OER), Eq. (2), which commonly occurs on a Nickel-based or an Iridium-based electrocatalyst.



The CO<sub>2</sub> electrolyzer is not only producing CO, but also H<sub>2</sub> via the competing hydrogen evolution reaction (HER), Eq. (3). However, the CO<sub>2</sub> electrolyzer is generally optimized to increase the CO selectivity and decrease the H<sub>2</sub> selectivity in the product stream. Thus, we supply additional H<sub>2</sub> from a proton exchange membrane water

electrolyzer (PEMWE) that oxidizes H<sub>2</sub>O to produce O<sub>2</sub> and protons (H<sup>+</sup>), Eq. (4). The protons then pass a cation exchange membrane (CEM) as they travel to the cathode side to be electrochemically reduced to H<sub>2</sub> via HER, Eq. (3).





### 2.2. Process models

A hydroxide-based DAC plant was modeled in Aspen Plus based off of Keith et al.'s design [24]. This model provides the mass and energy balances required to estimate the equipment and operational costs of the plant. In addition, we use our own CO<sub>2</sub> electrolysis model that calculates the mass and energy balances of a specified electrolysis system. For further details about the process models, we refer the reader to our previous publication [21].

We design the CO<sub>2</sub> electrolysis system based off of Wen and Ren et al.'s experimental results that achieved 90% Faradaic efficiency of CO (FE<sub>CO</sub>) at 612 mA per cm<sup>2</sup> and 3.3 V (energy efficiency (EE) ≈ 40%) in a 400-cm<sup>2</sup>, 4-cell electrolyzer stack [40]. These performance values are used in our baseline scenario (Table 1). We additionally consider an optimistic scenario in which we assume achievement of a current density and a cell voltage of 1,500 mA per cm<sup>2</sup> and 2.0 V (EE ≈ 67%), respectively. Table 1 summarizes these assumptions.

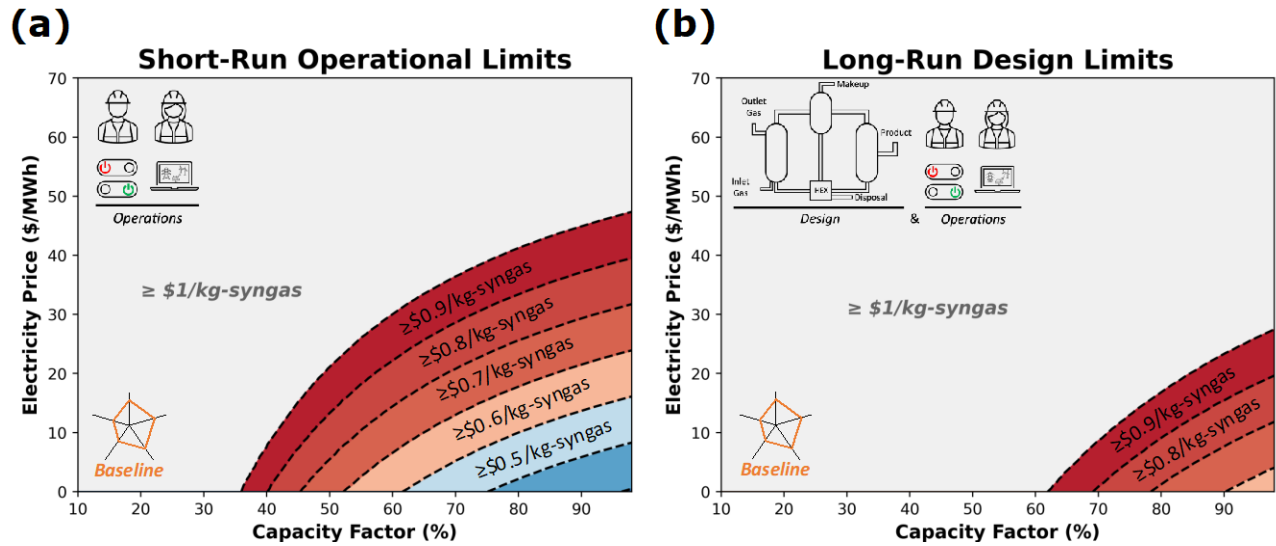
**Table 1.** Key baseline and optimistic assumptions of CO<sub>2</sub> electrolysis.

Metric	Baseline	Optimistic
Symbol		
Total Current Density (mA per cm <sup>2</sup> )	612	1,500
Cell Voltage (V)	3.3	2.0
FE <sub>CO</sub>	90%	90%
Single-Pass Conversion	27%	27%

### 2.3. Techno-economic model

Our techno-economic model was developed to take the results of our process models as inputs and estimate the equipment and operational costs according to Towler and Sinnott's methodology [41]. In addition, we estimate the gross profit, Eq. (5), and the net income, Eq. (6), while varying the capacity factor and average electricity price at a constant syngas price. The gross profit is used as an indicator of the system operations whereas the net income is used as an indicator of both the design and operation of the system. In Eq. (5) and (6), OPEX, CAPEX, and CRF refer to operational cost, capital cost, and capital recovery factor, respectively. The CRF, Eq. (7), is used





**Figure 2.** Capacity factor and average electricity price limits for (a) positive gross profit and (b) positive net income at different syngas prices of the baseline case (cell voltage = 3.3 V, current density = 612 mA per cm<sup>2</sup>, FE<sub>CO</sub> = 90%, CO<sub>2</sub> single-pass conversion = 27%).

to annualize the capital cost of the system, where  $i$  is the interest rate and  $N$  is the plant lifetime. For key assumptions and further details about our techno-economic model, we refer the reader to the SI.

$$\text{Gross Profit} = \text{Revenue} - \text{OPEX} \quad (5)$$

$$\text{Annual Net Income} = \text{Gross Profit} - \text{CAPEX} \cdot \text{CRF} \quad (6)$$

$$\text{CRF} = \frac{i(1+i)^N}{(1+i)^N - 1} \quad (7)$$

We use the techno-economic model outputs to define realistic targets for future renewably powered CO<sub>2</sub> electrolysis designs and techno-economic evaluations, especially in the context of renewably driven defossilized air-to-syngas pathways.

### 3. RESULTS & DISCUSSION

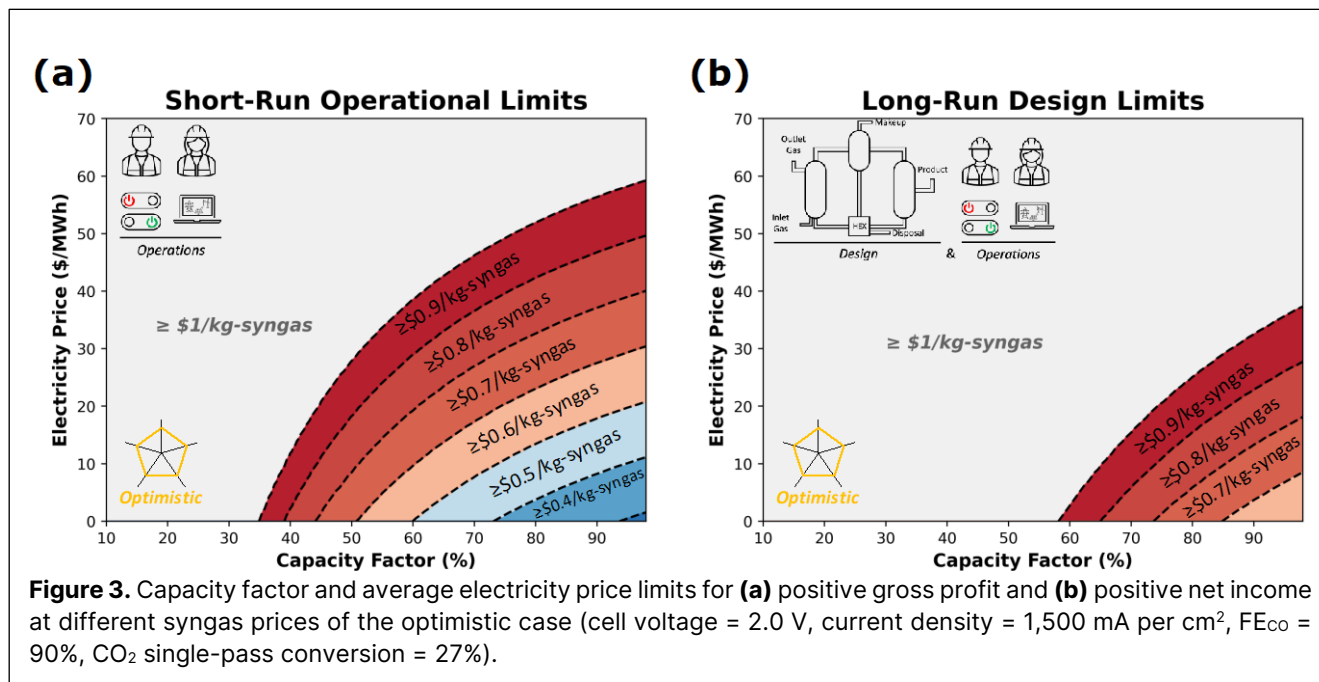
Two of the key assumed parameters in techno-economic assessment (TEA) calculations in electrolysis are the capacity factor and the average electricity price. In the past, a capacity factor of  $\geq 90\%$  and an average electricity price in the range of \$20-60 per MWh have been used in the CO<sub>2</sub> electrolysis field [21,42–47]. Such values are highly optimistic for interactions with a highly renewable decarbonized power system and might unrealistically benefit the economics of CO<sub>2</sub> electrolysis. To provide more realistic assumptions, a more-detailed power system model is needed to fill this gap. However, before this is pursued, it is important to understand the combinations of the capacity factor, average electricity price, syngas sale price, and CO<sub>2</sub> taxes that would allow the whole route to generate positive gross profit and positive net income.

We consider two cases in this effort. In the first

case, we focus on the operational limits of the presented DAC-electrolysis route (Fig. 1) by estimating the gross profit while varying the capacity factor and the average electricity price at syngas sale prices of \$0.3-1.0 per kg-syngas. We then identify the minimum capacity factor and average electricity price at each syngas price that allow the gross profit to stay positive. In the second case, we take the plant's design into consideration as well, meaning that we not only focus on the short-term economic decisions but the long-term ones as well. To accomplish this, we estimate the same parameters that enable the net income to stay positive at the different syngas sale prices. We conduct the same analysis for four scenarios: 1) Baseline without CO<sub>2</sub> tax credits (Fig. 2a and 2b), 2) Optimistic without CO<sub>2</sub> tax credits (Fig. 3a and 3b), 3) Baseline with a CO<sub>2</sub> tax credit of \$130 per t-CO<sub>2</sub> (Fig. 5a and 5b), and 4) Optimistic with a CO<sub>2</sub> tax credit of \$130 per t-CO<sub>2</sub> (Fig. 5c and 5d).

#### 3.1. Design and operational economic limits of the baseline DAC-CO<sub>2</sub>ER route without CO<sub>2</sub> tax credits

Figures 2a and 2b show the relationship between the capacity factor and the average electricity price at different syngas prices that would allow the gross profit and net income to be positive, respectively. We observe that a positive gross profit (Fig. 2a) at lower syngas prices requires a high-capacity factor and a low average electricity price. As the assumed syngas sale price increases from \$0.3 to \$1.0 per kg-syngas, more flexibility in the capacity factor and the average electricity price can be obtained, while remaining economically viable. For instance, at an assumed syngas sale price of \$0.7 per kg-



syngas, a capacity factor of  $\geq 95\%$  is required at an average electricity price of \$23 per MWh to generate a positive gross profit. However, at a higher assumed syngas sale price of \$1.0 per kg-syngas, more operational flexibility can be provided (e.g., capacity factor of 50%) at the same average electricity price (\$23 per MWh), while still generating a positive gross profit. This finding demonstrates the importance of capacity factor and average electricity price assumptions in determining the short-term operational economic feasibility of the investigated DAC-electrolysis route.

At longer timescales, the cost associated with the design of the plant (e.g., equipment, instrumentations) must be considered. We observe a red shift of the color-map that requires higher syngas prices at higher capacity factors and lower average electricity prices (Fig. 2b). This is caused by the addition of the annualized capital costs, which adds an economic restriction on the design and operations of the plant that must be considered and minimized at an early stage. However, it is worth noting that, in practice, this restriction does not impact the operational decisions but rather the design decisions.

Indeed, reducing the capital cost could significantly help achieve more flexible operations of the electrolysis unit such that the DAC-electrolysis route stays economically feasible throughout the broad range of dynamic electricity prices anticipated in the future [48]. Additionally, benefitting from incentives [49] that partially or fully pay off the capital cost of the plant could significantly hasten the deployment of these defossilized pathways, especially when considering the current average syngas market price (\$0.40 per kg-syngas, see the discussion in section 3.4 and calculation in the SI). However, it is worthwhile to re-iterate that this restriction does not

influence the operational decisions of the plant.

### 3.2. Design and operational economic limits of the optimistic DAC-CO<sub>2</sub>ER route without CO<sub>2</sub> tax credits

Improving the design of CO<sub>2</sub> electrolyzers such that they achieve high current densities ( $\geq 1,500$  mA per cm<sup>2</sup>) at low cell voltages ( $\leq 2.0$  V; EE  $\geq 67\%$ ) is still in progress [21,50]. Defining research targets that consider the economics of such pathways would be helpful to CO<sub>2</sub> electrolysis researchers at this developmental stage. Figure 3 shows the same relationships from Figure 2, however considering an optimistic CO<sub>2</sub> electrolysis performance that achieves 1,500 mA per cm<sup>2</sup> at 2.0 V (EE  $\approx 67\%$ ).

We find the average electricity price to gain more flexibility when considering the optimistic CO<sub>2</sub> electrolysis performance in our gross profit calculations (Fig. 3a). For instance, at an assumed syngas sale price of \$1.0 per kg-syngas, the maximum average electricity prices for the baseline and optimistic scenarios that enable positive gross profits at a capacity factor of 70% are approximately \$37 and \$46 per MWh, respectively. This finding is mainly due to the large influence of electricity costs on the economics of DAC-electrolysis routes, as driven by the cell voltage (or energy efficiency) [21,50].

In addition, we find the capacity factor limits of maintaining a positive gross profit to also change as the cell voltage and current density are optimized (Fig. 3a). At an average electricity price of \$30 per MWh and an assumed syngas sale price of \$1.0 per kg-syngas, the capacity factor limits of the baseline and optimistic scenarios that allow positive gross profits are estimated to be 60% and 52%, respectively. The higher flexibility with operating at lower capacity factors while still generating a

positive gross profit in the optimistic scenario is also driven by the lower electricity costs, which originate from achieving a lower cell voltage of 2.0 V.

Similarly, Figure 3b shows the same general results—i.e., higher flexibility of the average electricity price and the capacity factor limits to generating a positive net income. However, we observe a slightly higher flexibility in the capacity factor, resulting from the reduction of the cell voltage that reduces the power of the electrolyzer; and thus, the electrolyzer's capital cost. Therefore, higher flexibility in the capacity factor can be obtained in the optimistic scenario for operating with a positive net income.

Although assuming an optimistic scenario provides more flexibility in the average electricity price and capacity factor, high syngas prices of  $\geq \$0.6$  per kg-syngas are still needed to generate positive gross profit and net income (Fig. 3). In the following sections, we will explore the effect of CO<sub>2</sub> capture costs, CO<sub>2</sub> taxation on emitting technologies, and CO<sub>2</sub> tax credits for capture technologies in order for the current projections of the integrated process to make economic sense.

### 3.3. Design limits at different CO<sub>2</sub> capture costs without CO<sub>2</sub> tax credits

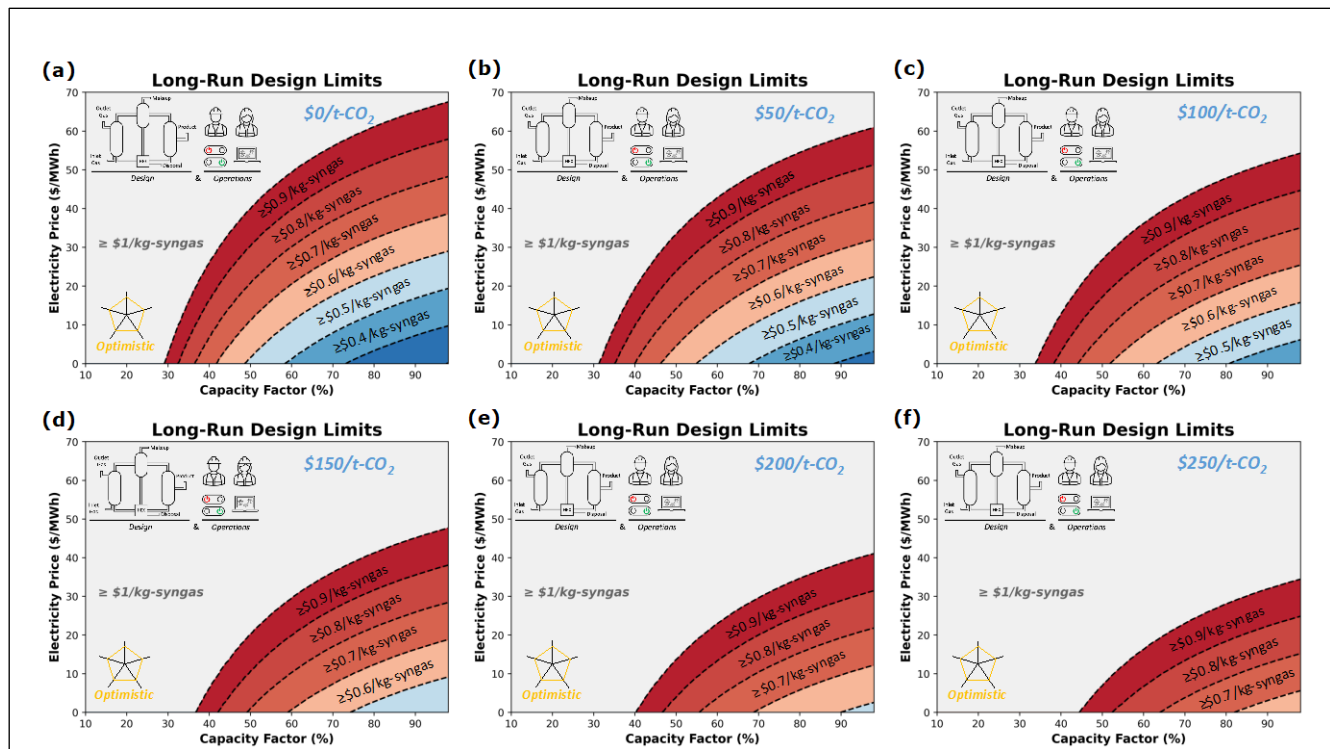
Sourcing the CO<sub>2</sub> from flue gases or from other non-fossil sources could have a significant influence on the

economic feasibility of the presented syngas production route. Thus, we estimated the long-run design limits at different capture costs ranging from \$0 to \$250 per t-CO<sub>2</sub>, as shown in Figure 4.

In general, as the CO<sub>2</sub> capture cost increases, the plot shifts to the red region. At a CO<sub>2</sub> capture cost of \$100 per t-CO<sub>2</sub>, consistent with the U.S. department of energy (DOE) carbon shot goals, we find the average electricity price to be \$19 per MWh at a capacity factor of 80% and a syngas sale price of \$0.7 per kg-syngas. Reducing the CO<sub>2</sub> capture cost to \$50 per t-CO<sub>2</sub>, consistent with sourcing CO<sub>2</sub> from PSCC, allows more flexibility in the average electricity price ( $\leq \$25$  per MWh) at the same capacity factor and syngas sale price. This result demonstrates the importance of reducing the CO<sub>2</sub> capture cost to less than \$100 per t-CO<sub>2</sub> or sourcing the CO<sub>2</sub> from an alternative source that would provide the carbon at a cheaper price. However, herein, we continue to focus on DAC to present an environmental process that allows the production of 2:1 syngas from air.

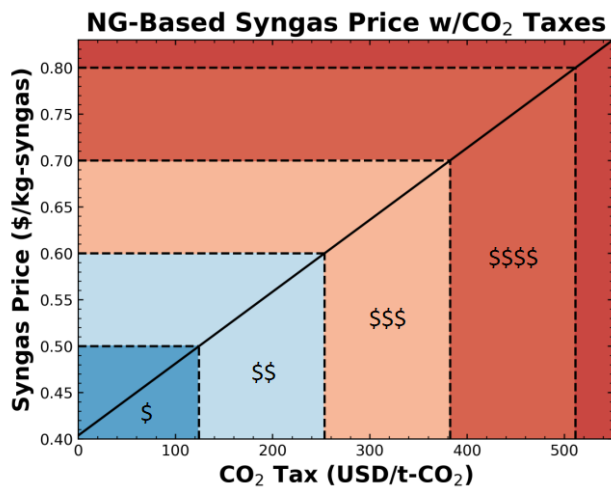
### 3.4. Conventional syngas market price with CO<sub>2</sub> taxation

The market syngas price is an important metric to compare against when evaluating emerging syngas production pathways. Previous literature have cited or estimated different syngas prices from natural gas (NG) feedstock processes, ranging from \$0.03 to \$0.74 per



**Figure 4. (a-f)** Capacity factor and average electricity price limits for positive net income at different syngas prices of the optimistic case (cell voltage = 2.0 V, current density = 1,500 mA per cm<sup>2</sup>, FE<sub>CO</sub> = 90%, CO<sub>2</sub> single-pass conversion = 27%) and at various CO<sub>2</sub> capture prices ranging from \$0 to \$250 per t-CO<sub>2</sub>.

kg-syngas [21,33,44,51,52]. We use the average NG-based syngas price of these estimates as the market price in our analysis. We assume a NG-based plant that is composed of dry methane reforming (DMR) for 1:1 syngas production and steam methane reforming (SMR) for additional H<sub>2</sub> supply to produce 2:1 H<sub>2</sub>:CO syngas, according to the design of Rezaei and Dzuryk [33]. Figure 5 shows the re-calculated NG-based syngas price with CO<sub>2</sub> taxes of \$0-550 per t-CO<sub>2</sub>. We assume a 90% capacity factor to estimate the process and energy-associated CO<sub>2</sub> emissions from the SMR-DMR pathway. Note that we are not accounting for any emissions outside the scope of the SMR-DMR process (e.g., natural gas processing).



**Figure 5.** NG-based syngas price as a function of CO<sub>2</sub> taxes. The color codes from blue (\$) to orange (\$\$\$\$) correspond to the ranges  $\leq \$0.5$ ,  $\leq \$0.6$ ,  $\leq \$0.7$ , and  $\leq \$0.8$  per kg-syngas cases, respectively. Note that we are not performing a full life-cycle assessment here; we are only accounting for the emissions from the DMR-SMR process, ignoring the emissions associated with raw materials' processing.

The linear relationship shows that the syngas price would increase to \$0.79 per kg-syngas with a CO<sub>2</sub> tax of \$500 per t-CO<sub>2</sub>. Although this would give the DAC-electrolysis pathway more flexibility in the capacity factor and average electricity price choices, it suggests a CO<sub>2</sub> tax that is three times higher than the highest CO<sub>2</sub> tax implemented today (\$155 per t-CO<sub>2</sub> in Uruguay [53]). Thus, it is likely not realistic to consider reaching a CO<sub>2</sub> tax rate of \$500 per t-CO<sub>2</sub> by mid-century.

However, considering a CO<sub>2</sub> tax rate of \$155 per t-CO<sub>2</sub>, the re-calculated NG-based syngas price increases from \$0.40 to approximately \$0.53 per kg-syngas (Fig. 5). At this price, the gross profit of the baseline and optimistic cases can be positive, although at restrictive capacity factor and average electricity price values, whereas the net income can only be negative. The only

exception to this conclusion is if the CO<sub>2</sub> capture costs dropped to  $\leq \$100$  per t-CO<sub>2</sub>, assuming no tax credits are provided to CO<sub>2</sub> capture (note that we will explore CO<sub>2</sub> tax credits in the next section). This finding, along with our previous results [21], suggest that a capital cost reduction of electrolyzers and air contactors would be necessary to make the economics of DAC-electrolysis viable. In addition, deployment incentives that partially or fully pay the capital expenses of building DAC-electrolysis plants could make a difference in deploying such emerging technologies and allowing them to compete with existing NG-based syngas production methods. In parallel, CO<sub>2</sub> tax credits could enormously help the net income to become positive by paying off some of the annual capital cost payments.

### 3.5. Design and operational economic limits of the baseline and optimistic DAC-CO<sub>2</sub>ER route with a CO<sub>2</sub> tax credit of \$130 per t-CO<sub>2</sub>

CO<sub>2</sub> tax credits can be paid to CO<sub>2</sub> capture plants whether they capture the gas from point sources or from air. In the U.S., the 45Q tax credit code [49] pays DAC plants a tax credit of \$180 per tonne of captured and geologically stored CO<sub>2</sub>, and \$130 per tonne of captured and used CO<sub>2</sub>. In our analysis, we therefore use a CO<sub>2</sub> tax credit of \$130 per t-CO<sub>2</sub> for the captured CO<sub>2</sub> by DAC to be used in the production of 2:1 syngas.

Figure 6 shows the capacity factor and average electricity price limits of the baseline (Fig. 6a and 6b) and optimistic (Fig. 6c and 6d) CO<sub>2</sub> electrolyzer cases with a CO<sub>2</sub> tax credit of \$130 per t-CO<sub>2</sub>. The optimistic case enables more flexibility in the capacity factor and average electricity price at all syngas prices. For example, a price of \$0.70 per kg-syngas at a capacity factor of 80% can generate a positive gross profit at average electricity prices of  $\leq \$32$  and  $\leq \$40$  per MWh for the baseline and optimistic cases, respectively. For the same case and same conditions, the maximum average electricity price lowers to \$7 and \$13 per MWh for the two cases, respectively, to generate a positive net income. These results clarify that a CO<sub>2</sub> tax of \$155 per t-CO<sub>2</sub> and a CO<sub>2</sub> tax credit of \$130 per t-CO<sub>2</sub> are insufficient for the considered DAC-electrolysis route to be economically feasible, especially when considering annual capital cost payments.

For the optimistic case to generate positive gross profit and net income at a reasonable average electricity price and competitive syngas sale price, two conditions must be met. First, a higher tax rate implemented on NG-based syngas production is necessary. For instance, a CO<sub>2</sub> tax rate of \$383 per t-CO<sub>2</sub> would increase the syngas market sale price to \$0.70 per kg-syngas, as shown in Figure 5. This rate is more than double the highest CO<sub>2</sub> tax rate today—\$155 per t-CO<sub>2</sub>—but is necessary from a business standpoint to strengthen the economic



competition of the presented syngas production pathway.

Second, the CO<sub>2</sub> tax credits must be increased. Assuming an equal CO<sub>2</sub> tax credit of \$383 per t-CO<sub>2</sub>, the average electricity price limit of the optimistic scenario at a syngas price of \$0.70 per kg can increase from \$4 to \$47 per MWh to generate a positive net income, while staying at a 90% capacity factor limit (Fig. S.1b). At this average electricity price and scenario, there will be more freedom in the operational decisions because the gross profit can stay positive even at a low-capacity factor of 60%, albeit at an average electricity price of \$18 per MWh (Fig. S.1a). Although discouraged from a design perspective, this scenario would allow the continued and economic operation of the plant.

To put these results into context, NREL's annual technology baseline (ATB) analysis predicts a levelized cost of energy (LCOE) of utility-scale solar PV with battery storage in the range of \$35-90 per MWh in 2050 [54]. The analysis presented in this section highlights the economic and policy challenges faced by the presented DAC-electrolysis pathway. In other words, it will be difficult for the DAC-electrolysis system to compete with conventional syngas production methods without a CO<sub>2</sub> emission tax rate and a CO<sub>2</sub> capture tax credit on the order of \$400 per t-CO<sub>2</sub>. Even in this scenario, the plant must operate most of the year at a capacity factor of  $\geq 80\%$  and the average electricity price must be in the lower range of the NREL ATB's LCOE predictions (i.e.,  $\leq$  \$40 per MWh). Thus, further design and energy efficiency improvements as well as policy incentives are necessary before this pathway can be commercialized at scale.

#### 4. SUMMARY AND FUTURE TARGETS

In this work, we focused on the capacity factor and average electricity price limits at several syngas prices that enable the gross profit and net income of a DAC-electrolysis route to be positive. We considered hydroxide-based DAC as well as gaseous CO<sub>2</sub> and liquid PEM H<sub>2</sub>O electrolysis systems to produce 2:1 H<sub>2</sub>:CO syngas. We additionally considered an SMR-DMR conventional route to re-calculate the syngas price after the addition of CO<sub>2</sub> taxes. We analyzed baseline and optimistic scenarios of CO<sub>2</sub> electrolysis and accounted for a tax credit payment of \$130 per t-CO<sub>2</sub>.

Our findings suggest that deploying the presented DAC-electrolysis route for syngas production will be economically challenging, even after considering current CO<sub>2</sub> taxes and tax credits. Indeed, our results show that the very best-case scenario—which considers improved CO<sub>2</sub> electrolysis performance, a CO<sub>2</sub> tax rate and tax credit of \$383 per t-CO<sub>2</sub>, and a capacity factor of  $\geq 80\%$ —still requires an average electricity price of \$40 per MWh to generate both positive gross profit and net income. This

price is in the lower range of the 2050 LCOE predicted average electricity prices by NREL's ATB analysis for a utility-scale solar PV with battery storage, and it is about 33% lower than the average 2023 wholesale electricity prices in U.S. markets [55]. Thus, several targets must be pursued before this pathway can be deployed as a defossilized syngas production route.

For CO<sub>2</sub> electrolysis, the main challenge is to achieve high current densities ( $\geq 1,500$  mA per cm<sup>2</sup>) at low cell voltages ( $\leq 2.00$  V) to lower the capital cost while improving the energy efficiency of the process. However, special attention must be paid to the durability and stability of the electrolysis system as achieving the performance metric without stability would not allow industrial deployment at the high capacity factors required for profitability [56]. In addition, flexible operation is still an open problem in the low-temperature CO<sub>2</sub> electrolysis field, requiring experimental tests of interrupted electrolysis operations for long durations.

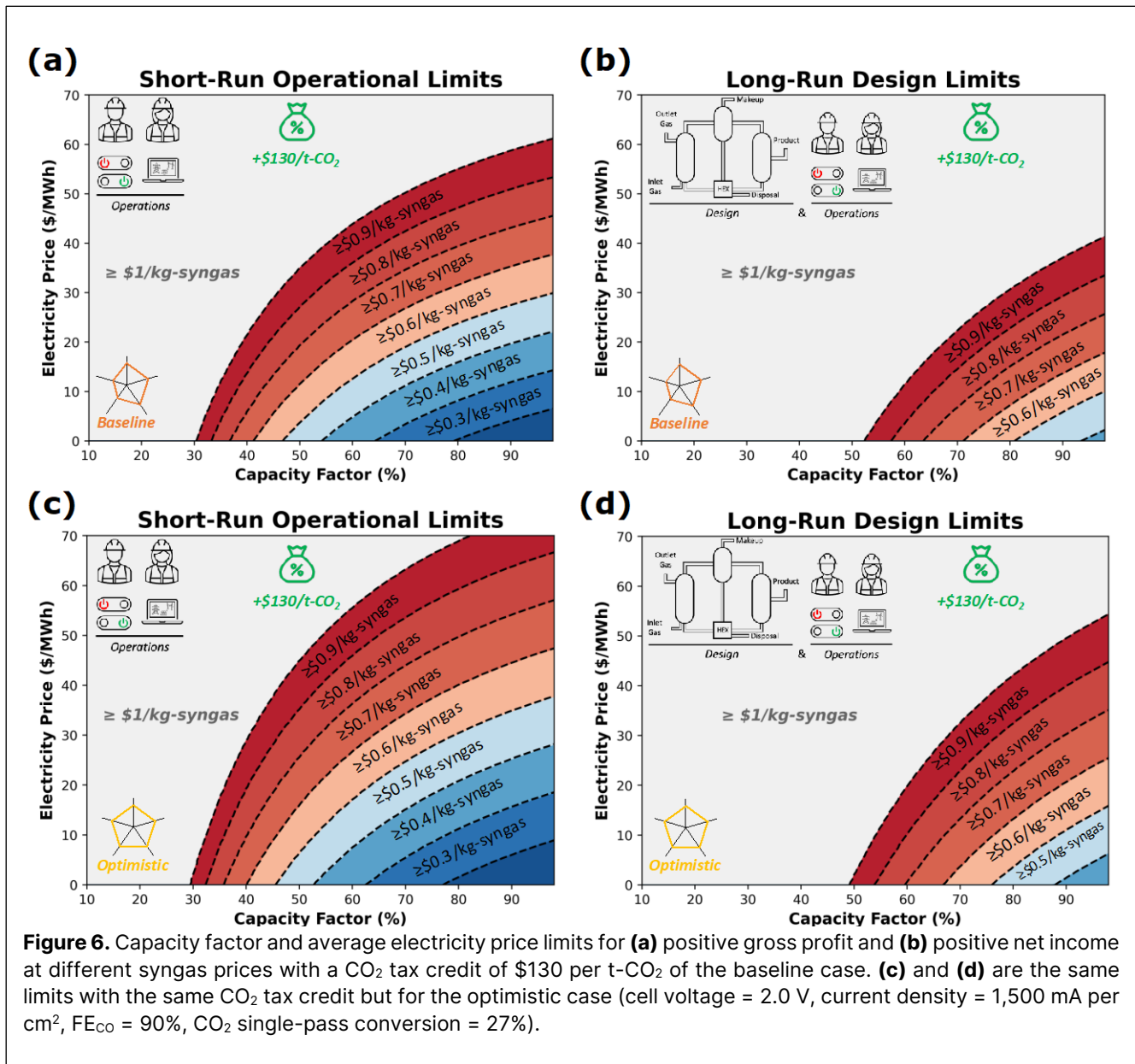
Hydroxide-based DAC plant designs must also be optimized to lower the energy consumption and capital costs. Specifically, the air contactor capital cost needs to be reduced to maximize the net income of the entire integrated route. Alternative low-temperature regeneration methods could help reduce the capital cost as there would not be any need to using heat exchanger networks, which accounts for approximately 21% of the total capital cost of the presented route [21].

Outside of technology development, high carbon taxes on CO<sub>2</sub>-emitting technologies and high carbon tax credits for CO<sub>2</sub>-capturing technologies would likely be necessary. Our results highlighted the importance of these policies in the shift from fossil-based to fossil-free production of syngas. For the defossilized air-to-syngas route presented here to be economically viable, CO<sub>2</sub> taxes and tax credit on the order of \$400 per t-CO<sub>2</sub> are required. These numbers are unquestionably too high. However, early and fast deployment is necessary to enable learning-by-doing and economies-of-scale to help reduce the total cost of these emerging defossilized pathways; and thus, their required CO<sub>2</sub> taxes and tax credits.

Future work should explore several avenues, one of which is the integration of a detailed power system model that considers the dynamics of renewable electricity generation and prices as well as energy storage. In this avenue, a comparison with wholesale electricity supply should be considered to understand the best-case scenario for the DAC-electrolysis route to be economically feasible. In addition, alternative air-to-syngas routes should be explored with a similar methodology as presented here. Indeed, the most feasible integrated air-to-syngas design is still unknown, motivating future works to further explore this research path.

Finally, it is worthwhile to note that the present





**Figure 6.** Capacity factor and average electricity price limits for (a) positive gross profit and (b) positive net income at different syngas prices with a CO<sub>2</sub> tax credit of \$130 per t-CO<sub>2</sub> of the baseline case. (c) and (d) are the same limits with the same CO<sub>2</sub> tax credit but for the optimistic case (cell voltage = 2.0 V, current density = 1,500 mA per cm<sup>2</sup>, FE<sub>CO</sub> = 90%, CO<sub>2</sub> single-pass conversion = 27%).

study particularly considered sourcing CO<sub>2</sub> from air. However, there are other sources of CO<sub>2</sub>, such as the ocean, flue gases, and CO<sub>2</sub> process emissions, which were not explored here. Specifically, the latter two sources are likely to enhance the process economics of the whole integrated pathway, potentially allowing stronger competition with conventional NG-based syngas production methods. However, one should be careful with environmental concerns of upstream processes that generate these point-source gases to assure the design of a sustainable, circular, and environmental process for syngas production.

## DIGITAL SUPPLEMENTARY MATERIAL

The supplementary information file includes details

about the techno-economic assessment calculations and supplementary figures.

## ACKNOWLEDGEMENTS

H.A. acknowledges support from the Saudi ministry of education, sponsored by the Saudi Arabian Cultural Mission (SACM) in the United States. This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the US Department of Energy (DOE) under contract no. DE-AC36-08GO28308. This work was supported by the Laboratory Directed Research and Development (LDRD) program at NREL. The views expressed in the article do not necessarily represent the views of the DOE, the US government, or the Saudi government.

## REFERENCES

1. International Energy Agency (September 19) Putting CO<sub>2</sub> to Use, France, IEA Publications.
2. IPCC (2021) Climate Change 2021: The Physical Science Basis: Summary for Policymakers, Cambridge, UK and New York, NY, USA, Cambridge University Press.
3. IPCC (2022) Climate Change 2022: Mitigation of Climate Change: Summary for Policymakers, Cambridge, UK and New York, NY, USA, Cambridge University Press.
4. Esqué A, Mitchell A, Rastogi K, et al. (2022) Decarbonizing the Aviation Sector: Making Net Zero Aviation Possible.
5. Darton RC, Yang A (2018) Removing Carbon Dioxide from the Atmosphere – Assessing the Technologies. *Chemical Engineering Transactions* 69: 91–96.
6. Biniek K, Henderson K, Rogers M, et al. (2020) Driving CO<sub>2</sub> Emissions to Zero (and beyond) with Carbon Capture, Use, and Storage. *McKinsey & Company* 8.
7. Beuttler C, Charles L, Wurzbacher J (2019) The Role of Direct Air Capture in Mitigation of Anthropogenic Greenhouse Gas Emissions. *Frontiers in Climate* 1: 10.
8. Ben-Iwo J, Manovic V, Longhurst P (2016) Biomass resources and biofuels potential for the production of transportation fuels in Nigeria. *Renewable and Sustainable Energy Reviews* 63: 172–192.
9. Tursi A (2019) A review on biomass: importance, chemistry, classification, and conversion. *Biofuel Research Journal* 6: 962–979.
10. Saxena RC, Adhikari DK, Goyal HB (2009) Biomass-based energy fuel through biochemical routes: A review. *Renewable and Sustainable Energy Reviews* 13: 167–178.
11. Rosa L, Sanchez DL, Mazzotti M (2021) Assessment of carbon dioxide removal potential via BECCS in a carbon-neutral Europe. *Energy Environ Sci* 14: 3086–3097.
12. Roe S, Streck C, Obersteiner M, et al. (2019) Contribution of the land sector to a 1.5 °C world. *Nat Clim Chang* 9: 817–828.
13. Uden S, Dargusch P, Greig C (2021) Cutting through the noise on negative emissions. *Joule* 5: 1956–1970.
14. Fuhrman J, McJeon H, Patel P, et al. (2020) Food–energy–water implications of negative emissions technologies in a +1.5 °C future. *Nat Clim Chang* 10: 920–927.
15. Dooley K, Kartha S (2018) Land-based negative emissions: risks for climate mitigation and impacts on sustainable development. *Int Environ Agreements* 18: 79–98.
16. Aleta P, Refaie A, Afshari M, et al. (2023) Direct ocean capture: the emergence of electrochemical processes for oceanic carbon removal. *Energy & Environmental Science* 16: 4944–4967.
17. de Lannoy C-F, Eisaman MD, Jose A, et al. (2018) Indirect ocean capture of atmospheric CO<sub>2</sub>: Part I. Prototype of a negative emissions technology. *International Journal of Greenhouse Gas Control* 70: 243–253.
18. Eisaman MD, Rivest JLB, Karnitz SD, et al. (2018) Indirect ocean capture of atmospheric CO<sub>2</sub>: Part II. Understanding the cost of negative emissions. *International Journal of Greenhouse Gas Control* 70: 254–261.
19. Yan L, Bao J, Shao Y, et al. (2022) An Electrochemical Hydrogen-Looping System for Low-Cost CO<sub>2</sub> Capture from Seawater. *ACS Energy Lett* 7: 1947–1952.
20. Zhuang W, Song X, Liu M, et al. (2023) Potential capture and conversion of CO<sub>2</sub> from oceanwater through mineral carbonation. *Science of The Total Environment* 867: 161589.
21. Almajed HM, Guerra OJ, Smith WA, et al. (2023) Evaluating the techno-economic potential of defossilized air-to-syngas pathways. *Energy Environ Sci* 16: 6127–6146.
22. McQueen N, Gomes KV, McCormick C, et al. (2021) A review of direct air capture (DAC): scaling up commercial technologies and innovating for the future. *Prog Energy* 3: 032001.
23. National Academies of Sciences, Engineering, and Medicine (2019) Direct Air Capture, *Negative Emissions Technologies and Reliable Sequestration: A Research Agenda*, Washington, DC, The National Academies Press.
24. Keith DW, Holmes G, St. Angelo D, et al. (2018) A Process for Capturing CO<sub>2</sub> from the Atmosphere. *Joule* 2: 1573–1594.
25. Lackner K, Ziock H-J, Grimes P (1999) Carbon Dioxide Extraction from Air: Is It An Option?, Los Alamos National Lab. (LANL), Los Alamos, NM (United States).
26. Zeman FS, Lackner KS (2004) Capturing Carbon Dioxide Directly from the Atmosphere. *World Resource Review* 16: 157–172.
27. Zeman F (2007) Energy and Material Balance of CO<sub>2</sub> Capture from Ambient Air. *Environ Sci Technol* 41: 7558–7563.
28. Sabatino F, Grimm A, Gallucci F, et al. (2021) A comparative energy and costs assessment and optimization for direct air capture technologies. *Joule* 5: 2047–2076.
29. Fasihi M, Efimova O, Breyer C (2019) Techno-economic assessment of CO<sub>2</sub> direct air capture

- plants. *Journal of Cleaner Production* 224: 957–980.
30. Long-Innes R, Struchtrup H (2022) Thermodynamic loss analysis of a liquid-sorbent direct air carbon capture plant. *Cell Reports Physical Science* 3: 100791.
  31. Marchese M, Giglio E, Santarelli M, et al. (2020) Energy performance of Power-to-Liquid applications integrating biogas upgrading, reverse water gas shift, solid oxide electrolysis and Fischer-Tropsch technologies. *Energy Conversion and Management: X* 6: 100041.
  32. Alakurtti S, Arasto A, Hannula I, et al. (2019) The Carbon Reuse Economy: Transforming CO<sub>2</sub> from a pollutant into a resource, VTT Technical Research Centre of Finland.
  33. Rezaei E, Dzuryk S (2019) Techno-economic comparison of reverse water gas shift reaction to steam and dry methane reforming reactions for syngas production. *Chemical Engineering Research and Design* 144: 354–369.
  34. Joo O-S, Jung K-D, Moon I, et al. (1999) Carbon Dioxide Hydrogenation To Form Methanol via a Reverse-Water-Gas-Shift Reaction (the CAMERE Process). *Ind Eng Chem Res* 38: 1808–1812.
  35. Twelve (2023) Twelve, 2023. Available from: <https://www.twelve.co>.
  36. CO<sub>2</sub> utilization technologies | GIGKarasek Available from: <https://www.gigkarasek.com/en/renewable-carbon-solutions>.
  37. CO<sub>2</sub> for a clean performance | Evonik Industries Available from: <https://corporate.evonik.com/en/cosub2-sub-for-a-clean-performance-118401.html>.
  38. CO<sub>2</sub> Electrolyzers | Dioxide Materials Available from: <https://dioxidematerials.com/technology/co2-electrolysis/>.
  39. Kas R, Yang K, Yewale GP, et al. (2022) Modeling the Local Environment within Porous Electrode during Electrochemical Reduction of Bicarbonate. *Ind Eng Chem Res* 61: 10461–10473.
  40. Wen G, Ren B, Wang X, et al. (2022) Continuous CO<sub>2</sub> electrolysis using a CO<sub>2</sub> exsolution-induced flow cell. *Nat Energy* 7: 978–988.
  41. Towler G, Sinnott R (2013) Capital Cost Estimating, *Chemical Engineering Design*, Elsevier, 307–354.
  42. Shin H, Hansen KU, Jiao F (2021) Techno-economic assessment of low-temperature carbon dioxide electrolysis. *Nat Sustain* 4: 911–919.
  43. Jouny M, Luc W, Jiao F (2018) General Techno-Economic Analysis of CO<sub>2</sub> Electrolysis Systems. *Ind Eng Chem Res* 57: 2165–2177.
  44. Moreno-Gonzalez M, Berger A, Borsboom-Hanson T, et al. (2021) Carbon-neutral fuels and chemicals: Economic analysis of renewable syngas pathways via CO<sub>2</sub> electrolysis. *Energy Conversion and Management* 244: 114452.
  45. Debergh P, Gutiérrez-Sánchez O, Khan MN, et al. (2023) The Economics of Electrochemical Syngas Production via Direct Air Capture. *ACS Energy Lett* 3398–3403.
  46. Sisler J, Khan S, Ip AH, et al. (2021) Ethylene Electrosynthesis: A Comparative Techno-economic Analysis of Alkaline vs Membrane Electrode Assembly vs CO<sub>2</sub>-CO-C<sub>2</sub>H<sub>4</sub> Tandems. *ACS Energy Lett* 6: 997–1002.
  47. De Luna P, Hahn C, Higgins D, et al. (2019) What would it take for renewably powered electrosynthesis to displace petrochemical processes? *Science* 364: eaav3506.
  48. Brancucci Martinez-Anido C, Brinkman G, Hodge B-M (2016) The impact of wind power on electricity prices. *Renewable Energy* 94: 474–487.
  49. Congressional Research Services (2022) Tax Provisions in the Inflation Reduction Act of 2022 (H.R. 5376).
  50. Salvatore D, Berlinguette CP (2020) Voltage Matters When Reducing CO<sub>2</sub> in an Electrochemical Flow Cell. *ACS Energy Lett* 5: 215–220.
  51. Pei P, Korom SF, Ling K, et al. (2016) Cost comparison of syngas production from natural gas conversion and underground coal gasification. *Mitig Adapt Strateg Glob Change* 21: 629–643.
  52. Daniel T, Masini A, Milne C, et al. (2022) Techno-economic Analysis of Direct Air Carbon Capture with CO<sub>2</sub> Utilisation. *Carbon Capture Science & Technology* 2: 100025.
  53. The World Bank (2023) Carbon Pricing Dashboard.
  54. NREL 2023 Annual Technology Baseline (ATB) Cost and Performance Data for Electricity Generation Technologies [data set].
  55. U.S. Energy Information Administration (2023) Wholesale Electricity and Natural Gas Market Data.
  56. Guerra OJ, Almajed HM, Smith WA, et al. (2023) Barriers and opportunities for the deployment of CO<sub>2</sub> electrolysis in net-zero emissions energy systems. *Joule* 7: 1111–1133.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Towards Sustainable Supply Chains for Waste Plastics through Closed-Loop Recycling: A case-study for Georgia

Elisavet Anglou<sup>a</sup>, Riddhi Bhattacharya<sup>a</sup>, Patricia Stathatou<sup>b</sup>, and Fani Boukouvala<sup>a\*</sup>

<sup>a</sup> Department of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA

<sup>b</sup> Renewable Bioproducts Institute, Georgia Institute of Technology, Atlanta, GA 30332 USA

\* Corresponding Author: [fani.boukouvala@chbe.gatech.edu](mailto:fani.boukouvala@chbe.gatech.edu)

## ABSTRACT

Sustainable and economically viable plastic recycling methodologies are vital for addressing the increasing environmental consequences of single-use plastics. In this study, we evaluate the plastic waste management value for the state of Georgia, US and investigate the potential of introducing novel depolymerization methods within the network. An equation-based formulation is developed to identify the optimum supply-chain design given the geographic location of existing facilities. Chemical recycling technologies that have received increasing attention are evaluated as candidate technologies to be integrated within the network. The optimum supply-chain design is selected based on environmental and economic objectives. The designed network of pathways uses a mix of different technologies (chemical and mechanical recycling) in a way that are both economically environmentally sound.

**Keywords:** recycling, supply chain, plastics, waste management, optimization

## 1. INTRODUCTION

Plastic materials have revolutionized our daily lives gradually replacing materials used for centuries such as wood, glass, or steel. Unfortunately, a significant amount of the plastics used ends up in landfills or marine environments with a very small percentage of them currently being recycled. For context, 35.7 MT of plastics were generated in the US in 2018 while approximately 26.9 MT ended up to landfill (75.6%) [1]. Consequently, the transition to circular economies (CE) in which waste materials will be effectively re-used stands as one of the most prevalent challenges of our times particularly in context of waste plastics [2].

The plastic waste management routes can be categorized into four categories: pre-consumer, mechanical recycling, chemical recycling, and energy recovery pathways. Currently, mechanical recycling is the primary method for recycling due to its low cost and simplicity, however, the material properties of the plastic degrade during processing (each plastic can be recycled 2-6 times during each lifetime) [3, 4]. As a result, solely depending on mechanical recycling impedes the realization of a closed-loop recycling economy [4, 5].

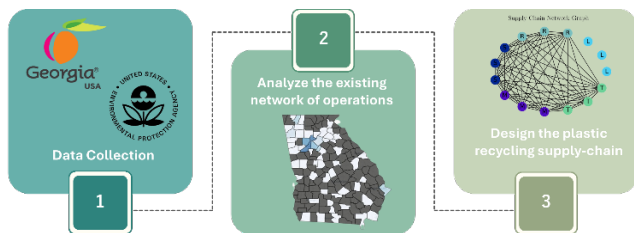
Chemical recycling of polymers through depolymerization pathways has garnered increased attention in the last decade as a promising alternative strategy. This is because chemical recycling enables the breakdown of polymers to constituent monomers thus, bypassing material degradation issues [3, 6]. A wide variety of methodologies have been proposed to chemically recycle waste plastics that differ considerably in terms of efficiencies, reaction pathways and maturity levels thus, ranking technologies and making decisions is nontrivial [7].

In combination to the already complex network of waste management options (e.g., landfilling, energy recovery, recycling etc.), design of economic and sustainable solutions that would maximize circularity is an open challenge. Previous studies have evaluated how to effectively manage waste systems through integrated supply-chains [8-11]. Ma et al. [4] developed a mixed integer linear program (MILP) approach to study the performance of thermochemical technologies at a regional-scale for low-density polyethylene (LDPE) and polypropylene (PP) waste. Recently, Badejo et al. [12] examined multiple technologies for managing high-density polyethylene (HDPE) waste using an MILP framework focusing on the



United States East Coast.

In this work, we model the supply chain of two types of waste plastics (i.e., polyethylene terephthalate (PET) and HDPE) through an equation-based framework using geospatial data for the state of Georgia, US. Geographic locations of the existing waste processing facilities (e.g., landfilling, mechanical recycling, etc.) are integrated within the network. Subsequently, we assess the potential of integrating novel chemical recycling methodologies into the superstructure for processing PET and HDPE waste. Literature data are incorporated into the formulation, to enable capturing the economic and emissions trade-offs between alternative technologies. Further, we employ multi-objective optimization to identify the Pareto-optimal solutions. In summary, our study presents a computational framework based on a realistic system representation that can be utilized to compare alternative pathways and enable the design of cost-effective supply-chains for PET and HDPE waste plastics.



**Figure 1.** Overview of our analysis including data collection, formulation and optimization of the supply chain model.

## 2. MODEL DESCRIPTION

### 2.1. Model overview and problem statement

An equation-based formulation is established to systematically model the value chain of plastics. The primary objectives are: (a) to determine the most efficient path for plastic waste from various points of collections ultimately to the end-user; and (b) the mix of transformation technologies that would maximize circularity. The model is constructed as a graph where its nodes represent the source, transformation, and demand facilities. Each node is characterized by the geographic coordinates of each facility in Georgia, US (GA). The edges of the graph represent the material flowing between the nodes. This manuscript specifically studies the plastic waste management in the state of Georgia; however, the framework is generalizable and can be adapted for use in other regions that can vary in size.

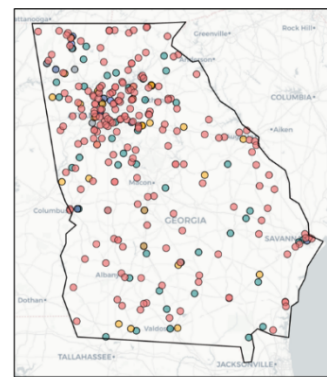
The supply network in Georgia consists of distinct operations: (a) the plastic waste (HDPE and PET) is collected locally at transfer stations; (b) transported to landfills or recycling facilities; (c) transformed to usable forms through mechanical or chemical recycling; (d) mixed with virgin plastics to meet the market demands; and (e)

transferred to the end-users. A simplistic graph of the route that plastics follows through the designed network is depicted in Figure 2.

Data sourced by the state of Georgia and the Environmental Protection Agency (EPA) [1, 13] are used as inputs to the model. Figure 3 visually presents the geographic locations of the system's nodes. The databases were further refined to include only those locations that process or utilize PET and HDPE. A total of 201 collection sites, 44 landfills, 29 recycling facilities (with 15 processing HDPE and 14 PET), and 30 market sites (11 and 19 sites that demand HDPE and PET, respectively) are considered.



**Figure 2:** Overview of the potential routes for plastic waste upcycling.



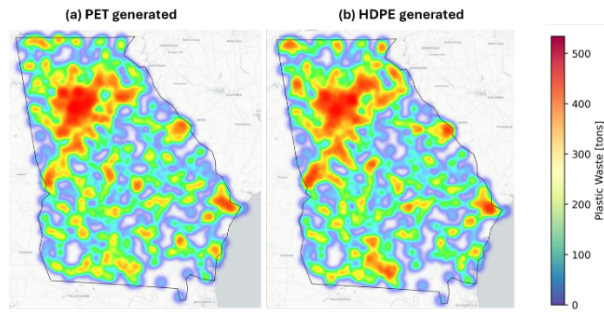
● Transfer Stations ● Landfills ● PET Recycling ● HDPE Recycling ● Market

**Figure 3:** Map of the state of Georgia, US including all nodes of the PET and HDPE supply chain.

Figure 4 (a) and (b) depicts the spatial distribution of plastic waste generation for PET and HDPE in Georgia, respectively sourced from EPA database[1]. The amounts of HDPE and PET waste generated at each zip-code, are assigned as input rates to the closest collection site. It is assumed that the waste is already separated to HDPE and PET at each collection. The demand capacities for the each of the market locations is also assumed to be equal to the required amount of plastic at the closest zip-code [1].



At the existing geographic locations of mechanical recycling nodes, hypothetical chemical recycling facilities are introduced. It is assumed that each recycling node can process plastic waste either chemically or mechanically, provided that the capacity constraints are not exceeded. Various candidate technologies are compared for PET, including dissolution, enzymatic hydrolysis, glycolysis and methanolysis based on literature data [7]. For HDPE the investigation is focused only on dissolution as the primary chemical recycling technology.



**Figure 4:** Heatmap of annual (a) PET (b) HDPE waste generated in Georgia, US as extracted by the EPA database.

## 2.2. Model formulation

As mentioned earlier, the supply chain model adopts a graph network representation. The set of all nodes is denoted as  $N$ . Collection sites ( $T$ ), landfill ( $L$ ), mechanical recycling ( $MR$ ), chemical recycling ( $CR$ ), and market ( $M$ ) sets are introduced to describe the sets of distinct type of facility. Therefore, the entire network is described as  $N = T \cup L \cup MR \cup CR \cup M$ . Furthermore, for each of the sets introduced, a subset for PET and HDPE nodes are also established such that for example  $MR = MR_{PET} \cup MR_{HDPE}$ . The same holds for each of the sets introduced apart from  $L$  that can accept both HDPE and PET waste. Each chemical recycling node is characterized by the candidate depolymerization technology. A binary variable  $y_{i,j,tech}$  is introduced to specify the type of chemical recycling technology implemented at each CR node. Finally, every node is characterized by the material flow  $m_{ij}$ , from the starting node  $i$  to the destination node  $j$ .

The model is formulated as a MILP with the objective of minimizing the economic or environmental impact. The environmental (Equation (1)) and economic (Equation (2)) objectives are defined as follows:

$$E = \sum_{i=1}^N \sum_{j=1}^N d_{ij} m_{ij} E_{mile_{ij}} + \sum_{j=1}^N \sum_{i=1}^N m_{ij} E_{i,j} \quad (1)$$

$$C = \sum_{i=1}^N \sum_{j=1}^N d_{ij} m_{ij} C_{mile_{ij}} + \sum_{j=1}^N \sum_{i=1}^N m_{ij} C_{i,j} \quad (2)$$

Here,  $d_{ij}$  represents the distance between node  $i$  and  $j$ , while  $m_{ij}$  signifies the material flow between the nodes.  $E_{ij}$  and  $C_{ij}$  denote the  $CO_2$ -eq emitted and the cost

of production at the specific node with an input flow of  $m_{ij}$ , while  $C_{mile}$  and  $E_{mile}$  the cost and the emissions of transporting plastic through the edges of the network. The first term in Equations (1) and (2) represents the impact of the transportation to the objectives while the second term the impact of processing the plastic within each node. The material balances of flows entering and exiting each node are given by Equation (3) with  $\alpha_{ij}$  representing the conversion factors. Different conversion factors are assigned based on the technology selected, processing node and material type. Equation (3) should be satisfied for all nodes.

$$\sum_{i=N_{in}} \sum_{j=N_{in}} m_{i,j} \alpha_{ij} = \sum_{i=N_{out}} \sum_{j=N_{out}} m_{ij} \quad (3)$$

Furthermore, demand and source constraints are imposed through Equations (4)-(5). Equation (4) guarantees that the market needs are fulfilled. The demands are satisfied by mixing recycled with virgin, petroleum-derived plastic denoted as  $vPET$  and  $vHDPE$ , respectively. This is included because the current market requirements cannot be exclusively satisfied through recycling, even if all collected plastic is recycled. Equation (5) ensures that all waste gathered is effectively managed either through recycling or disposed at the landfills. Moreover, Equation (6) introduces a capacity constraint within the formulation, to prevent material flows from exceeding the maximum processing capacity of each facility.

$$\sum_{i=CR} \sum_{j=M} m_{i,j} + v_{Plastic} = \sum_{i=M} Demand_i \quad (4)$$

$$\sum_{i=T} \sum_{j=L \cup CR \cup MR} m_{i,j} = \sum_{i=T} Source_i \quad (5)$$

$$\sum_{i=CR \cup MR} \sum_{j=M} m_{i,j} \leq \sum_{i=IM} Capacity_i \quad (6)$$

Equations (7) and (8) set a quality constraint for the final product. This ensures that mechanically recycled plastic can be used in the production of new materials only if blended at a maximum threshold of 50% with virgin or chemically recycled plastic. Finally, Equation (9) guarantees that at each CR node, only one technology can be active.

$$\sum_{i=R} \sum_{j=M} m_{i,j} \leq 0.5 (\sum_{i=CR} \sum_{j=M} m_{ij} + \sum_{i=M} v_{PET_i}) \quad (7)$$

$$\sum_{i=R} \sum_{j=M} m_{i,j} \leq 0.5 (\sum_{i=CR} \sum_{j=M} m_{ij} + \sum_{i=M} v_{HDPE_i}) \quad (8)$$

$$\sum_{tech} \sum_{i=CR} \sum_{j=M} y_{i,j,tech} \leq 1 \quad (9)$$

The emission and cost factors utilized in Equations (1)-(2) are depicted in Table 1. All the environmental factors apart from the chemical recycling steps were obtained from the ecoinvent database v3.10, following the ReCiPe 2016 methodology [14, 15]. The emissions and cost factors for the investigated chemical recycling technologies are taken from a recent study conducted by Uekert et al. [7] in which the authors performed rigorous simulations of various chemical recycling processes. The

cost factors for the rest of the steps are taken from the literature and published industry data [7, 16, 17].

**Table 1:** Emissions and cost factors of processing, feed, and mechanical recycling nodes

Process	Price (\$/kg)	Emissions (kg CO <sub>2</sub> -eq/kg)
Landfill HDPE	0.08	0.138
Landfill PET	0.08	0.095
MR HDPE	0.63	0.737
MR PET	0.54	1.337
Virgin HDPE	0.79	3.277
Virgin PET	1.19	4.086
Transportation	0.04/km	0.196/metric ton km

**Table 2:** Emissions and cost factors for chemical recycling technologies investigated

Process	Price (\$/kg)	Emissions (kg CO <sub>2</sub> -eq/kg)
Dissolution PET	0.87	4.49
Enzymatic Hydrolysis PET	2.01	3.95
Glycolysis PET	0.96	1.32
Methanolysis PET	1.05	4.19
Dissolution HDPE	1.10	2.40

### 2.3. Solution Strategy

The problem defined by Equations (1)-(9) is formulated as an MILP problem. The mathematical problem is solved in Pyomo with CPLEX v22.2. First, the two single-objective problems are solved independently, and the optimal network configurations and technologies of choice are obtained. In the case that the single-objective configurations are different, it indicates a conflict between the two solutions. This implies that there are trade-offs between economically sound and environmentally friendly solutions. Next, to analyze the trade-offs between the two solutions, the multi-objective optimization problem is formulated and solved using the  $\epsilon$ -constraint methodology (Equation 10). The mathematical problem is transformed to a single-objective, bounded by an additional constraint that corresponds to the other objective [18, 19].  $C_{\min}$  and  $C_{\max}$  correspond to the minimum and maximum values of the cost objectives as identified by the single-objective problems.

$$\min E \quad (10a)$$

$$\text{s. t. } C < \epsilon, \text{ where } C_{\min} \leq \epsilon \leq C_{\max} \quad (10b)$$

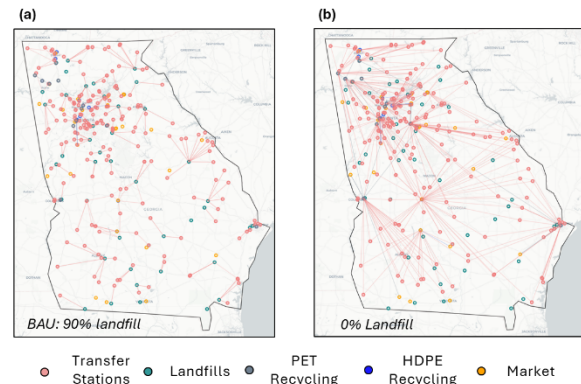
## 3. RESULTS AND DISCUSSION

To evaluate the potential of integrating chemical recycling within the existing network we analyze: (a) the

environmental impacts (Section 3.1), (b) the economic implications (Section 3.2), and (c) both objectives simultaneously (Section 3.3).

### 3.1. Environmental considerations of the GA recycling network

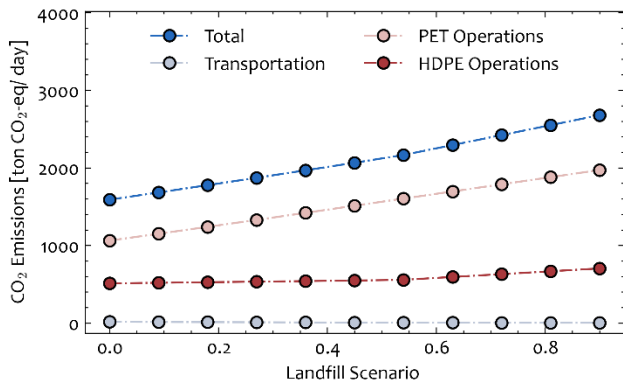
The spatial solution considering the environmental objective is depicted in Figure 5 (a). This solution is representative of the Business-As-Usual (BAU) scenario, where 10% of the plastic waste is recycled. A very small portion of the generated waste plastic are directed to recycling facilities. For PET, mechanical recycling nodes remain inactive and all of plastics are processed through chemical recycling. This is because of the higher conversion rates and relatively similar emissions rates between mechanical recycling and glycolysis (e.g., candidate technology with lower emission factor). Strictly only considering emissions, chemical recycling is favored in the BAU case (e.g., 90% landfill). The opposite is true for HDPE, where all plastics that is not landfilled are mechanically processed. This is attributed to the relatively lower emissions of MR as compared to CR for HDPE. It is worth acknowledging that the emission data for chemical recycling processes are based on experiments and simulations hence, the results may change as closed loop depolymerization recycling methods are further explored and optimized at larger scales.



**Figure 5:** Solution for (a) 90% landfill scenario; (b) 0% landfill scenario under the environmental objective.

To further explore the potential of recycling, sensitivity analyses are carried out to investigate the scenarios in which more plastic waste is recycled. The percentage of waste directed to landfills is changed to represent hypothetical scenarios. As landfilling is increasingly banned across Europe and some US states, those scenarios are important to consider. The network configuration for the zero-waste scenario is depicted in Figure 5 (b). The solution showcases an interconnected network of nodes with both chemical and mechanical recycling facilities activated.

The total CO<sub>2</sub>-eq emissions of the different scenarios are highlighted in Figure 6. It is observed that as landfill is reduced, the total emitted CO<sub>2</sub>-eq are also reduced, even though the emissions of landfilling are minimal. This is outweighed by the fact that plastics diverted from landfills are transformed into usable forms which, consequently, reduces the virgin plastic amount required to satisfy the market demands. Moreover, this is supported by the relatively small emission factors of glycolysis as compared to fossil production. The total emissions reported for HDPE are far lower than those of PET only because the input waste amount of PET is approximately double that of HDPE waste.

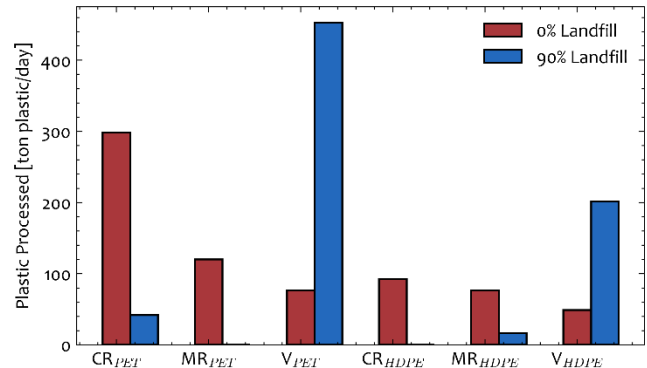


**Figure 6:** Emitted CO<sub>2</sub>-eq for different landfilling scenarios categorized by distinct operations

Furthermore, more recycling nodes are activated as landfilling is reduced since more waste needs to be processed while virgin requirements are reduced. The mix of technologies utilized to fulfill the market demands are visually represented in Figure 7 for the BAU and zero-waste scenarios. In the BAU scenario, the demand of new plastic is fulfilled mainly through the production of virgin materials while, in the zero-waste case, mainly through the transformation of waste to useful products. The different solutions can be attributed to quality requirements set by Equations (7) and (8), the capacity constraints set for recycling as well as the amount of plastic diverted from the landfills to processing nodes.

In terms of the investigated technologies, glycolysis outperforms all the other candidate methods and is chosen as the most promising recycling solution in all the investigated scenarios. This is highly correlated with the fact that emissions of glycolysis are considerably lower compared to all others and very similar to those of mechanical recycling. The trade-off is that glycolysis has lower conversion than some of the other candidate technologies, however, this does not outweigh the higher environmental impact. Moreover, the use of plastic sourced from fossil resources is reduced in all those hypothetical cases, as highlighted in Figure 7. As more plastic waste is available for re-processing, the market demands do

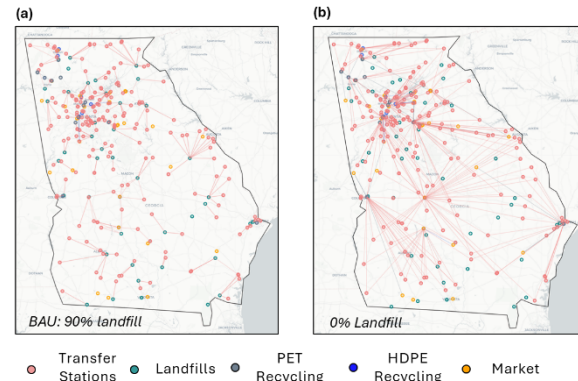
not have to be met with virgin quantities and the emissions are reduced. However, even under the zero-waste scenario, some virgin plastic is still required to meet the demands due to material losses occurring at intermediate nodes within the value-chain.



**Figure 7:** Amount of plastic processed through chemical and mechanical recycling compared with virgin requirements for 0% and 90% landfill scenarios.

### 3.2. Economic considerations of the GA recycling network

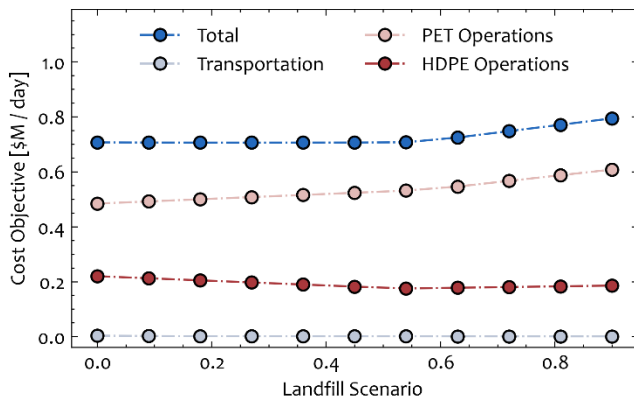
The spatial solution when considering the economic implications of the network are highlighted for the BAU and the zero-waste cases in Figure 8 (a)-(b). The effect of the same circularity scenarios to the economic objective are depicted in Figure 9.



**Figure 8:** Optimal network configuration under the (a) BAU; and (b) zero-waste scenario considering the economic objective.

It is observed that the evaluated cost objective is relatively constant throughout the different scenarios because of the similar cost factors in the competing scenarios (e.g., fossil production compared to recycling). The use of recycling technologies in the zero-waste scenarios does not reduce the cost relatively to the BAU that needs fossil – derived plastic to meet the market demands. More specifically, the cost of virgin PET is set at

\$1.19/kg compared to \$0.87/kg for the dissolution that has the lowest cost factors amongst the competing technologies, while the cost of landfill is minimal. These two relatively close values in conjunction with the zero-waste constraint, drive the cost of recycling up, and even though recycling is selected, the cost is not substantially reduced. Moreover, the HDPE operation costs increase as the share of landfilling is reduced due to the higher costs of chemical recycling when compared to virgin or MR production. This trend is influenced by the quality constraints set by Equation (8) which enforces a certain amount of HDPE waste to be processed through CR, because otherwise the quality of the final product will be inadequate.



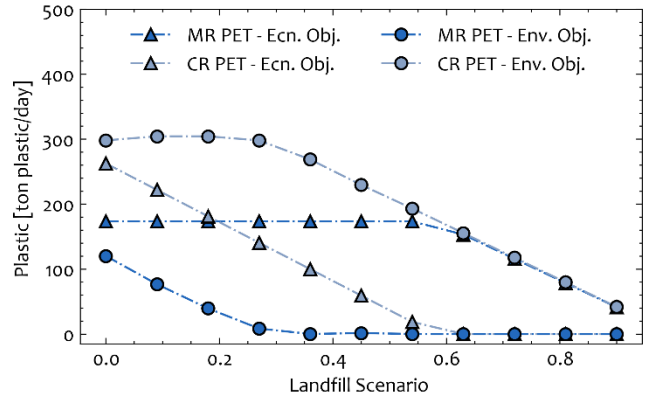
**Figure 9:** Cost objective for different landfilling scenarios as identified for the different technologies.

In terms of the competing technologies for PET, dissolution outperforms all the other options, being different from the environmentally favorable recycling methodology (i.e., glycolysis). This result is attributed to the relatively low-cost factors used for dissolution. In addition, the virgin requirements follow similar trends with the environmentally friendly solution discussed earlier with less fossil-derived plastic required, as landfilling is reduced. All in all, the solution identified for the cost minimization scenario corresponds to a different network with different chemical recycling technologies activated compared to the green-house gas (GHG) minimization scenario, as depicted in Figure 10. This shows that the solutions are in fact in contrast.

### 3.3. Identifying trade-offs between environmentally friendly and cost-effective solutions.

In this section, we investigate the trade-offs between the economically friendly and environmentally sound solutions. The values of the two objectives are depicted in Table 3 for the two distinct landfilling scenarios (0%, 90%). This, along with the recommended technologies (e.g., glycolysis is chosen as the best environmental

case, dissolution for the most economical solution) indicates that the two objectives are in conflict meaning that different value chain networks are optimal. Moreover, the amounts of PET processed through chemical and mechanical recycling are very different for the two single-objective problems as illustrated in Figure 10.



**Figure 10:** vPET required for different landfilling scenarios and technologies to meet the market demands.

The BAU solution is an exception since the identified configuration is very similar for the two problems. This is because very small amounts of plastic waste are recycled and almost all of it is diverted to landfills. The difference between the two cases is how the 10% of waste will be processed. For the zero-waste case studied, the spatial solution and the selected technologies are different, as highlighted in Table (3). Therefore, the trade-offs between the two objectives for this scenario are evaluated by following the procedure outlined in the Solution Strategy section.

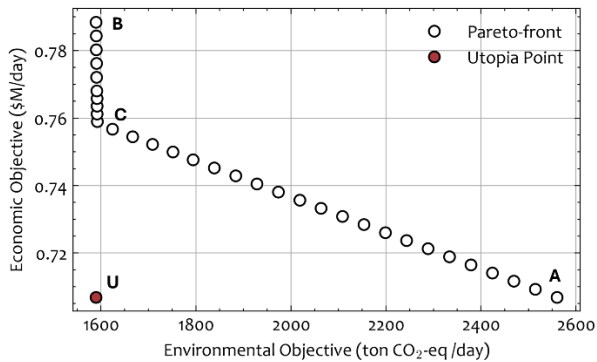
**Table 3:** Values of the environmental and economic for the two extreme landfilling scenarios

Objectives	Economics (M\$/day)	Emissions (ton-CO <sub>2</sub> -eq/ day)
<b>90% Landfill</b>		
min Env	0.82	2679
min Ecn	0.79	2682
<b>0% Landfill</b>		
min Env	0.79	1590
min Ecn	0.71	2561

The Pareto front is depicted in Figure 11 for the case of 0% landfilling. Point A represents the optimal configuration corresponding to the minimum cost objective, regardless of the environmental impact. Similarly, point B represents the optimal configuration from an environmental perspective. The exact values of the objectives are presented in Table 3. All other Pareto solutions that lie in between points A and B signify the trade-offs between the two objectives based on the level of



importance between the values.



**Figure 11:** Pareto front and utopia point for the multi-objective optimization problem for the zero-waste case.

Point U, also referred to as Utopia point, reflects an ideal solution, in which the values obtained from the two single-objective problems are plotted. This is a hypothetical scenario that can never be reached for those two conflicting scenarios. The different solutions observed as we move along the Pareto-front stem from changes in the configuration of the optimal waste management network in the studied region and the technology chosen. The most significant difference between the two configurations is that the economically friendly solution only dissolution is chosen as a candidate technology, while in the environmentally sound value chain glycolysis is favored. Glycolysis and dissolution technologies are activated with different quantities processed as we move between points A and C. After point C, the cost objective increases at a different rate than the environmental objective. This is attributed to fact that after this point, only glycolysis is the active CR technology with some amounts processed through MR. As we move between point C to B, waste is diverted from MR to CR which increases the processing costs but has minimal impact to the emissions. This is because the emissions of MR and glycolysis for PET are very similar, while the opposite is true for the cost factors.

#### 4. CONCLUSIONS

In this work, we discuss the design of optimal value chains of plastic recycling for a specific set of collection sites, transformation facilities and market. A superstructure network model was formulated to describe the existing recycling chain in the state of Georgia in the US. Our analysis evaluates the potential of integrating different depolymerization recycling technologies for PET and HDPE waste within the existing network. Different spatial arrangements and technologies are selected and designed for environmentally and economically friendly solutions. This contribution focuses only on the comparison of different depolymerization methods. However, to

holistically evaluate the supply-chain of plastic recycling, it is essential to compare other technologies not solely based on depolymerization, such as energy recovery or feedstock recycling. The landscape of plastic waste management is highly dynamic and as new technologies are advanced and new research is conducted, the emission and cost factors used in this analysis may change. In terms of the HDPE feedstock, this assessment only accounted for the use of dissolution as a chemical recycling technology. Future work will delve deeper into the integration of additional chemical recycling technologies beyond PET and HDPE. We anticipate that this approach will allow for the evaluation and comparison of multiple waste management pathways along with more realistic feedstocks.

#### ACKNOWLEDGEMENTS

This work is supported by the U.S. National Science Foundation—Emerging Frontiers in Research and Innovation program under grant 2028998. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

#### REFERENCES

1. Environmental Protection Agency: <https://epa.maps.arcgis.com/apps/webappviewer/index.html?id=fefb46fd31a14b80a836bd0bd4d788e2>. Accessed: 12/2023].
2. Avraamidou, S., et al., Circular Economy - A challenge and an opportunity for Process Systems Engineering. *Computers & Chemical Engineering*, 2020. **133**: p. 106629-106629.
3. Tricker, A.W., et al., Stages and kinetics of mechanochemical depolymerization of poly (ethylene terephthalate) with sodium hydroxide. *ACS Sustainable Chemistry & Engineering*, 2022. **10**(34): p. 11338-11347.
4. Ma, J., et al., Economic Evaluation of Infrastructures for Thermochemical Upcycling of Post-Consumer Plastic Waste. 2022.
5. Thakker, V. and B.R. Bakshi, Designing Value Chains of Plastic and Paper Carrier Bags for a Sustainable and Circular Economy. *ACS Sustainable Chemistry & Engineering*, 2021. **9**(49): p. 16687-16698.
6. Luo, Y., et al., Economic and Environmental Benefits of Modular Microwave-Assisted Polyethylene Terephthalate Depolymerization. *ACS Sustainable Chemistry & Engineering*, 2023. **11**(10): p. 4209-4218.
7. Uekert, T., et al., Technical, economic, and



- environmental comparison of closed-loop recycling technologies for common plastics. *ACS Sustainable Chemistry & Engineering*, 2023. **11**(3): p. 965–978.
8. Rathore, P., et al., Towards a sustainable organic waste supply chain: a comparison of centralized and decentralized systems. *Journal of Environmental Management*, 2022. **315**: p. 115141.
  9. Saif, Y., S. Griffiths, and A. Almansoori, Municipal solid waste supply chain management under an integrated optimization of sustainability targets. *Computers & Chemical Engineering*, 2022. **160**: p. 107725.
  10. Allman, A., et al., Biomass waste-to-energy supply chain optimization with mobile production modules. *Computers & Chemical Engineering*, 2021. **150**: p. 107326.
  11. Brandão, R., et al., Reverse supply chain conceptual model for construction and demolition waste. *Waste Management & Research*, 2021. **39**(11): p. 1341-1355.
  12. Badejo, O.A., et al., Design of Sustainable Supply Chains for Managing Plastic Waste: A Case Study for Low Density Polyethylene. 2024.
  13. Georgia State Website: <https://map.georgia.org/explore/#recycling>. Accessed: 10/2023].
  14. Wernet, G., et al., The ecoinvent database version 3 (part I): overview and methodology. *The International Journal of Life Cycle Assessment*, 2016. **21**(9): p. 1218-1230.
  15. Huijbregts, M.A.J., et al., ReCiPe2016: a harmonised life cycle impact assessment method at midpoint and endpoint level. *The International Journal of Life Cycle Assessment*, 2017. **22**(2): p. 138-147.
  16. Clayton County Landfill <https://www.claytoncountyga.gov/government/transportation-and-development/landfill/#:~:text=All%20Landfill%20Scale%20Rates%20are,for%20the%201st%201%2C000%20lbs>. 12/2023].
  17. Freight rates: Trucking rates per mile 2024 <https://www.method.me/pricing-guides/trucking-rates-per-mile/#:~:text=To%20answer%20your%20question%20on,are%20at%20%243.14%20per%20mile>. Accessed 12/2023].
  18. Mavrotas, G., Effective implementation of the  $\epsilon$ -constraint method in multi-objective mathematical programming problems. *Applied mathematics and computation*, 2009. **213**(2): p. 455-465.
  19. Giannikopoulos, I., et al., Multi-objective optimization of production cost and carbon loss in the US petrochemicals industry, in *Computer Aided Chemical Engineering*. 2022, Elsevier. p. 547-552.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Environmental Impact of Simulated Moving Bed (SMB) on the Recovery of 2,3-Butanediol on an Integrated Biorefinery

Marco E. Avendano<sup>a</sup>, Jianpei Lao<sup>a</sup>, Qiang Fu<sup>a</sup>, Sankar Nair<sup>a</sup>, and Matthew J. Realff<sup>a\*</sup>

<sup>a</sup> Georgia Institute of Technology, School of Chemical and Biomolecular Engineering, Atlanta, GA, US

\* Corresponding Author: [matthew.realff@chbe.gatech.edu](mailto:matthew.realff@chbe.gatech.edu).

## ABSTRACT

2,3 butanediol (BDO) has garnered recent interest due to the high titer concentrations that can be obtained through biochemical routes and its potential for efficient conversion into long-chain hydrocarbons. BDO separation, however, is challenging given its low volatility and high affinity towards water. In this study, two BDO separation pathways were compared, single distillation and combined simulated moving bed (SMB) adsorption with distillation. The separations were incorporated into a 2018 biorefinery design developed by the National Renewable Energy Laboratory (NREL) to produce renewable fuels from corn stover, with BDO as an intermediate and adipic acid as the co-product. The comparison was performed on the basis of sustainability, using lifecycle greenhouse gas (GHG) emissions as the metric. It was found that using a single distillation column gives GHG emissions of 48 g<sub>CO<sub>2e</sub></sub>/MJ for the renewable fuel. This is lower than 93 g<sub>CO<sub>2e</sub></sub>/MJ for petroleum fuel but is higher compared to the SMB-based process which achieves 21 g<sub>CO<sub>2e</sub></sub>/MJ. Additionally, the minimum fuel selling price (MFSP) of each pathway was computed. Single distillation gave a minimum MFSP of \$2.54/GGE (gallon of gasoline equivalent) of fuel, while SMB reached \$2.45/GGE. The SMB's MFSP is lower than the Department of Energy's (DOE) target of \$2.50/GGE, demonstrating this pathway is both an economic and sustainable alternative and a sound separation candidate that can enable the viability of the entire biorefinery. The effect of BDO fermentation titer was also considered through a sensitivity analysis.

**Keywords:** Adsorption, Biofuels, Distillation, Life Cycle Analysis, Technoeconomic Analysis.

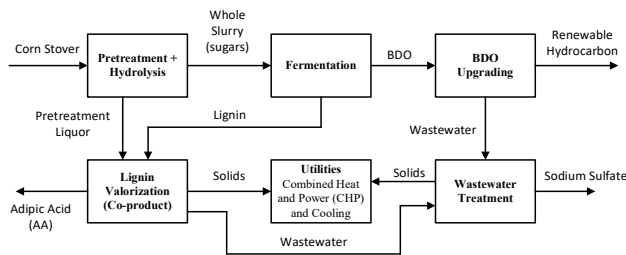
## INTRODUCTION

2,3 butanediol (BDO) is a molecule of high interest in the chemical and energy industries that is commonly used as an intermediate for added-value products. BDO is commonly produced by the catalytic conversion of C4 components from hydrocarbon mixtures [1]. Although high conversions can be achieved, these pathways have expensive operating costs and still rely on fossil fuels. This has kept BDO from becoming a sustainable and widespread chemical. Recent advances in the biochemical production of BDO from biomass-derived sugars, and high fermentation titers (50-110g/L), have led to an increased attention to this molecule [2].

BDO is seen as a platform to help decarbonize hard-to-abate sectors, such as heavy transportation (trucking

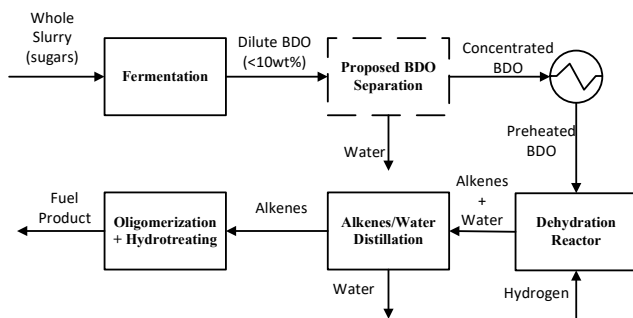
and aviation). It has a higher heating value and is less volatile than more commonly produced compounds like ethanol or butanol. Thus, at high purity values it can be used as blendstock or even directly as a drop-in fuel. BDO can also be an intermediate to produce renewable hydrocarbons. Its four carbons and two hydroxyl groups allow the dehydration into high carbon alkenes, which could facilitate subsequent oligomerization and hydrotreating steps. Most notably, in 2018 the National Renewable Energy Laboratory (NREL) [2] developed a process design that details the production of a renewable 50/50 diesel+naphtha fuel product from corn stover. BDO obtained from fermentation is the key intermediate of this process and adipic acid is the main co-product. This design will be referred to as the state-of-technology (SOT). **Figure 1** shows a high-level diagram of the SOT. The inside-

battery limit (ISBL) includes alkaline pretreatment, hydrolysis, fermentation, upgrading to fuels and co-product trains. This is an integrated biorefinery, where utilities and other services are provided internally, including combined heat and power (CHP), cooling and wastewater facilities in the outside-battery limit (OSBL).



**Figure 1.** High level diagram of the NREL SOT process. Adapted from Davis et al, 2018 [2]

As seen, the ISBL in the SOT lacks a separation step that recovers BDO post-fermentation. This dilute broth (~100g/L BDO titer) is directly sent to the dehydration reaction. In reality, however, is unlikely for the reactor to efficiently upgrade BDO at dilute conditions, making the separation a required step in the plant. **Figure 2** shows a more detailed diagram of the modified upgrading portion of the SOT, with the proposed separation located between fermentation and dehydration to concentrate BDO.



**Figure 2.** Block diagram based on NREL SOT's [2] main processes downstream of the fermentation

Rejecting water reduces stream size, which directly affects other units. This includes the preheater before dehydration, the dehydration reactor itself and the post-dehydration separation (alkenes/water distillation). The decrease in stream size reduces equipment size, leading to lower capital costs. Most importantly, the BDO pre-heating energy is also reduced, significantly decreasing operating costs and fuel usage. This leads to much smaller CO<sub>2</sub>e emissions that can make the fuel product a sustainable substitute of fossil hydrocarbons.

Nevertheless, BDO separation from a dilute aqueous mixture is challenging. BDO's high boiling point (177°C) and affinity towards water make distillation highly energy

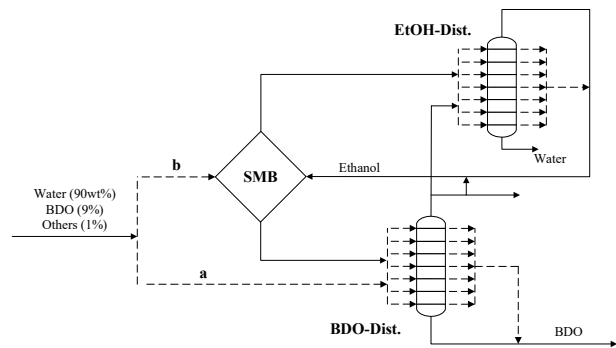
intensive. Simulated moving bed (SMB) adsorption is proposed as an alternative. Adsorption is a material-based separation, and SMB can enrich BDO to the required purity levels, without incurring in large operating expenses.

The main objective of this study is to compare the lifecycle of two pathways that recover BDO from a fermentation broth: 1) simple distillation and 2) SMB + distillation. Greenhouse gas (GHG) emissions was chosen as the sustainability metric, which is minimized for each pathway through deterministic optimization. To ensure the biorefinery is viable, the resulting design is subject to a minimum economic performance.

## PROCESS DESCRIPTION

### Single Column Distillation

**Figure 3** shows the possible separation alternatives to recover BDO. Single distillation bypasses the SMB and sends the broth directly to the column. This process operates under mild vacuum (0.1-0.5 bar) to prevent high temperatures that could lead to side reactions and sugar degradation. A tray efficiency of 80% was assumed and the number of theoretical stages was fixed at 20. This number was found to offer an optimal balance between capital and operating costs through a sensitivity analysis of the total annualized cost.



**Figure 3.** High level process diagram of the proposed BDO separation alternatives: a) single distillation column and b) SMB + distillation

Water primarily exits the condenser as distillate, while BDO is recovered from the bottoms. The broth impurities leave the column through either the distillate or bottoms, depending on the volatility. Organic acids (malic, lactic and acetic) and inorganic ions (Cl<sup>-</sup> and SO<sub>4</sub><sup>2-</sup>) are recovered from the distillate. Acetoin, glycerol, xylitol and residual sugars (arabinose, xylose and maltose) are retrieved from the bottoms. The concentration of these impurities is very small (<1%) in all streams and their presence is not expected to considerably affect the cost.

### Simulated Moving Bed + Distillation

The core of the proposed novel separation is the simulated moving bed (SMB) adsorption unit. SMB adsorption is a technology that mimics movement of solid throughout the bed by fixing the adsorbent and rotating instead the inlet and outlet ports. A 4-4-4-4 SMB is considered, where each column is packed with pure silica MFI-type zeolite, a highly hydrophobic material that is very selective towards BDO. Ethanol was chosen as the SMB desorbent. It has enough affinity with the adsorbent to efficiently displace the adsorbed BDO, and is highly volatile, facilitating its recovery via distillation. Following the SMB are two distillation columns, one to recover BDO from the extract and the other to recover ethanol from the raffinate.

A case study for PAREX [3], a well-known industry-scale SMB process, was used to guide column size (length and diameter), maximum column velocity, Peclet number and other high level operating decisions. Both distillation columns were modelled with the same approach as the single column. A tray efficiency of 50% was assumed for ethanol/water. Through a total annualized cost sensitivity analysis, 30 theoretical stages were found to offer the optimal trade-off between capital and operating costs for both columns.

## METHODOLOGY

### Physical Parameters

Figure 4 shows the stepwise approach that was followed to perform the comparative study. For either case, the first step was obtaining the key physical parameters needed to model either separation unit. For distillation, UNIQUAC was used to compute activity coefficients. The ASPEN Plus V.12. database [4] was used to retrieve the interaction parameters of the binary pairs. These coefficients are based on experimental vapor liquid equilibrium (VLE) data. This database was also used to retrieve the coefficients needed to calculate vapor pressure, heat capacity and reference enthalpy of each component.

For adsorption, a linear driving force (LDF) assumption was used to model the mass transfer from the bulk liquid into the solid adsorbent. The equilibrium in the adsorbent was represented with a Mixed Linear + Langmuir (MLL) isotherm. The relevant parameters are then the apparent mass transfer coefficient,  $K_{app,i}$  Henry's linear constant,  $H_i$  saturation capacity,  $q_{m,i}$ , and Langmuir's affinity,  $K_i$  where  $i$  represents each component. These parameters were obtained by fitting experimental data of a lab-scale SMB pilot plant.

### Mathematical Models

Next, the obtained parameters were used to solve the mathematical equations that represent each system. Distillation was rigorously modeled, meaning mass and energy balances were performed in each stage, while

enforcing phase equilibrium for each component. For adsorption, the mass balance, the convection-diffusion equation was used with axial dispersion values estimates from the Peclet number. It was assumed that the SMB stayed constant at 50°C, as this was the temperature of the pilot-plant experiments. The boundary conditions, flow balances at each node and enforcement of cyclic steady state (CSS) conditions constitute the remainder of the SMB equations.

This led to a system of partial differential algebraic equations (PDAEs) that were modelled as a non-linear programming (NLP) problem. Pyomo.DAE 6.5. [5] was used to discretize the PDAEs and the system was solved using Ipopt 3.12.13 [6].

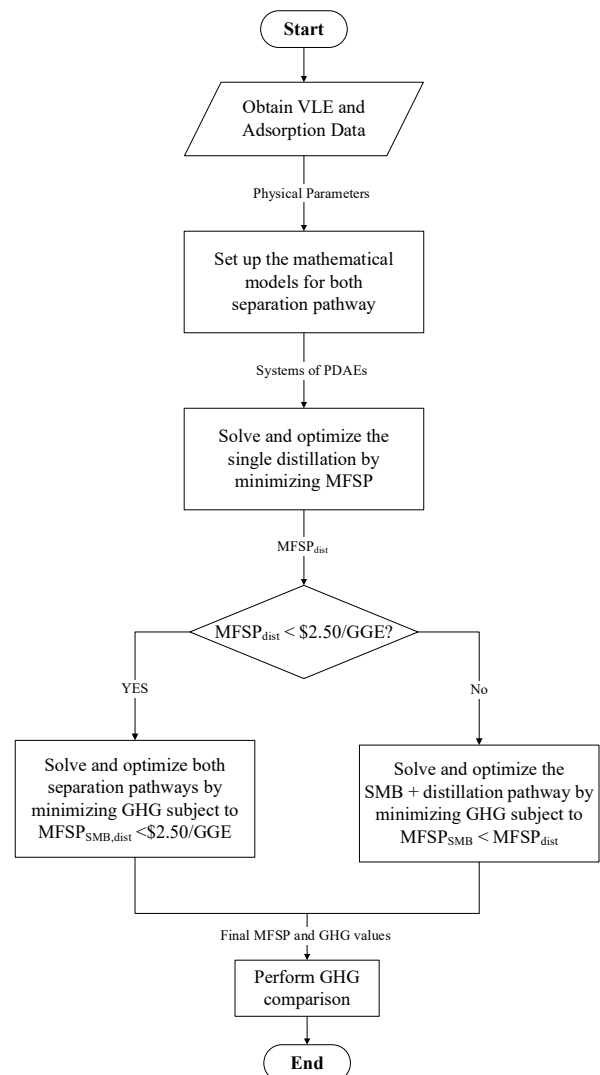


Figure 4. Proposed stepwise approach to develop the comparative study between separation pathways

### Economic Analysis

### Separation Scheme

After solving the PDAEs, the resulting mass and energy balances were used to size the equipment and estimate operating expenses. The distillation columns were sized as tray towers under vacuum conditions. The columns had constant diameter, which was estimated using a correlation based on internal vapor flow. A stage spacing of 1.5 ft was assumed. For SMB, each adsorption column had the same fixed diameter of 4m. This value is in accordance with other large-scale SMB designs found in literature [7,8]. The length of the columns was set as a free variable to be determined from the optimization. After sizing, the columns of distillation and SMB were both costed as empty vertical vessels [9].

The condensers and reboilers were sized as shell-and-tube heat exchangers. The area was estimated assuming counter-current operation and a constant overall heat transfer coefficient. A single stage liquid-ring pump was used to generate vacuum in the distillation columns and centrifugal pumps were used for pressure rise. All aforementioned correlations for purchased and installed costs were obtained from Seider et al. [9], except for the rotary valve, which was obtained from a case study of a large-scale petrochemical plant [10].

Besides utilities, which are mostly satisfied internally, the main operating expenses are related to the purchase of make-up ethanol and adsorbent. The unit cost of adsorbent was based on the cost of its synthesis raw materials. This cost was scaled to the biorefinery level using the six-tenths factor rule. The obtained price was \$20/kg and a replacement time of 5 years was assumed. Given the discrepancy in size between the lab bench and the biorefinery, there is an inherent degree of uncertainty in these values that could be addressed through sensitivity analysis. However, this was considered outside of the scope of this work.

Next, the “Discounted Cash Flow Analysis” model from NREL [11] was implemented in the NLP problem. This model is based on U.S. Department of Energy (DOE) recommendations for renewable energy systems. It includes values for interest rate, equity, depreciation, plant lifetime and taxes. The most important metric from the model is the minimum fuel selling price (MFSP) of the fuel. This is a key metric used to assess the viability of biofuels. DOE has set a MFSP target of \$2.50/GGE (gallon of gasoline equivalent) or less to be reached by 2030. With the current assumptions, the SOT reaches a MFSP value of \$2.47/GGE.

### Utilities of the Biorefinery

As mentioned, the main effect of BDO separation is reducing operating expenses of the biorefinery. Most notably, the broth is raised to 250°C and 60 atm prior to the dehydration reactor, and by removing water and reducing the BDO stream size, the energy load also significantly decreases. This saves around 90% of the high-

pressure steam (HPS) internally produced, which can instead be used to generate additional power in the turbine and reduce imported electricity. Altogether, this leads to a power output of 32.9 MW and a ~50% import reduction from the grid compared to the SOT. This value was corroborated by Liu et al. [12] in a previous study.

Next, heat duties associated with the separation scheme were considered. The distillation reboilers are the main consumers of energy in either pathway. These columns use low-pressure steam (LPS), which requires withdrawal from the turbine system. Without changes to the existing design, this withdrawal would reduce the turbine’s power output. Thus, as a simplification, natural gas is imported in order to generate additional steam needed to meet these duties and still maintain the turbine output of 32.9 MW.

### Greenhouse Gas Emissions System Boundary

The fuel’s lifecycle analysis was then considered, with GHG emissions (in  $g_{CO_2e}/MJ_{fuel}$ ) as the key metric. A technical report in 2020 was developed by the Argonne National Lab (ANL) [13], where the SOT’s well-to-wheel emissions were calculated. The system boundary in this report encompasses all processes between the corn stover collection in the field and the production of fuel and co-product from the biorefinery. The functional unit of the system is a GGE of fuel.

For this study GHGs were estimated using the displacement method. Through this method lifecycle emissions are attributed entirely to the fuel and none to the co-product. The fuel, however, receives credit for the emissions associated with the co-product’s conventional fossil pathway. This remains consistent with the economic model of the SOT, where adipic acid production revenue is credited to the fuel’s MFSP. The ANL 2020 report was used as a guideline and the GREET tool [14] was used to develop expressions for GHG emissions. The reader is encouraged to read this report for more details regarding the lifecycle assumptions.

### Optimization

Below are the expressions of the NLP problem. Pyomo and Ipopt were also used for the optimization.

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{y} \\ \text{s.t.} \quad & \mathbf{f}(\mathbf{u}, \mathbf{x};) = \mathbf{0} \\ & \mathbf{g}(\mathbf{u}, \mathbf{x};) \geq \mathbf{0} \\ & \mathbf{y} = \mathbf{h}(\mathbf{u}, \mathbf{x}; \theta) \\ & \mathbf{x}_{min} \leq \mathbf{x} \leq \mathbf{x}_{max} \end{aligned}$$

The objective function  $\mathbf{y}$  can be either MFSP or GHG minimization. The equality constraints,  $\mathbf{f}$ , are represented by the system of PDAEs and the discounted cash flow model. The only inequality constraints,  $\mathbf{g}$ , in this problem are the maximum allowable MFSP and minimum BDO purity post-distillation. The main decision variables,  $\mathbf{x}$ , of the SMB are the length of each adsorption column, the



velocity in each zone and step time. For distillation variables are pressure, reflux ratio and reboiler duty. With appropriate initialization, the stage location of each feed stream can also be determined without resorting to MINLP. The state variables are represented by  $u$  and the physical parameters by  $\theta$ .

## RESULTS

### Baseline GHG Emissions

Using GREET and the displacement method, it was found that the lifecycle GHG of the SOT's fuel product is 23.0 g<sub>CO<sub>2e</sub>}/MJ. From this value 121.4 g<sub>CO<sub>2e</sub>}/MJ is attributed to the fuel and 98.4 g<sub>CO<sub>2e</sub>}/MJ is credit from producing adipic acid. This represents a significant decrease from 93.0 g<sub>CO<sub>2e</sub>}/MJ for petroleum-derived diesel [14]. It should be noted that we assumed a carbon footprint of adipic acid (AA) of 4.3 kg<sub>CO<sub>2e</sub>}/kg<sub>AA</sub>. This does not consider GHG-100 emissions from nitrous oxide (N<sub>2</sub>O) and is a more conservative estimate than the ANL 2020 report, which assumed 11.6 kg<sub>CO<sub>2e</sub>}/kg<sub>AA</sub> and led to net emissions of -148 g<sub>CO<sub>2e</sub>}/MJ for the SOT.</sub></sub></sub></sub></sub></sub></sub>

Next, an expression was developed to relate GHG with the relevant mass and energy flows (natural gas import and make-up ethanol). First, GHG emissions were re-calculated based on the updated grid electricity import. As mentioned, turbine output increases from 11.8 MW to 32.9 MW, reducing import from 41.5 MW to 20.4 MW. This resulted in 103.3 g<sub>CO<sub>2e</sub>}/MJ attributed to fuel and after considering adipic acid credit, a net value of 4.9 g<sub>CO<sub>2e</sub>}/MJ was obtained. Then, natural gas (NG) and make-up ethanol (EtOH) import were varied in GREET to estimate how GHG changes, resulting in the following linear expression:</sub></sub>

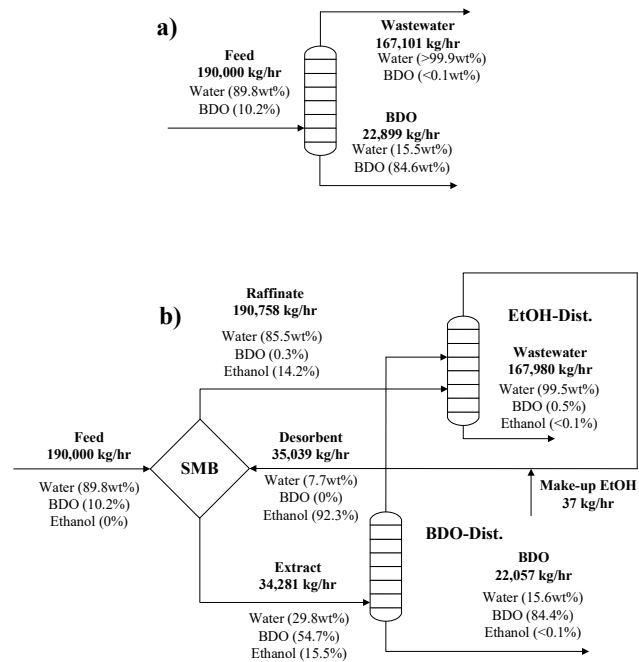
$$GHG = 4.9 \times \frac{1}{R} + 0.1362 \times NG + 0.0023 \times EtOH \quad (1)$$

R represents BDO recovery. This ranges from 0 to 1 and serves as a scaling factor to make-up for BDO lost during separation. GHG is expressed in g<sub>CO<sub>2e</sub>}/MJ, NG in MMBTU/hr and EtOH in kg/hr. This was the expression used in the objective function for minimizing GHG. It should be clarified that NG and EtOH are the added imports that result from incorporating the separation to the biorefinery. Natural gas (58 MMBTU/hr) and ethanol (37 kg/hr) are both imported in the SOT, but those are already accounted in the 4.9 g<sub>CO<sub>2e</sub>}/MJ calculation.</sub></sub>

### Approach Implementation

The proposed approach was implemented and for both pathways, the binary feed is 10wt% BDO (~100g/L titer), while the minimum purity requirement of the exit stream is 85 wt% BDO. **Figure 5a** shows the mass balances of the optimal configuration of the single distillation pathway. This configuration achieved a MFSP<sub>dist</sub> of \$2.54/GGE. **Figure 5b** shows the obtained optimal

configuration for the SMB pathway. This configuration was obtained with a MFSP<sub>SMB</sub> <= \$2.54/GGE constraint.



**Figure 5.** Resulting optimal process designs of: a) the single distillation column and b) SMB + distillation

The optimization problems were solved using the true moving bed (TMB) approximation model for the SMB and an IPOPT convergence tolerance of  $1 \times 10^{-3}$ . It should also be noted that the distillate of the EtOH-dist column is ~92 wt% ethanol. Thus, the ethanol/water azeotrope at this vacuum pressure (~96 wt%) is not expected to be encountered.

### Comparison of the Proposed Pathways

**Table 1** shows the comparison between the distillation columns for both pathways. As seen the most significant difference is the duty of the reboiler and condenser units for each pathway. Most notably, it is shown that the combined reboiler duty from the two columns in the SMB pathway is much lower than single distillation. This further establishes adsorption as a far more effective separation alternative in this context.

In all cases a lower boundary of 0.10 bar was imposed to the column pressure. For single distillation, the optimal pressure was slightly higher at 0.22 bar, since this raises the temperature of the wastewater from the tops, which was used for heat integration with the feed.

**Table 1.** Results operating conditions of the single distillation, and BDO and EtOH recovery columns

Distillation Column	Single Dist.	BDO Dist.	EtOH Dist.
Number of Stages	20	30	30
Feed Stage(s)	19	29	10/18
Pressure (bar)	0.22	0.10	0.30
Reflux Ratio (mole)	0.06	0.02	1.77
Reboiler Duty (MW)	118.2	6.5	24.3
Condenser Duty (MW)	116.2	6.2	26.5

**Table 2** shows the comparison between the alternatives evaluated in this study. The ideal separation represents the SOT with BDO recovery, but without incurring in capital or operating costs yet (i.e., the maximum potential of BDO separation [12]).

**Table 2.** Results operating conditions of the single distillation, and BDO and EtOH recovery columns

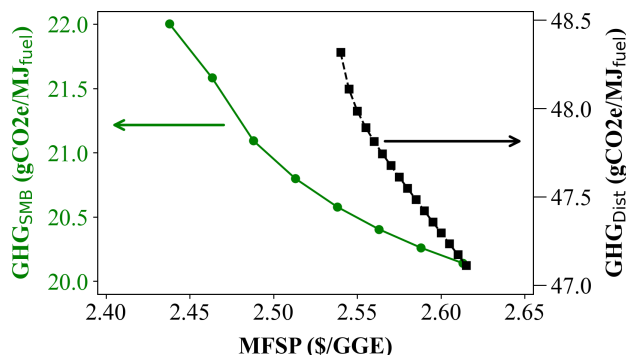
Process	SOT	Ideal Sep.	Single Dist.	SMB+ Dist.
Natural Gas Import (MMBTU/hr)	58	58	379	174
Grid Import (MW)	41.5	20.4	20.4	20.4
MFSP (\$/GGE)	2.47	2.11	2.54	2.54
GHG (gCO <sub>2e</sub> /MJ)	23.0	4.9	48.7	20.5

Single distillation, as expected, had poor economic and sustainability performance. Its large MFSP and GHG values render the entire process unfeasible, and thus, simple distillation cannot be considered a viable alternative to recover BDO in this biorefinery. In comparison, SMB+distillation gave lower GHG, demonstrating the effectiveness of the proposed hybrid separation and its potential to be incorporated in the biorefinery.

### Sensitivity Analyses

In the proposed approach, either MFSP or GHG is minimized, but not both simultaneously. This does not provide a complete outlook and Pareto fronts were built to allow a more thorough comparison. **Figure 6** shows the Pareto optimal curves for the single distillation and SMB processes. In the case of single distillation, the leftmost point represents the result from minimizing MFSP without GHG restriction, while the rightmost point is the minimization of GHG without MFSP restrictions.

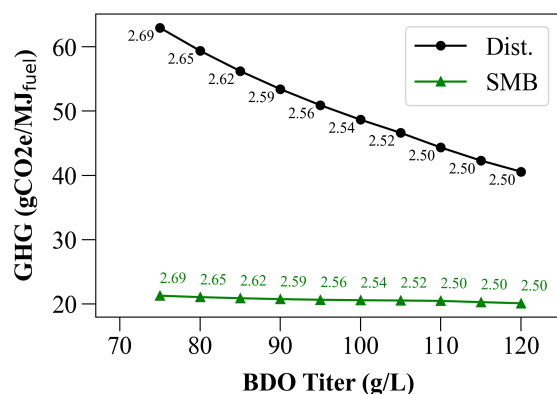
For the SMB process, the leftmost point is also MFSP minimization without GHG constraint. However, the rightmost end is not bounded, given that the SMB can be as large as needed. This indefinitely increases the quantity of adsorbent, decreasing desorbent use and energy load. Thus, GHG also decrease indefinitely but at the expense of a prohibitive adsorbent purchasing cost.



**Figure 6.** Pareto Fronts of the Separation Alternatives

It should be noted, that for either front, GHG is not very sensitive to the MFSP. The most important observation, however, is that the SMB-process can improve the SOT's economic and sustainability performance. As shown, for any MFSP under \$2.47/GGE, GHG remains below 23 gCO<sub>2e</sub>/MJ, proving again the viability of this separation alternative.

Finally, a sensitivity analysis was developed by varying BDO titer. The fermentation in the SOT is an ongoing research topic and is not clear that a titer of 100g/L is optimal for the overall flowsheet. **Figure 7** shows the effect of BDO titer on GHG and MFSP by applying the proposed approach. As seen, GHG emissions of single distillation considerably increase with decrease in titer. This is an expected result, given that a large and dilute stream leads to a higher energy demand that increases both GHG and MFSP.



**Figure 7.** Effect of BDO titer concentration on the GHG emissions for both separation pathways (the annotated values represent the MFSP at each point in \$/GGE)

Remarkably, however, the change of GHG emissions in SMB is almost negligible. To understand this better we can further explore the 75 g/L titer point. After optimizing the NLP problem with GHG as the objective function, the obtained MFSP is \$2.69/GGE, and the quantity of adsorbent is 1392 tonnes. However, the at this titer level MFSP minimization gives \$2.46/GGE, and a much lower

adsorbent quantity of 342 tonnes. Implementing a very large quantity of adsorbent is clearly a suboptimal economic decision, but it greatly reduces desorbent use and energy duties, all of which help improve GHG. Thus, SMB size is allowed to increase almost indefinitely in order to accommodate varying titers, and only marginal changes in GHG are observed.

## CONCLUSIONS

A comparative study was performed between two separation technologies to recover BDO from a fermentation broth. Single distillation was compared to a novel SMB+distillation separation scheme on the basis of GHG emissions. Both separations were incorporated in a biorefinery design developed by NREL that produces renewable fuels with corn stover as feedstock and dilute (10wt%) BDO as an intermediate.

The comparison was done by following a stepwise optimization approach where GHG is minimized. The resulting design achieved GHG emissions of 20 g<sub>CO<sub>2e</sub></sub>/MJ for SMB+distillation, lower than 48 g<sub>CO<sub>2e</sub></sub>/MJ for single distillation and 93 g<sub>CO<sub>2e</sub></sub>/MJ for petroleum fuel.

The effect of BDO fermentation titer was also considered. As expected, for single distillation, decreasing titer significantly increased energy consumption and, thus, worsened MFSP and GHG emissions. However, for SMB, the GHG emissions were not sensitive to titer.

## ACKNOWLEDGEMENTS

This study was supported by the U.S. Department of Energy (DOE) Bioenergy Technology Office (BETO) through the DOE Award Number: DE-EE0009263

## REFERENCES

1. Haider J, Harvianto GR, Qyyum MA, Lee, M. Cost- and energy-efficient butanol-based extraction-assisted distillation designs for purification of 2,3-butanediol for use as a drop-in fuel. *ACS Sustain. Chem. Eng.* 6:14901-14910 (2018).
2. Davis, RE, Grundl NJ, Tao L, Bidy MJ, Tan EC, Beckham GT, Humbird D, Thompson DN, Roni, MS. Process design and economics for the conversion of lignocellulosic biomass to hydrocarbon fuels and coproducts: 2018 biochemical design case update; biochemical deconstruction and conversion of biomass to fuels and products via integrated biorefinery pathways. United States (2018) <https://doi.org/10.2172/1483234ase>
3. Minceva, M. Rodrigues, AE. Two-level optimization of an existing SMB for p-xylene separation. *Comput Chem Eng* 29:2215-2228 (2005).
4. Nicholson B, Siirola JD, Watson JP, Zavala VM,

Biegler, LT. pyomo.dae: a modeling and automatic discretization framework for optimization with differential and algebraic equations. *Math. Prog. Comp.* 10: 187-223 (2018).

5. Data from: Aspen Properties® Aspen Technology Inc. 2022. Accessed April 1, 2023.
6. Wachter A, Biegler, LT. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Program.* 106:25-57 (2006).
7. Chiang YD, Bassas HN, Lively RP, Nair S. Separation and Purification of 2,5-Dimethylfuran: Process Design and Comparative Technoeconomic and Sustainability Evaluation of Simulated Moving Bed Adsorption and Conventional Distillation. *ACS Sustain. Chem. Eng.* 8:12482-12492 (2020).
8. Chiang YD, Liang WW, Yang SW, Bond CR, You WR, Lively RP, Nair S. Separation and purification of furans from n-butanol by zeolitic imidazole frameworks: Multicomponent Adsorption Behavior and Simulated Moving Bed Process Design. *ACS Sustain. Chem. Eng.* 7:16560-16568 (2019).
9. Seider WD, Lewin DR, Seader JD, Widagdo S, Gani R, Ng KM. Product and process design principles: synthesis, analysis and evaluation. Wiley (2016).
10. Do TX, Lim Y, Lee J, Lee W. Techno-economic analysis of petrochemical complex retrofitted with simulated moving-bed for olefins and aromatics production. *Chem Eng Res Des* 106: 222-241 (2016).
11. Humbird D, Davis R, Tao L, Kinchin C, Hsu D, Aden A, Schoen P, Lukas J, Olthof B, Worley M, Sexton D, Dudgeon D. Process design and economics for biochemical conversion of lignocellulosic biomass to ethanol: dilute-acid pretreatment and enzymatic hydrolysis of corn stover. United States (2011) Web. doi:10.2172/1013269.
12. Liu J, Dempsey JW, Li S, Jiang Y, Snowden-Swan LJ, Kubic WL, Tan EC, Freeman CJ. Methodology for assessing the maximum potential impact of separations opportunities in industrial processes. *Front. Sustain* 3: (2022).
13. Cai H, Ou L, Wang M, Tan EC, Davis RE, Dutta A, Tao L, Hartley D, Roni MS, Thompson DN. Supply chain sustainability analysis of renewable hydrocarbon fuels via indirect liquefaction, ex situ catalytic fast pyrolysis, hydrothermal liquefaction, combined algal processing, and biochemical conversion: update of the 2019 state-of-technology cases. United States (2020).
14. Wang MQ. GREET 1.0 -- Transportation fuel cycles model: methodology and use. United States (1996) Web doi:10.2172/230197

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative

commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Design and Optimization of Methanol Production using PyBOUND

Prapatsorn Borisut<sup>a,b</sup>, Bianca Williams<sup>a</sup>, Aroonsri Nuchitprasittichai<sup>b</sup>, and Selen Cremaschi<sup>a\*</sup>

<sup>a</sup> Department of Chemical Engineering, Auburn University, Auburn, AL, United States

<sup>b</sup> School of Chemical Engineering, Suranaree University of Technology, Nakhon Ratchasima, Thailand

\* Corresponding Author: [szc0113@auburn.edu](mailto:szc0113@auburn.edu).

---

## ABSTRACT

In this paper, we study the design optimization of methanol production with the goal of minimizing methanol production cost. One challenge of methanol production via carbon dioxide (CO<sub>2</sub>) hydrogenation is the reduction of operating costs. The simulation of methanol production is implemented within the Aspen HYSYS simulator. The feeds are pure hydrogen and captured CO<sub>2</sub>. The process simulation involves a single reactor and incorporates recycling at a ratio of 0.995. The methanol production cost is determined using an economic analysis. The cost includes capital and operating costs, which are determined through the equations and data from the capital equipment-costing program. The decision variables are the pressure and temperature of the reactor contents. The optimization problem is solved using a derivative-free algorithm, pyBOUND, a Python-based black-box model optimization algorithm that uses random forests (RFs) and multivariate adaptive regression splines (MARS). The predicted minimum methanol production cost by pyBOUND is \$1396.56 per tonne of methanol, which corresponds to the pressure of 68.82 bar and temperature of 192.23°C while the actual cost is \$1393.95 per tonne of methanol at these conditions. The cost breakdown of methanol production is 75% hydrogen price, 11% utility cost, 8% capital cost, 5% carbon dioxide price, and 1% operating cost.

---

**Keywords:** Methanol, Carbon Dioxide, Process Synthesis, Process Design, Simulation, Optimization, pyBOUND

## INTRODUCTION

Methanol is utilized as a feedstock for power generation, transportation fuel, and wastewater treatment, as well as for producing various value-added chemicals [1]. Methanol is environmentally friendly, and the use of methanol has lower risks, lower emissions, and higher performance than gasoline [2]. The data from the methanol market indicate that the methanol market is expected to be worth \$38.0 billion by 2028, growing at a compound annual growth rate of 4.2% during the forecast period [3]. Formaldehyde production from methanol is expected to be the largest demand in the forecasted period [3].

Methanol can be produced by two main processes: methanol from syngas and methanol from CO<sub>2</sub> hydrogenation. Traditionally, methanol is produced from syngas. The first commercial methanol plant was established in 1923 by BASF. Conventionally, methanol is produced

from petroleum products via hydrogenation of CO and CO<sub>2</sub> and reverse water gas shift reaction. In the commercial production of methanol from syngas, a catalyst known as CuO/ZnO/Al<sub>2</sub>O<sub>3</sub> is widely employed [4-5]. The choice of copper-based catalysts for converting CO<sub>2</sub> to methanol is primarily driven by their affordability and high catalyst activity [6]. Nowadays, 80% of worldwide methanol is produced from natural gas [7].

Finding new and more sustainable energy sources is crucial to address the environmental crisis caused by the widespread use of fossil fuels. The largest source of CO<sub>2</sub> emissions among human activities that contribute to greenhouse gases (GHGs) is burning fuels for power and heat, accounting for more than 42% of the estimated anthropogenic CO<sub>2</sub> emissions [8]. Carbon capture utilization and storage is the most common technique to reduce anthropogenic CO<sub>2</sub> emissions. Methanol production from CO<sub>2</sub> hydrogenation (using CO<sub>2</sub> captured and pure hydrogen) is considered an important route to reduce CO<sub>2</sub>



emissions [9]. The impact of direct CO<sub>2</sub> hydrogenation to methanol on the environment is lower than production of methanol from syngas [10]. Many studies focus on the optimization of CO<sub>2</sub> hydrogenation process and present various heat integration within the methanol conversion unit [11-12]. Yousaf et al. summarized computational studies on techno-economic analysis of CO<sub>2</sub> hydrogenation [13]. Most studies on the methanol production design using capture CO<sub>2</sub> and pure hydrogen attempt to reduce the production cost due to the high hydrogen price.

We optimize methanol production via CO<sub>2</sub> hydrogenation in this work. The goal is to minimize the methanol production cost by adjusting the pressure and temperature of the reactor contents. The hydrogenation process is in Aspen HYSYS, and the optimization problem is solved using pyBOUND, a Python-based black-box optimization algorithm using random forests [14]. The paper is organized as follows: The next section provides an overview of the general pyBOUND framework. The Case study section provides information on the simulation and economic analysis of the CO<sub>2</sub> hydrogenation process. The optimization formulation section describes an objective function and constraints of the case study. The Results and Discussion section presents the optimum operating conditions and costs. Our concluding remarks and future directions are summarized in the last section.

## A PYTHON-BASED BLACK-BOX OPTIMIZATION USING RANDOM FORESTS, PYBOUND [14]

pyBOUND is a two-stage algorithm. The first stage employs random forests (RFs) and generates better decision variable bounds, shrinking the search space. The second stage refines the solution using multivariate adaptive regression splines (MARS). A general pyBOUND framework is depicted in Figure 1.

### pyBOUND Stage 1: Shrinking the Search space using RFs

For the first stage, the algorithm reduces the size of the original search space by solving a series of global deterministic subproblems using RF models. This stage consists of sampling, construction of an RF model, global optimization of the constrained approximation problem, and collection of new sampling points. The initial data set is generated using a Sobol sequence design. The overall constrained approximation model, which consists of the RFs model objective equation, the RF model constraints, and variable bounds, is solved to optimality using deterministic global optimization methods. The decision variable bounds are generated from the solution of the constrained approximation optimization problem. The constrained approximation models with the RF models are

MILPs, which are solved with CPLEX (version 22.1.1).

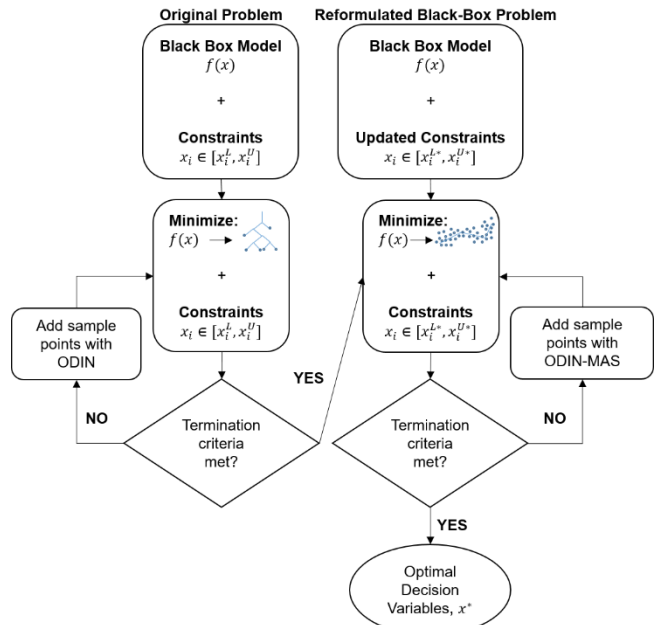


Figure 1. General pyBOUND framework [14].

The algorithm determines a single set of bounds based on the threshold bounds given by every RF tree. The three methods for reducing variable bounds called the cutting methods were studied for bounds reduction. The methods are selecting the widest set of bounds (Wide), selecting the bounds where at least two trees intersect each other (Intersection), and averaging the bounds given by each tree for a single set of bounds (Average). The “Intersection” cutting method is selected for optimization problems with less than ten decision variables, and the “Average” cutting method is selected and implemented for problems with ten or more decision variables.

The ODIN sampling method [15] is selected to update the RF model training set with a termination criterion that depends on the size of the decision variable vector and maximum reduction in the search space.

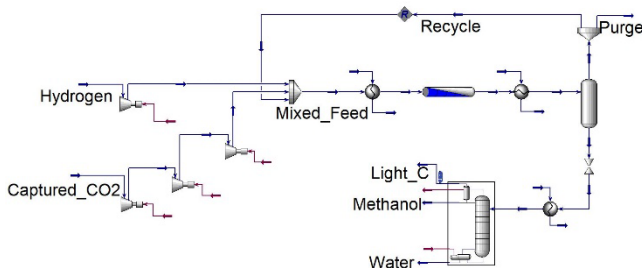
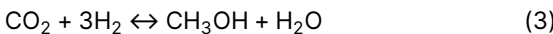
### pyBOUND Stage 2: Refinement of Solution using MARS Models

This stage refines the solution with a local search in the reduced search space, constructs the model, and improves the model with adaptive sampling. Based on the final form of the trained MARS surrogate model, a deterministic optimization problem is formulated and solved to refine the solution. MARS models result in MINLP optimization models, which are solved with ANTIGONE [16]. New sample points are added in the second stage using adaptive sampling on a hybrid of the ODIN sampling method and Mixed Adaptive Sampling (MAS) [17], referred to as ODIN-MAS [15].

## CASE STUDY: METHANOL PRODUCTION

### Process Simulation

The process flow diagram for methanol production implemented within the Aspen HYSYS simulator (version 14) is depicted in Figure 2. The feeds of the process are pure hydrogen and captured CO<sub>2</sub> with a molar flow rate of 3000 kmoles per hour and 1000 kmoles per hour, respectively. Pure hydrogen is fed at 30 bar and 25°C and then compressed to the desired pressure. Captured CO<sub>2</sub> is fed at 2 bar and 28°C through the three-stage compressors to increase pressure before being combined with pure hydrogen. The mixed feed is heated up and enters the reactor, which is modeled as a plug flow reactor and catalytic reaction. The reactions of methanol production via CO<sub>2</sub> hydrogenation follows Equations (1) – (3). The reactor is simulated using a kinetic model [18]. After that, the stream of produced methanol and unreacted feed enters the cooler before it is separated. The vapor stream is recycled and purged with a 0.995 recycle ratio. The liquid phase stream is sent to the distillation column after cooling. The specifications of the distillation column are a 1.52 reflux ratio and 40°C of distillate temperature. The methanol product purity is specified at 99.1% by weight.



**Figure 2.** Methanol production via CO<sub>2</sub> hydrogenation.

### Economic Evaluation

The methanol production cost is determined using an economic analysis. The cost includes both capital and operating costs. These costs are determined through the equations and data from the capital equipment-costing program [19]. The data was adjusted for inflation from the year 2001-2023 by using values of the Chemical Engineering Plant Cost Index, CEPCI. The value of CEPCI in 2001 and 2023 is 297 and 799.1, respectively [20]. The capital cost is calculated using the costing technique. This technique depends on the specific equipment type, the specific system pressure, and the specific material of

construction. In this work, the equipment costs are estimated based on gross root cost ( $C_{GR}$ ), which involves total module cost ( $C_{TM}$ ), Equations (4)-(5), where  $n$  represents the total number of equipment, and bare module cost ( $C_{BM}$ ), Equation (6).  $F_M$  and  $F_p$  are the material factor and pressure factor, respectively.

$$C_{GR} = C_{TM} + 0.50 \sum_{i=1}^n C_{BM,i}^0 \quad (4)$$

$$C_{TM} = \sum_{i=1}^n C_{TM,i} + 1.18 \sum_{i=1}^n C_{BM,i} \quad (5)$$

$$C_{BM} = C_p^0 F_{BM} = C_p^0 [B_1 + B_2 F_p F_M] \quad (6)$$

where  $C_p^0$  is the purchased cost for a base condition which can be calculated following Equation (7).

$$\log_{10} C_p^0 = K_1 + K_2 \log_{10}(A) + C_3 (\log_{10}(A))^2 \quad (7)$$

$A$  is the capacity or size parameter for the equipment. The overall assumptions used for the cost estimation are as follows:

- The plant operates for 8400 hours per year.
- The plant is expected to have 20 years of plant operation and a 10% interest rate.
- Direct supervisory and clerical labor are taken 18% of labor cost.
- Maintenance and repairs are taken 6% of the fixed capital investment.
- Operating supplies are taken 0.9% of the fixed capital investment.
- Laboratory charges are taken 15% of labor cost.
- Local taxes and insurance are taken at 3.2% of the fixed capital investment.
- Plant overhead cost is the summation of 70.8% of labor cost and 3.6% of the fixed capital investment.
- Administration cost is the summation of 17.7% of labor cost and 0.9% of the fixed capital cost.

### OPTIMIZATION FORMULATION

We consider the methanol production via CO<sub>2</sub> hydrogenation as a black-box optimization problem of the form given in Equations (8)-(9),

$$\min_x f(x) \quad (8)$$

$$\text{subject to } x_i \in [x_i^L, x_i^U] \quad \forall i \in \{1,2\} \quad (9)$$

The objective of the optimization problem is to minimize the methanol production cost where  $f(x)$  represents the methanol production cost with the decision

variables of the process with finite bounds  $[x_i^L, x_i^U]$ ,  $x_i$  represents the decision variables:  $x_1$  is the pressure and  $x_2$  is the temperature of the reactor contents. The range of each decision variable is given in Table 1.

**Table 1:** Initial range of decision variables.

Decision variables	Lower bound	Upper bound
Pressure (bar)	50	80
Temperature (°C)	190	210

## RESULTS AND DISCUSSION

The goal is to minimize methanol production cost by applying a derivative-free optimization algorithm, pyBOUND. From the results of the RF stage, the “Intersection” cutting method is used to shrink the search space of the optimization problem. The results at each iteration of the RF stage are shown in Table 2.

**Table 2:** The reduction of search space.

Iteration	1	2	3
Volume of reduction (%)	53.13	72.90	80.13
Number of sample points	20	15	15
Lower bound	50, 190	62.66, 190	62.66, 190
Upper bound	80, 199.38	80, 199.38	80, 196.88

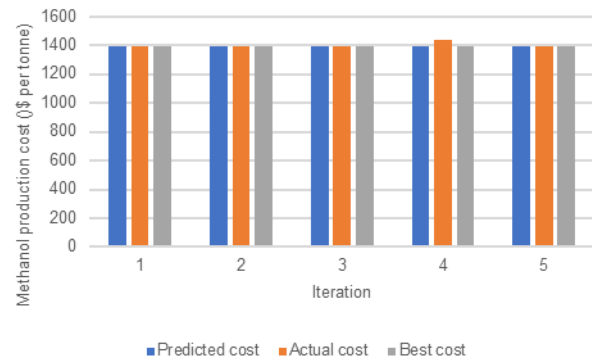
The criterion for the termination of the RF stage is specified as the maximum volume reduction of search space of 0.75. The bound of the pressure is reduced to the range 62.66 – 80 bar and the bound for temperature is reduced to the range 190 – 196.88°C. Eighty percent of the search space is reduced compared to the original search space. The third row of Table 2 shows the number of sample points. Originally, the number of sample points was generated with ten times the number of decision variables (20 sample points). After applying pyBOUND, the final number of sample points in the RF stage is 15, while the number of function evaluations is 30, which is due to some sample points being cut during the RF stage iterations. Fifteen sample points are then used as the initial sample set for the MARS stage.

The results of the MARS stage at each iteration are shown in Table 3. Figure 3 also shows the results of the predicted, actual, and best methanol production cost at each iteration. The best methanol production cost is the lowest cost in the training data set. In the MARS stage, the sample points are added at each iteration using ODIN-MAS algorithm. The pyBOUND algorithm stopped at 144 sample points while the R-squared value of the

model increased after increasing the number of sample points. For the first iteration, the model can not fit well with only twenty sample points. The value of R-squared at the first iteration shows a negative value. At the second iteration, the total number of sample points used to construct the MARS model increased to 37, and the value of R-squared is a positive value.

**Table 3:** MARS model and optimal results

Iteration	Function evaluation	R-squared	Predicted optimum location (bar, °C)
1	20	-0.05	68.84, 192.26
2	37	0.26	68.80, 192.34
3	58	0.50	68.68, 192.24
4	90	0.86	72.51, 192.32
5	144	0.88	68.82, 192.23

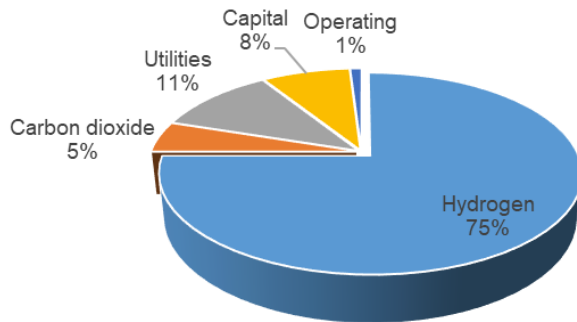


**Figure 3.** The values of minimum methanol production cost at each iteration.

The predicted methanol production costs at each iteration are close to the actual methanol production costs. Finally, the pyBOUND algorithm stopped with the criteria of R-squared value more than 0.6 and the Euclidean distance between the best objective values of the last three iterations less than 0.0004. The error between predicted and actual methanol production costs is 0.19%. The optimum methanol production cost is \$1396.56 per tonne of methanol, corresponding to a pressure of 68.82 bar and a temperature of 190.23°C, while the actual methanol production cost at these conditions is \$1393.95 per tonne of methanol. We found that the best objective value is \$1393.80 per tonne of methanol. This value is close to the actual objective value and the best location is similar to the predicted optimal location. The Euclidean distance between the best and actual objective values is 0.02 at the last iteration. The algorithm run time is 458.5 min, which is divided into 79.92 min for the RF stage and 378.58 min for the MARS stage. From the total run time, the time for pyBOUND is 17.32 min, with the remaining

time spent by Aspen HYSYS process simulation.

At the optimum, the contribution of various factors to the cost is illustrated in Figure 4. The contribution of raw materials, including CO<sub>2</sub> and H<sub>2</sub> prices, are 5% and 75%, respectively. The utilities make up 11% of the overall cost. The capital and other operating expenses are 7% and 1%, respectively. The H<sub>2</sub> price is identified to be the major cost of this process.



**Figure 4.** Cost breakdown.

## CONCLUSION AND FUTURE WORK

The pyBOUND algorithm has been successfully applied to the optimization of methanol production via CO<sub>2</sub> hydrogenation process. The minimum cost of methanol production is determined. The algorithm, pyBOUND, demonstrates that the algorithm can estimate the optimum value of the methanol production cost with a small number of sample points. Future work on pyBOUND will focus on applying the algorithm to problems with high number of decision variables and several chemical engineering problems. The optimum value obtained by pyBOUND is comparable to other optimization methods. pyBOUND might exhibit improvements when compared to existing algorithms for both high dimensionality problems and for locating the true optimal decision variable values. We will consider the hyperparameters of the RF stage for consistency of generating variable bounds and investigate the impact of the stopping criteria of the pyBOUND algorithm.

## ACKNOWLEDGEMENTS

The authors would like to acknowledge the Department of Chemical Engineering, Auburn University for providing the facilities and software to support this work. The author, Prapatsorn Borisut, was funded by the National Research Council of Thailand (NRCT) via the Royal Golden Jubilee (RGJ) Ph.D. scholarship, batch number 21, and grant number PHD/0214/2561 for the student intern program.

## REFERENCES

1. Dalena, F., Senatore, A., Marino, A., Gordano, A., Basile, A. Chapter 1 – Methanol production and applications: an overview, in *Methanol*, A. Basile and F. Dalena, Editors. Elsevier. 3-28 (2018)
2. Bill, A. Carbon dioxide hydrogenation to methanol at low pressure and temperature. [Ph.D. Thesis, EPFL, Lausanne, Switzerland]. (1998) <https://infoscience.epfl.ch/record/32203>
3. MARKETSANDMARKETS. Methanol Market by Feedstock (Natural Gas, Coal), Derivative (Formaldehyde, MTO/MTP, Gasoline, MTBE, MMA, Acetic Acid, DME, Biodiesel), Sub-Derivative, End-use Industry (Automotive, Construction, Electronics), and Region – Global Forecasts to 2028. <https://www.marketsandmarkets.com/Market-Reports/methanol-market-425.html>
4. Sayah, A. K., Hosseinabadi, S.H., Farazar, M. CO<sub>2</sub> abatement by methanol production from flue-gas in methanol plant. *World Acad. Sci. Eng. Techno. Int. J. Chem. Molecul. Eng.* 4:9 (2010)
5. Jadhav, S. G., Vaidya, P. D., Bhanage, B. M., Joshi, J. B. Catalytic carbon dioxide hydrogenation to methanol: a review of recent studies. *Chem Eng Res Design* 92:2557–2567 (2014)
6. Ali, K. A., Abdullah, A. Z., Mohamed, A. R. Recent development in catalytic technologies for methanol synthesis from renewable sources: a critical review. *Renew. Sustain. Energ. Rev.* 44:508–518 (2015)
7. Chein, R.-Y, Chen, W.-H., Chyuan Ong, H., Loke Show, P., Singh, Y. Analysis of methanol synthesis using CO<sub>2</sub> hydrogenation and syngas produced from biogas-based reforming processes. *Chemical Engineering Journal* 426:130835 (2021)
8. Meunie, N., Chauvy, R., Mouhoubi, S., Thomas, D., De Weireld, G. Alternative production of methanol from industrial CO<sub>2</sub>. *Renewable Energy* 146:1192–1203 (2020)
9. Leonzio, G., E. Zondervan, P.U. Foscolo. Methanol production by CO<sub>2</sub> hydrogenation: Analysis and simulation of reactor performance. *International Journal of Hydrogen Energy* 44(16):7915-7933 (2019)
10. Artz, J., Müller, T.E., Thenert, K., Kleinekorte, J., Meys, R., Sternberg, A., Bardow, A., Leitner, W. Sustainable Conversion of Carbon Dioxide: An Integrated Review of Catalysis and Life Cycle Assessment. *Chemical Reviews* 118(2):434-504 (2018)
11. Pérez-Fortes, M., Schöneberger, J. C., Boulamanti, A., Tzimas, E. Methanol synthesis using captured CO<sub>2</sub> as raw material: Techno-economic and environmental assessment. *Applied Energy* 161:718-732 (2016)

12. Abdelaziz, O.Y., Hosny, W. M., Gadalla, M. A., Ashour, F. H., Ashour, I. A., Hulteberg, C. P. Novel process technologies for conversion of carbon dioxide from industrial flue gas streams into methanol. *Journal of CO<sub>2</sub> Utilization* 21:52-63 (2017)
13. Yousaf, M., A. Mahmood, A. Elkamel, M. Rizwan, M. Zaman. Techno-economic analysis of integrated hydrogen and methanol production process by CO<sub>2</sub> hydrogenation. *International Journal of Greenhouse Gas Control* 115:103615 (2022)
14. Williams, B. From Data to Decisions: Development of Surrogate Models for Process Optimization [Doctoral dissertation, Auburn University]. <https://etd.auburn.edu//handle/10415/8348>
15. Smith, J. Computationally assisted biofuel production: hydrodynamics, optimization, and heuristics. [Doctoral dissertation, The University of Tulsa]. <https://www.proquest.com/docview/1761846406>
16. Misener, R., C.A. Floudas, ANTIGONE: Algorithms for continuous/integer global optimization of nonlinear equations. *Journal of Global Optimization* 59(2):503-526 (2014)
17. Eason, J., S. Cremaschi. Adaptive sequential sampling for surrogate model generation with artificial neural networks. *Computers & Chemical Engineering* 68:220-232 (2014)
18. Van-Dal, É. S., C. Bouallou. Design and simulation of a methanol production plant from CO<sub>2</sub> hydrogenation. *Journal of Cleaner Production* 57:38-45 (2013)
19. Turton R, Shaeiwitz JA, Bhattacharyya D, Whiting WB. Analysis, Synthesis, and Design of Chemical Processes. Prentice Hall (2012)
20. Charles M. <https://toweringskills.com/financial-analysis/cost-indices/>

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Biomanufacturing in Space: New Concepts and Paradigms for Process Design

Brenda Cansino-Loeza<sup>a</sup>, Vernon McIntosh<sup>b</sup>, Krista Ternus<sup>b</sup>, Victor M. Zavala<sup>a\*</sup>

<sup>a</sup>University of Wisconsin-Madison, Department of Chemical and Biological Engineering, Madison, WI, United States of America

<sup>b</sup>Signature Science, LLC, Austin, TX, United States

\* Corresponding Author: [zavatejeda@wisc.edu](mailto:zavatejeda@wisc.edu).

---

## ABSTRACT

One of the main challenges to support life in space is the development of sustainable, circular processes that reduce the high cost of resupply missions. Space biomanufacturing is an emerging paradigm that aims to reduce the need for resources, enabling on-demand manufacture of products. The cost of installing biomanufacturing systems in space depends on the cost of transporting the system components, which is directly proportional to their mass/weight. From this perspective, the system mass is a critical factor that dictates process design, and this has important implications in how we can approach such design. For instance, mass constraints require circular use of resources and tight process integration (to minimize resupply) and restricts the type of resources and equipment needed. In this work, we evaluate the lactic acid bioproduction design using *Escherichia coli*, *Saccharomyces cerevisiae*, and *Pichia pastoris*. We use the Equivalent System Mass (ESM) metric as a key design measure. ESM allows the quantification of different physical properties of the system in a common mass basis. Our analysis reveals that 97.7 kg/year of lactic acid can be produced using *Saccharomyces cerevisiae* in a 10 L stainless steel fermenter. Furthermore, considering that stainless steel is the design material and quantifying the mass of 1 g/cm<sup>2</sup> of shielding material, the total system mass was 19 kg. This modeling framework also identified the critical system elements responsible for the highest system mass and launch cost. Overall, our analysis reveals how focusing on system mass can bring new design perspectives that can aid the design of traditional manufacturing systems.

---

**Keywords:** Space manufacturing, Equivalent System Mass, Process Design, Circularity, Sustainability

## INTRODUCTION

Since 2000, the International Space Station (ISS) has served as a research laboratory in low Earth orbit, supporting several research studies in different scientific fields [1]. During these years, space-based research that study systems under unique, non-terrestrial conditions have produced strong impacts and diverse technological and social benefits. However, the cost of supporting the ISS is still an important factor that influences the development of new projects. The ISS Program's annual budget is about \$3 billion, of which, \$1.8 billion is related to the transportation of crew and cargo services making this the most expensive element of the program [2]. This is mainly because fuel is the dominant cost of space missions [3], and although lower launch costs have been

achieved in recent years, the current price for *delivering one kilogram* of cargo to low Earth orbit *is approximately \$10,000*, which is a high transportation cost [4]. Therefore, one of the main challenges in current space research is to create sustainable, closed-loop (circular) systems for future long-space duration missions to reduce the need for resupply missions.

Water, oxygen, energy, and food are some essential resources to sustain life in space but launch capacity for the ISS is approximately 1 ton of water [5]. Consequently, there is a need for flexible and low-cost space manufacturing processes that enable on-site production of essential chemicals and resources [6]. Over the past few years, research has been carried out to develop advanced technologies that convert human waste and local resources into pure consumables to satisfy the capacity

launch limitations and reduce the high cost of resupply missions. In this regard, in-situ resource utilization has offered new possibilities to develop circular processes that allow for the on-demand production of value-added chemicals and materials for the construction, maintenance, and repair of mission systems [7]. This is crucial in space exploration, because long-duration space missions require circular systems that are capable of manufacturing their essential items [8]. Furthermore, biomanufacturing is a promising approach to support sustainable and circular systems for space exploration [9]. This relies on the use of biological systems that are engineered to produce and manufacture value-added products and objects on demand. Hence, biomanufacturing is a convenient alternative to design systems in remote locations or where supply chains for consumables cannot operate reliably [10].

The use of biomanufacturing in space has diverse applications. Significant research has focused on regenerative medicine applications (e.g., tissue engineering and disease modeling) [11]. Lately, the production of nutrients and food through biomanufacturing has been crucial to support life in space. These studies have focused solely on technical and operational features, without considering transportation cost restrictions (which are critical to implement systems in space).

The traditional paradigm for process design is objective-driven, in the sense that the system is engineered to maximize/minimize specific performance goals. When designing a system in space, however, the mass/size of the system is a critical factor that takes priority due to strict limitations of transportation and resources. As such, the design of space systems is *constraint-driven*. Moreover, space systems are subject to disturbances/externalities (e.g., gravity and cosmic radiation) that pose non-obvious and often unknown constraints. For instance, the fermentation reactor (a key component of a biomanufacturing system), is strongly affected by storage conditions, gravity, and radiation, which can affect the performance and survival of microorganisms employed to produce value-added products.

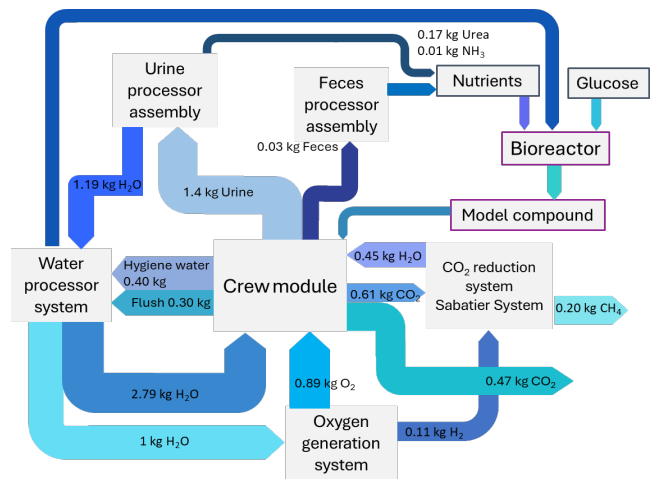
Considering this, new approaches are needed to support the design of biomanufacturing systems that consider both operational restrictions and mass restrictions. One approach to solve this problem involves the use of the Equivalent Systems Mass (ESM) analysis. ESM is a metric that allows the consideration of different physical resources such as power, cooling, and equipment using mass as a common unifying metric [12]. The focus on mass can reveal interesting/hidden aspects of process components; for instance, mass constraints can limit the type of energy vectors used. In this work, we present a computational design framework to evaluate the design of space biomanufacturing systems to produce lactic acid (one of the most valuable platform

chemicals) based on the mass of the system using *Escherichia coli*, *Saccharomyces cerevisiae* and *Pichia pastoris* as microorganisms. Our case study aims to evaluate different configurations for the design of biomanufacturing systems in space to identify the specific components of the system that are responsible for the highest impacts in terms of mass.

## METHODOLOGY

This work evaluates system design scenarios by quantifying system mass using different engineered organisms (*Escherichia coli*, *Saccharomyces cerevisiae*, and *Pichia pastoris*) to produce lactic acid. Lactic acid bioproduction has been widely studied on Earth because it serves as a platform chemical with diverse applications (i.e., chemical, pharmaceutical, food, cosmetics, and plastic industries) [13]. Recently, lactic acid has gained interest due to its capacity to be transformed into polylactic acid (PLA), a biodegradable polymer with applications in the production of packaging and new components through 3D printing, which is very convenient in space missions [14].

Because the performance of the biomanufacturing production may be impacted due to the high radiation levels that exist in the space environment, the addition of a shielding material could be necessary to maintain genetic stability in space fermentation systems. Therefore, these factors are considered to propose different bioprocess designs which can be coupled with the equipment of the life support system of the ISS to take advantage of the in-situ resource utilization.



**Figure 1:** Daily mass balance per crewmember in the ISS.

The life support system involves various chemical and biological processes, allowing the recycling of waste into valuable resources. In this way, the system operates in a closed-loop life support system. However, some waste resources that are not recovered (i.e., urine, fecal

feces) or are commonly vented into space (i.e., CH<sub>4</sub>, CO<sub>2</sub>) can serve as inputs in the bioproduction of lactic acid. For instance, nutrients that are required in the fermentation process can be obtained from human waste and urine. Figure 1 shows the daily mass balance per crewmember in the ISS. Commonly, there are 6 astronauts aboard the ISS, and based on this, the resources required to satisfy their basic living needs and the wastes generated can be quantified from the daily mass balance in the ISS. In this case, the available water that can be supplied to the biomanufacturing system is approximately 5 L per day. Therefore, fermentation volumes less than 5 L can use the available resources in the Life Support System to create a more circular system. Otherwise, for fermentation volumes greater than 5 L, the system will require external resources from resupply missions to be able to carry out the fermentation process. In this work, different fermentation volumes are evaluated to estimate their required mass, which is linked to the launch cost.

### Equivalent System Mass Analysis

The evaluation of a system by means of the ESM allows for the reduction of several physical quantities to a single parameter expressed in the units of mass (kilograms) for direct comparison. ESM has been employed as a transportation cost measure in studies to avoid the technical and political complications of using dollar costs for comparisons. In space missions, the cost of transporting a payload is proportional to the mass of that payload; therefore, a mass-based measure such as ESM is appropriate to quantify the launch cost of the system. In this regard, through ESM analysis, it is possible to identify which of several options can meet all specified requirements and have the lowest launch cost in relation to mass, volume, power, and cooling (Equation 1).

$$ESM = M + (P \times P_{eq}) + (C \times C_{eq}) \quad (1)$$

ESM can be calculated using Equation 1, where  $P_{eq}$  and  $C_{eq}$  are the power and cooling conversion factors (87 kg/kW and 146 kg/KW, respectively) to transform the non-mass quantities to their mass equivalent. These factors were derived by the NASA Life Support Baseline and Assumptions Document (BVAD) [15] and relates the quantity of mass and cost of installation and components per rate of power delivered. The term  $M$  refers to the mass of components that compose the system (e.g., chemicals, glass, stainless steel, polyethylene),  $P$  represents the power consumption (i.e., electrical power demand) and  $C$  is the thermal demand for rejecting from the system (Figure 2).

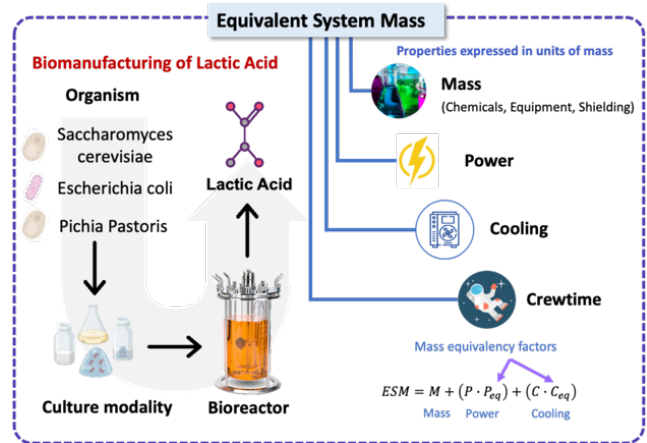


Figure 2: Equivalent System Mass framework.

In this work, the biomanufacturing system design is evaluated to compare the mass of different system designs considering the following assumptions:

- The system is composed of a spherical fermenter for the production of lactic acid.
- The materials employed for the fermenter are glass and stainless steel. The density of the stainless steel and glass are 2500 and 7500 kg/m<sup>3</sup>, respectively.
- The shielding material is polyethylene, which has a density of 930 kg/m<sup>3</sup>.
- The concentration of glucose is 200 g/L.
- In agreement with Rojdev et al. (2009) [16], the typical values for spacecraft structures have areal density thicknesses of approximately 1 to 20 g/cm<sup>2</sup>. In this work, we compared systems with 1 and 5 g/cm<sup>2</sup> areal density thicknesses.

Based on these specifications, the mass of the system involves the mass of the materials or chemicals employed in the lactic acid fermentation, the mass of the equipment and the mass of the shielding needs. The scenarios proposed for the different bioprocess design configurations labeled using the following format:

PX\_A\_B

Where X represents the volume of the fermenter in Liters, A represents the material design, which can be glass or stainless steel (G is used for glass and S represents stainless steel), and B indicates the polyethylene shielding thickness employed in the system. In this sense, some of the scenarios generated are:

- Standard scenario where shielding is not included (P10).
- Scenario with fermentation volume of 10 L using glass as the system design material and including

shielding with 1 cm of thickness (P10\_G\_1).

- Scenario with fermentation volume of 10 L using glass as system design material and including shielding with 5 cm of thickness (P10\_G\_5).
- Scenario with fermentation volume of 10 L using stainless steel as system design material and including shielding with 1 cm of thickness (P10\_S\_1).
- Scenario with a fermentation volume of 10 L using stainless steel as the system design material and including shielding with a 5 cm of thickness (P10\_S\_5).

In this work, fermentation volumes of 5, 10, 15 and 20 L were evaluated for the production of lactic acid. It is worth mentioning that the scenarios evaluated correspond to systems that are not exposed to radiation doses. However, for future work, we will evaluate how the growth of organisms is affected by different radiation doses. This can provide information on how fermentation time is impacted and how this may modify the amount of input resources needed in biomanufacturing processes.

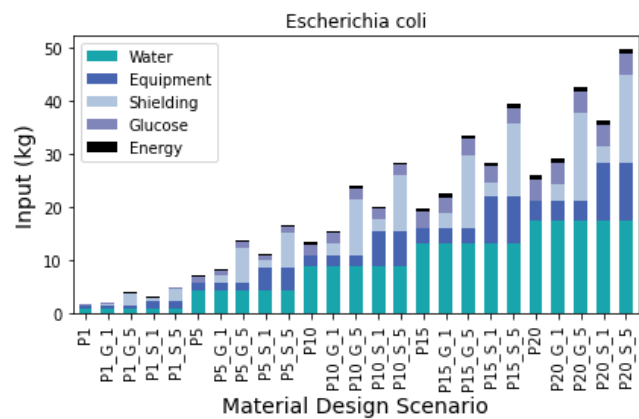
## RESULTS

ESM analysis was performed for the production of lactic acid from glucose using *Escherichia coli*, *Saccharomyces cerevisiae* and *Pichia pastoris*. The analysis was carried out for various design configurations and production capacities.

### Equivalent System Mass

The system mass considers both the mass of the design and protection materials and the production of lactic acid, which in turn is restricted by the fermentation volume. As a result, the system mass among the different organisms is similar. Figure 3 shows the results of ESM analysis using *Escherichia coli* as the host organism for lactic acid production. The results obtained indicate that for a fermentation volume of 10 liters, the system mass is approximately 12.4 kg for a standard scenario (P10). When shielding is added to the system (P10\_G\_1), its mass increases to 14.5 kg. At a shielding of 5 cm thickness, the mass increases further to 22.8 kg (P10\_G\_5). Furthermore, it can be seen that the equipment material design plays an important role in the system mass as a three-fold increase in equipment mass is obtained when stainless steel is used.

On the other hand, if the volume of the fermenter increases to 20 L, the system mass increases to 24 kg for a standard scenario. For the scenarios P20\_S\_1 and P20\_S\_5 the mass of the system is 34 kg and 48 kg, respectively.



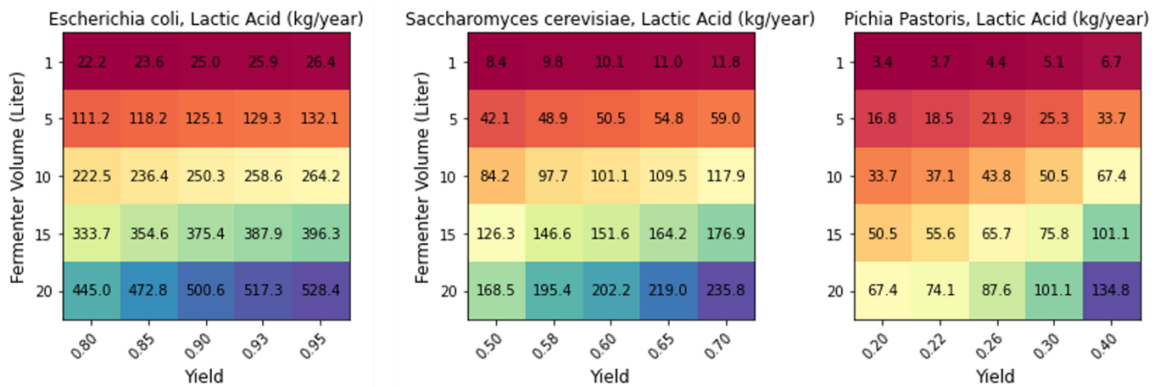
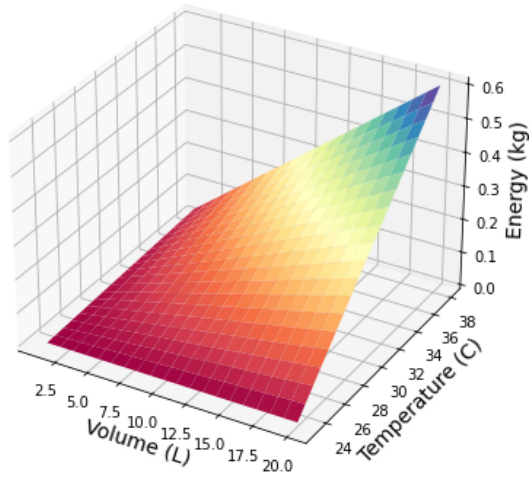
**Figure 3:** Equivalent System Mass Analysis for different scenarios per batch.

Results show that the mass of the system does not change significantly in terms of the organism selected for each batch fermentation system; however, the production of lactic acid is favored when *Escherichia coli* is employed as better yield is achieved. However, as can be seen in Figure 4, in a 10 L fermenter with a glucose concentration of 200 g/L, the total mass of the system using *Escherichia coli* is about 1425 kg/year. The mass of the *Saccharomyces cerevisiae* system is approximately 900 kg/year, 37% less mass than the mass required to operate a system using *Escherichia coli*. For both cases, the mass related to the equipment and shielding is negligible since represents less than 1% of the mass of the system.

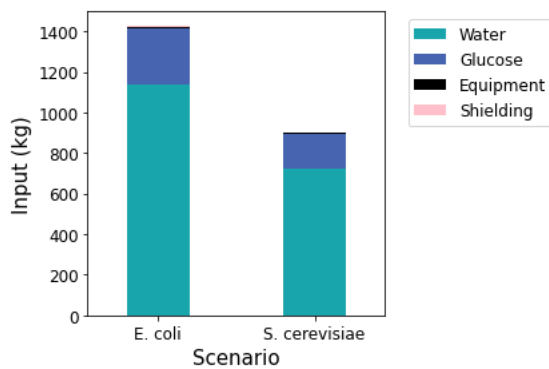
### Energy

The energy required to maintain the temperature of the fermentation process to produce lactic acid in one year was evaluated. Considering that the temperature inside the ISS is approximately 23 °C, the energy requirements in terms of mass were evaluated for different fermentation volumes (1 L – 20 L) and operation temperatures (24 °C – 37 °C) (Figure 5). Once we evaluated the energy required to maintain the temperature in the bioreactor, it was multiplied by the factor 87 kg/kW to convert the units of energy into units of mass. The energy required in terms of mass is equal to the factor 0.053 kg per liter per °C for *Escherichia coli*, 0.021 for *Saccharomyces cerevisiae* and 0.01 kg per liter per °C for *Pichia pastoris*. Our findings show that for a temperature equal to 36 °C and a volume fermenter equal to 1 L, the mass per year related to the energy required in the fermentation system is equal to 0.68 kg for *Escherichia coli*; however, the mass increases to 8.48 kg for a volume equal to 20 L.

**Figure 5:** Energy in terms of mass required for the lactic acid fermentation in one year.



**Figure 6:** Annual lactic acid production projections using different organisms



**Figure 4:** Equivalent System Mass Analysis for different scenarios per year.

hosts using glucose as a substrate.

Organism	Production rate (g/l/h)	Yield (g/g)	Reference
<i>Escherichia coli</i>	2.33	0.93	[17]
<i>Saccharomyces cerevisiae</i>	2.34	0.58	[18]
<i>Pichia Pastoris</i>	0.0363	0.22	[19]

### Lactic acid Production

A sensitivity analysis was performed to evaluate the impact of productivity on the production of lactic acid using each of the organisms. Results show that according to the productivity reported for the studied organisms (Table 1), the production of lactic acid per year with *Escherichia coli* in a fermenter volume of 10 L is about 258 kg,

**Table 1:** L-lactic acid production of different microbial



and it is higher than the other organisms evaluated. In the case of *Saccharomyces cerevisiae*, the production of lactic acid is 97.7 kg per year, and using *Pichia pastoris* the production is about 37.1 kg per year. Figure 6 shows the production of lactic acid that can be achieved assuming different yield and fermentation volumes.

## Challenges and Open Questions

Addressing the design of systems under space conditions requires different perspectives and considerations that are difficult to evaluate and reproduce in terrestrial conditions. ESM is a promising tool to evaluate new systems. As a cost metric, ESM may not be capable of capturing how biological performance is impacted by different radiation levels, but it allows us to evaluate the mass of resources involved in the process at different operating conditions. Because of this, research must be carried out to understand how to optimally engineer organisms for space environments. On the other hand, additional analyses are required to identify the trade-offs between engineering hosts to have better performance in high radiation environments compared to the addition of more radiation shielding material. Moreover, other mechanisms to reduce the shielding requirements and organisms' protection against radiation (e.g., different preservation modalities) should be evaluated in the design of space bioprocess. In this regard, immobilized cell methods such as lyophilization and hydrogel systems are good alternatives to suspend cell cultures and sustain long-term metabolic processes. Therefore, for future work, this modeling framework may be complemented with biomanufacturing design alternatives that could reduce the impact of radiation. Such a study could help to better assess trade-offs between strain resistance, preservation modalities, and overall process economics.

## CONCLUSIONS

This study presented a computational design framework for the design of biomanufacturing systems in space using a mass perspective to get an estimate of the launch cost of the system. ESM has been used as a transportation cost metric that involves the principal factors that affect the performance of the biomanufacturing systems such as the mass of the equipment, and all these elements are expressed in units of mass to allow for comparisons among the different design configurations of biomanufacturing systems.

In this work, the bioproduction of lactic acid using *Escherichia coli*, *Saccharomyces cerevisiae* and *Pichia pastoris* were evaluated by considering different fermenter materials and volumes to identify the bottlenecks in the design configurations in terms of mass and yield. Our analysis reveals that lactic acid is favored when *Escherichia coli* is employed. About 258 kg of lactic acid

per year can be produced using *Escherichia coli* in a 10 L stainless steel fermenter. At a shielding material density of 1 g/cm<sup>2</sup> the total mass of the system obtained from the ESM analysis is about 19 kg. Our findings show that most of the mass of the system is related to the water required to produce lactic acid. In this regard, ESM analysis helps to identify the system components that contains greater mass and would result in high launch costs. For future work, we are interested in studying other alternatives to create closed-loop systems involving the available resources inside the ISS.

## ACKNOWLEDGEMENTS

The authors received funding from Defense Advanced Research Projects Agency B-SURE program (N660012324019) for this research. The views, opinions and/or findings expressed should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

## REFERENCES

1. Warren LE. International Space Station open-source data. *Patterns*, 1(9). (2020)
2. NASA, NASA's Management and utilization of the International Space Station. <https://oig.nasa.gov/docs/IG-18-021.pdf>
3. Alvarado KA, Martínez JB, Brown MM, Christodoulou X, Bryson S, Denkenberger DC. Food production in space from CO<sub>2</sub> using microbial electrosynthesis. *Bioelectrochemistry*, 149, 108320. (2023)
4. Sukhinov DV, Gotovtsev PM, Sergeeva YE. Phototrophic microorganisms in bioregenerative life support systems for long-term crewed expeditions: Prospects and challenges. *Acta Astronautica*. (2023)
5. Lindeboom REF, De Paepe J, Vanoppen M, Alonso-Fariñas B, Coessens W, Alloul A, Vlaeminck SE. A five-stage treatment train for water recovery from urine and shower water for long-term human Space missions. *Desalination*, 495, 114634. (2020)
6. Tang C, Wang L, Zang L, Wang Q, Qi D, Dai Z. On-demand biomanufacturing through synthetic biology approach. *Materials Today Bio*, 18, 100518. (2023)
7. Wang Y, Hao L, Li Y, Sun Q, Sun M, Huang Y, Xiao L. In-situ utilization of regolith resource and future exploration of additive manufacturing for lunar/martian habitats: A review. *Applied Clay Science*, 229, 106673. (2022)
8. Cowley A, Perrin J, Meurisse A, Micallef A, Fateri M, Rinaldo L, Sperl M. Effects of variable gravity conditions on additive manufacture by fused

- filament fabrication using polylactic acid thermoplastic filament. *Additive Manufacturing*, 28, 814-820. (2019)
9. Averesch NJ, Berliner AJ, Nangle SN, Zezulka S, Vengerova GL, Ho D, Arkin AP. Microbial biomanufacturing for space-exploration—what to take and when to make. *Nature Communications*, 14(1), 2311. (2023)
  10. Cilliers J, Hadler K, Rasera J. Toward the utilisation of resources in space: knowledge gaps, open questions, and priorities. *Microgravity*, 9(1), 22. (2023)
  11. Sharma A, Clemens RA, Garcia O, Taylor DL, Wagner, NL, Shepard KA, Wagner WR. Biomanufacturing in low Earth orbit for regenerative medicine. *Stem Cell Reports*, 17(1), 1-13. (2022)
  12. Levri J, Fisher JW, Jones HW, Drysdale AE, Ewert MK, Hanford AJ, Vaccari DA. Advanced life support equivalent system mass guidelines document (No. NASA/TM-2003-212278). (2003)
  13. de Oliveira RA, Komesu A, Rossell CE, Maciel Filho R. Challenges and opportunities in lactic acid bioprocess design—From economic to production aspects. *Biochemical Engineering Journal*, 133, 219-239. (2018)
  14. Prado-Rubio OA, Gasca-González R, Fontalvo J, Gómez-Castro FI, Pérez-Cisneros ES, Morales-Rodríguez R. Design and evaluation of intensified downstream technologies towards feasible lactic acid bioproduction. *Chemical Engineering and Processing-Process Intensification*, 158, 108174. (2020)
  15. Wójtowicz MA, Cosgrove JE, Serio MA, Lee JM. An Equivalent System Mass (ESM) Analysis for the Universal Waste Management System (UWMS) with and without the Torrefaction Processing Unit (TPU). 2020 International Conference on Environmental Systems. (2020)
  16. Rojdev K, Atwell W, Wilkins R, Gersey B, Badavi FF. Evaluation of multi-functional materials for deep space radiation shielding. In *National Space and Missile Materials Symposium* (No. JSC-CN-18378). (2009)
  17. Dien BS, Nichols NN, Bothast RJ. Recombinant *Escherichia coli* engineered for production of L-lactic acid from hexose and pentose sugars. *Journal of Industrial Microbiology and Biotechnology*, 27(4), 259-264. (2001)
  18. Lee JY, Kang CD, Lee SH, Park YK, Cho, KM. Engineering cellular redox balance in *Saccharomyces cerevisiae* for improved production of L-lactic acid. *Biotechnology and Bioengineering*, 112(4), 751-758. (2015)
  19. Yamada R, Ogura K, Kimoto Y, Ogino H. Toward the

construction of a technology platform for chemicals production from methanol: D-lactic acid production from methanol by an engineered yeast *Pichia Pastoris*. *World Journal of Microbiology and Biotechnology*, 35(2). (2019)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Techno economical assessment of a low-carbon hydrogen production process using residual biomass gasification and carbon capture.

E.J. Carrillo<sup>a</sup>, J. Lizcano-Prada<sup>b</sup>, V. Kafaro<sup>a</sup>, D. Rodriguez-Vallejo<sup>c</sup>, A. Uribe-Rodríguez<sup>d\*</sup>

a Research Center for Sustainable Development in Industry and Energy (CIDES), Universidad Industrial de Santander, 680002 Bucaramanga, Colombia

b TIP, Colombia, Km 7 + 400m Anillo vial Palenque, Diagonal Floridablanca No 22-31 - Bodega 11, Floridablanca, Colombia

c Pfizer, Chemical Research Development, Sandwich, CT13 9NJ, United Kingdom.

d Centre for Innovation and Technology Colombian Petroleum Institute, ECOPETROL, 681011 Piedecuesta, Colombia.

\* Corresponding Author: ariel.uribe@ecopetrol.com.co.

## ABSTRACT

Aiming to mitigate the environmental impact derived from fossil fuels, we propose an integrated carbon capture-biomass gasification process is proposed to produce low-carbon hydrogen as an alternative energy carrier. The process begins with the pre-treatment of empty fruit bunches (EFB), involving grinding, drying, torrefaction, and pelletization. The resulting EFB pellet is then fed into a dual gasifier, followed by a catalytic cracking of tar and water gas shift reaction to produce syngas, aiming to increase its H<sub>2</sub> to CO ratio. Subsequently, we explore two alternatives (DEPG and MEA) for syngas upgrading by removing CO<sub>2</sub>. Finally, a PSA system is modeled to obtain H<sub>2</sub> at 99.9% purity. The pre-treatment stage densifies the biomass from an initial composition (%C 46.47, %H 6.22, %O 42.25) to (%C 54.10, %H 6.09, %O 28.67). The dual gasifier operates at 800°C, using steam as a gasifying agent. The resulting syngas has a volume concentration (%CO 20.0, %CO<sub>2</sub> 28.2, %H<sub>2</sub> 42.2, %CH<sub>4</sub> 5.9). Next stages of the process focus on removing the CO<sub>2</sub> and increased H<sub>2</sub> through catalytic reactions from the syngas. Thus, the DEPG carbon capture process can decrease the CO<sub>2</sub> concentration to 2.9%, increasing the hydrogen to 95.6% in volume. In contrast, the MEA process reduces the concentration of CO<sub>2</sub> to 5.2% and increases the concentration of H<sub>2</sub> to 93.1%. Moreover, we estimate a levelized costs of hydrogen (LCOH) and carbon capture cost for each method (DEPG and MEA) (LCOC) and CO<sub>2</sub> avoided (LCCA). LCOH: 3.05 USD/kg H<sub>2</sub>, LCOC: 92 and 59 USD/t CO<sub>2</sub> and 183 and 119 USD/t CO<sub>2</sub>, for DEPG and MEA respectively.

**Keywords:** Empty fruit bunch, Gasification, Carbon capture, Torrefaction, Pre-treatment.

## INTRODUCTION

In recent years, processes that generate clean energy have gained relevance in academia and industry. Due to the growing energy demand, supplied by the oil industry, there has been a constant increase in greenhouse gas emissions. It is estimated that CO<sub>2</sub> emissions generated by industries are in the order of 8 billion tons per year, it is estimated that emissions will reach 10 billion tons per year [1].

For this reason, several countries such as Japan, Australia, Germany, and most of the countries of the

European Union have committed to initiating decarbonization processes, in addition to encouraging processes that lead to a circular economy. One of the energy transition strategies consists of the use of hydrogen as an energy carrier [2]. Currently, the demand for hydrogen has undergone a significant increase from 20 MT (million tons) in 1970 to more than 70 MT by 2018 globally. Most of the hydrogen produced, comes from fossil sources with a commercial value of 1 to 3 USD per kg [3].

In the Colombian context, there is a commitment to reduce CO<sub>2</sub> emissions by 51% by 2030 by promoting the production of hydrogen from sustainable sources [4].

The thermochemical processes are one of the most

attractive alternatives in the short term to transform biomass or carbonaceous fuel, into energy and chemical products or their building blocks such as syngas. Depending on the oxygen requirements, the operating temperature and the syngas composition, the thermochemical processes can be classified into pyrolysis, gasification, and combustion [5].

Currently, the predominant transformation process for the biomass is the combustion, which accounts for approximately 90% of its worldwide utilization, generating energy ranging from a few megawatts to 100 megawatts [6]. Nevertheless, nowadays, there is a growing interest in the gasification process due to the flexibility to produce syngas, hydrogen, and biogenic carbon dioxide. As reported by the International Energy Agency (IEA) Bioenergy, as of 2020, there are 686 operational gasifiers worldwide, where 272 plants are producing syngas on a large scale, boasting an energy capacity of approximately 200 *GWh*—equivalent to around 200 *MWh* per installed gasifier[7].

Biomass gasification is challenging due to its high moisture content, high *O/C* ratio, low bulk density, lower grindability, and heterogeneous nature, these features make it difficult to use biomass as fuel, for that reason is necessary to carry out a pre-treatment. Furthermore, higher tar formation during the gasification of raw biomass increases the downstream cost of gas separation. Pre-treatment of biomass through torrefaction could remove some of these limitations, making biomass a more suitable feedstock for gasification[8].

Carbon capture stands out as a promising technology to mitigate climate change. The integration of biomass gasification and carbon capture lead to negative net carbon emissions and may therefore provide an important technological alternative for meeting current greenhouse gas stabilization targets. To this end, syngas from biomass gasification combined with pre-combustion carbon capture can be used to produce either biofuels or electricity [9].

It is expected that the technological development of biomass gasification processes coupled with  $CO_2$  capture might satisfy the energy demand in such a way that they can be considered as a technical, economic and environmental feasible alternative. It is hoped that research such as the one presented here could contribute to the development of these processes.

## PROCESS DESCRIPTION

### Feedstock

The palm oil is a well-established industry in Colombia, this level of industrialization is providing a reliable source of raw material for further transformation into bioenergy, biofuels and biohydrogen among others. We select the empty fruit bunch (EFB), which is a residual

biomass from the palm oil extraction as raw material for the gasification process. The EFB characterization given by the ultimate and proximal analysis is shown in Table 1. In additionally, we consider that the EFB has a moisture content of 48% by weight, with a particle size distribution of 40 *cm* in diameter, following a normal distribution with a mean of 8.3 *cm*. The EFB gasification plant has a design capacity of 80 *t/h*. The biomass was modeled as no-conventional solid, using Aspen plus V12, for enthalpy and density are applied HCOALGEN and DCOALIGT correlations respectively, which are taken by the simulation. The HHV value of biomass moisture provided by simulation is compared to literature data [10].

**Table 1:** Ultimate and proximal analysis.

Ultimate	Value%	Proximate	Value%
C	46.47	Fixed carbon	16.07
H	6.22	Volatile	69.14
O	42.25	Moisture	48.00
S	0.08	Ash	3.86
N	1.12	HHV	18.5
			MJ/Kg
Ash	3.86		

## PRE-TREATMENT

This stage encompasses biomass grinding, drying, torrefaction, and pelletization processes. We select two fluid packages: SOLIDS for processes related to biomass, and Redlich-Kwong-Soave-Boston-Mathias for gases generated from the burning of gases and liquids resulting from torrefaction reactions.

The main objective of the pre-treatment is to increase the energy density of the biomass. Thus, the particle size distribution, *O/C* and *H/C* ratios can be considered as key performance targets of this stage (Figure 1). At the end of the process, 35.7 *T/h* torrefied pellets of EFB are produced, with a mass yield of 0.44 *Kg* pellet torrefied EFB/*Kg* moisture biomass and ratio of energy consumption of 0.16 *KW/Kg* pellet torrefied EFB.

### Grinding

We simulate this process using the correlations provided by Rosin Rammler and Sperling Bennet, which are retrieved from the Aspen plus database.

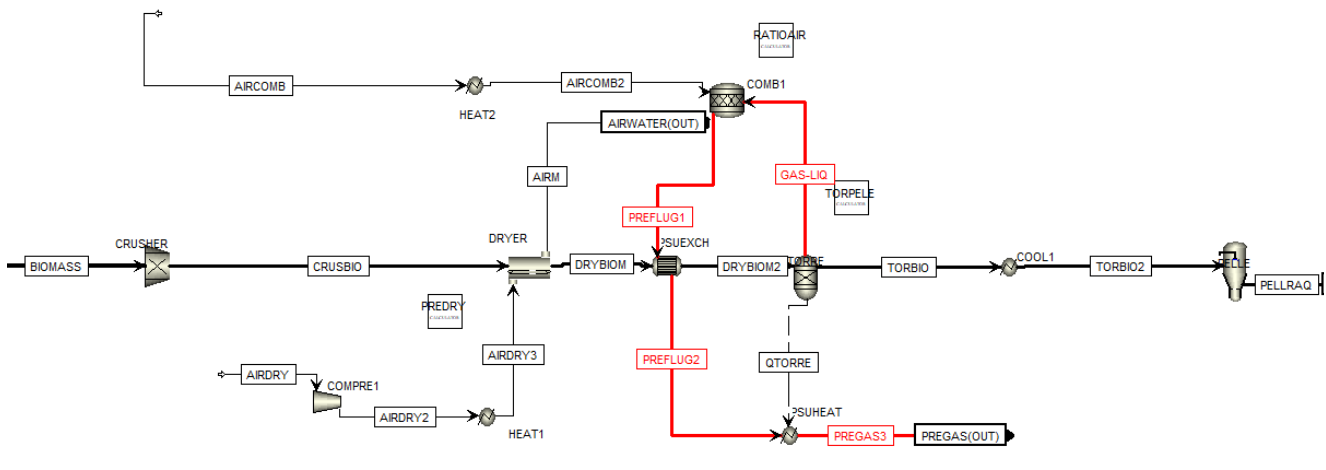
We define a target of 5 *mm* as the particle size, which is predicted as a function of the specific power and Bond work index or Hardgrove grindability index (HGI). The first parameter is given by Ruksathamcharoen with a value of 14 *Kwt/t* [11], and the second one is provided by Montealegre with a value of 12.6 HGI [12]. See the operation unit CRUSHER in the Figure 1.

### Dying

We adapt correlations proposed by Han [13]

**Table 3:** Comparison raw EFB with torrefied pellets of EFB (PTEFB)

Ultimate analysis	Raw EFB	PTEFB	Proximate	Raw EFB	PTEFB
C	46.47	54.10	Fixed carbon	16.07	23.76
H	6.22	6.09	Volatile	69.14	62.14
O	42.25	28.67	Moisture	48.00	5.35
S	0.08	0.11	Ash	3.86	9.43
N	1.12	1.76	HHV[MJ/Kg]	18.5	20.8
Ash	3.86	9.43			



through a CALCULATOR block to predict air demand and sizing for dried. This operation unit is represented by the DRYLER, which was set up to operate in co-current. See details in the Table 2.

**Table 2:** Operations condition and dimensions of DRYLER.

Parameter	value	Units
Length	20.94	m
Cross area	7.02	m <sup>2</sup>
Residence time	23	min
Ratio air biomass	7.89	Kg air/Kg EFB
Input air temperature	180	°C
Outlet moisture	9	Weight (%)

### Torrefaction

This process seeks to improve *O/C* and *H/C* ratios. Thus, the residual biomass can be considered as a vegetal coal, with a performance alike the mineral coal. We represent this operation unit by TORRE and the block calculator TORPELE. We predict yields of solid, liquid and gas and the composition of each phase based on Azir's work [8] [14]. According to the author, the optimum temperature and residence time for the torrefaction process are 250 °C and 40 minutes respectively, which are

considered in this work (see results in the Table 3). The gas and liquid streams of this process are used to provide the energy requirements for the torrefaction (see red line in Figure 1).

### Pelletization

The resulting torrefied EFB is cooled at 100 °C before it is fed to the pelletization stage, which is modeled by the GRANULADOR block called PELLE, defining the calculator TORPELE. It was set up an energy requirement of 30 Kwt/t of EFB to achieve a particle size of 10 mm [15]. Note: The effects of pelletization on the chemical properties of torrefied EFB are not considered.

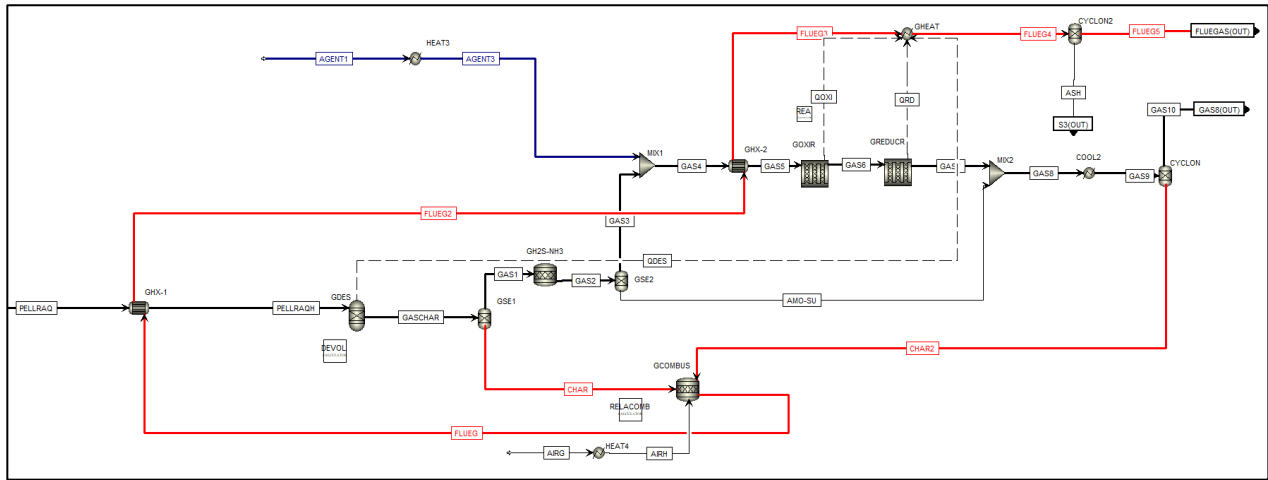
### GASIFICATION

We model a dual gasifier for EFB pellets following the methodology proposed by Puig [16] and Bach [17]. This process operates at atmospheric pressure and using steam as gasification agent. It is selected Peng Robison Boston- Mathias as fluid package for the simulation. In the Figure 2, we represent the gasification process using five blocks (GDES, DEVOL, GH2S-NH3, GOXIR and GREDUCR). The decomposition of torrefied pellets of EFB is modeled by the blocks GDES and DEVOL, using the



**Table 4:** Mass fractions for each part of the gasification process

Component	Decompositions	Gasification	Combustion chamber
C	0.18	0	0
O <sub>2</sub>	0	0	0.04
CO	0.30	0.16	0
CO <sub>2</sub>	0.30	0.35	0.24
H <sub>2</sub>	0.01	0.03	0
N <sub>2</sub>	0.01	0.002	0.70
CH <sub>4</sub>	0.04	0.03	0
C <sub>6</sub> H <sub>6</sub>	0.03	0.01	0
NH <sub>3</sub> (Kg/h)	0	380.5	0
H <sub>2</sub> S (Kg/h)	0	29.7	0
C <sub>2</sub> H <sub>4</sub>	0.07	0	0
H <sub>2</sub> O	0	0.42	0
S <sub>2</sub>	0.001	0	0
Ash	0.06	0	0.02
Flow rate (T/h)	34.516	59.39	94.50
Stream name	GASCHAR	GAS 10	FLUEG4
Ratio steam/Biomass	0	0.9	0



**Figure 2.** Simulation of dual gasifier.

correlations proposed by Neves [18], which are represented by the equations (1-11). As result, we predict the yield for each component ( $CO, CO_2, H_2, H_2O, CH_4, C_2H_4, Tar, C, Ash$ ). Note: The operating temperature is the independent variable. According to Lee, the tar fraction is represented as  $C_6H_6$  allowing this assumption to be made[19].

$$(\sum_j Y_{j,F} - Y_{ch,F} \sum_j Y_{j,ch}) * LHV_G = (Y_{tar,F} * Y_{H_2O,F}) * LHV_G + Y_{C_2H_4,F} * LHV_{C_2H_4} + Y_{CH_4,F} * LHV_{CH_4} + Y_{CO,F} * LHV_{CO} + Y_{H_2,F} * LHV_{H_2} \quad (1)$$

$$Y_{ch,F} = 0,106 + 2,43 * e^{(-0,66*10^{-2}*T)} \quad (2)$$

$$Y_{C,ch} = 0,93 + 0,92 * e^{(-0,42*10^{-2}*T)} \quad (3)$$

$$Y_{O,ch} = 0,07 + 0,85 * e^{(-0,48*10^{-2}*T)} \quad (4)$$

$$Y_{H,ch} = -0,41 * 10^{-2} + 0,10 * e^{(-0,24*10^{-2}*T)} \quad (5)$$

$$Y_{H_2,F} = 1,145 * (1 - e^{(-0,11*10^{-2}*T)})^{9,384} \quad (6)$$

$$Y_{CH_4,F} = -2,18 * 10^{-4} + 0,146 * Y_{CO,F} \quad (7)$$

$$Y_{CO,F} = \left( (3 * 10^{-4} + \frac{0,0429}{1 + (\frac{T}{632})^{-7,23}}) \right)^{-1} * Y_{H_2,F} \quad (8)$$

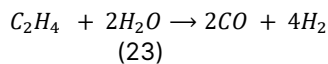
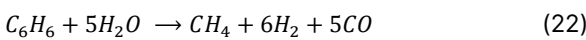
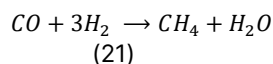
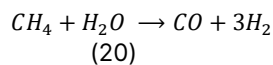
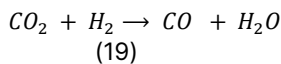
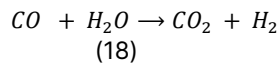
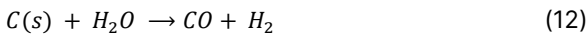
$$Y_{C,tar} = 1,14 * Y_{C,bio} \quad (9)$$

$$Y_{O,tar} = 0,8 * Y_{O,bio} \quad (10)$$

$$Y_{H,tar} = 1,13 * Y_{H,bio} \quad (11)$$

The carbon and ashes generated in the decomposition stage are sent to a combustion chamber called GCOMBUS, the remaining gas products from decomposition are directed to the stoichiometric reactor called GH2S-NH3, where impurities such as  $NH_3$  and  $H_2S$  are formed. These impurities are temporarily separated from the gas, with the aim of excluding them from the gasification reactions and thus maintaining equilibrium. Subsequently, the gas, free of impurities are mixed with steam and enters the gasification stage conducted in the gas reactor, in this case there is not oxygen in the mix, for that reason only the reduced reactions occur, producing a reduced gas. This reduced gas is then mixed with the impurities to obtain synthesis gas. The equations 11-23 represented the reduction reactions.

The red stream in Figure 2 illustrates the energy exchange between the gases generated by the combustion of the solid phase (carbon and ashes) and the chamber of the bubbling bed gasifier. Table 3 displays the mass composition of different stages of the process.



**Table 5:** Operations condition and dimensions of dual gasifier.

Parameter	value	Units
Temperature gasifier	800	°C
Temperature flue gas	847	°C
Diameter	2.5	m
Length	15	m
Ratio steam biomass	0.9	Kg steam/Kg EFB
Outlet moisture	9	Weight (%)

The proposed model suggests that unreactive char

and ash interchange heat with the flue gas as well as syngas thus, temperature of flue gas and solids should be the same. The temperatures of syngas and flue gas are compared with the experimental case. Table 5 shows the operations conditions of gasifier.

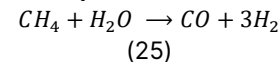
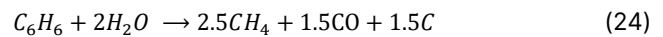
The results obtain in the gasification simulation are compared with the experimental data [20]. When comparing the results, errors of less than 15% are observed with respect to the volumetric composition of the products, except for tar, where the error is 20%. These discrepancies arise from differences in the raw materials used, as well as the absence of torrefaction processes in the evaluation performed by Smich. Another significant factor affecting result accuracy is the presence of transfer phenomena, which are not accounted for in the gasifier.

## GAS CLEANING AND UPGRATING

### Catalytic cracking tar (CCT)

The resulting gas contain quantities of tar ( $C_6H_6$ ) that can cause tube plugging. However, this tar can also be utilized to enhance the process yield. Therefore, it is crucial to eliminate tar through the catalytic cracking process, as depicted in Figure 3.

The diagram illustrates the necessity to raise the outlet gas temperature to 900 °C to achieve tar destruction using dolomite catalyst (see equation 24), Subsequently gas goes into other catalytic reactor to eliminate methane and transform into  $CO$  and  $H_2$  (see equation 25), the final gas contains high concentration of  $CO$ , see Table 6. Both reactions follow the Arrhenius model. The kinetic model is provided by Srinivas who investigates [21].



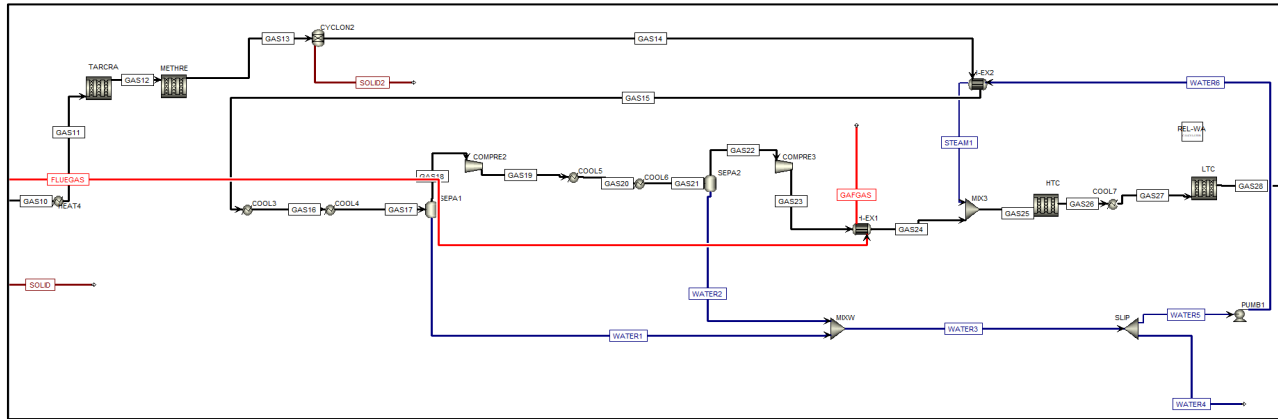
### Water gas shift reaction (WGSR)

Water gas shift reaction usually occurs in a catalytic environment and high pressures, the reactor operates at 28 bar and the process subdivided into two parts. High temperature shift catalytic (HTSC) and Low temperature shift catalytic (LTSC) see the equation (26-27). According to Reza [22] "if high purity of hydrogen is needed, the most common configuration for doing WGSR consist in two consecutive adiabatic fixed bed reactors".

The motive of this arrangement due to the reaction is kinetically favorable at high temperatures while it is thermodynamically at low temperatures. The first reactor used an iron-based catalyst and the second used copper-based catalyst, the Figure 3 shows simulation of

**Table 6:** Volumetric composition at 0 °C of dry syngas for each process

Component	Gasification	Tar cracking	Methane cracking	HTC	LTC
CO	20.0	20.7	23.0	6.9	1.0
CO <sub>2</sub>	28.2	27.4	22.4	32.8	36.5
H <sub>2</sub>	44.2	42.9	53.2	59.9	62.1
N <sub>2</sub>	0.3	0.3	0.2	0.2	0.2
CH <sub>4</sub>	5.9	7.4	0	0	0
C <sub>6</sub> H <sub>6</sub>	0.0002	0	0	0	0
NH <sub>3</sub> (Kg/h)	380.5	380.5	380.5	380.5	380.5
H <sub>2</sub> S (Kg/h)	29.7	29.7	29.7	29.7	29.7
Pressure (bar)	1.0	1.0	1.0	28	25
Temperature (°C)	800	900	900	316	189
Dry flow rate (T/h)	34.516	34.716	36.987	42.884	45.274
Stream name	GAS 11	GAS 12	GAS 13	GAS 26	GAS 28
Reactor volume (m <sup>3</sup> )	73.63	1.17	0.20	11.69	2.09



**Figure 3.** Gas cleaning and upgrading.

each reactor. In the Table 6 shows upgrading gas composition at the end of process.



$$R_{HTSC} = A_C \times 10^{2.845} \exp\left(\frac{-111}{RT}\right) P_{CO} P_{CO_2}^{-0.36} P_{H_2}^{-0.09} \left(1 - \frac{P_{CO_2} P_{H_2}}{K_e P_{CO} P_{H_2O}}\right) \quad (27)$$

$$R_{LTSC} = A_C \times 2.96 \times 10^5 \exp\left(\frac{-47400}{RT}\right) \left(P_{CO} P_{H_2O} - \frac{P_{CO_2} P_{H_2}}{K_e}\right) \quad (28)$$

$$K_e = \exp\left(\frac{-4577.8}{T} - 4.33\right) \quad (29)$$

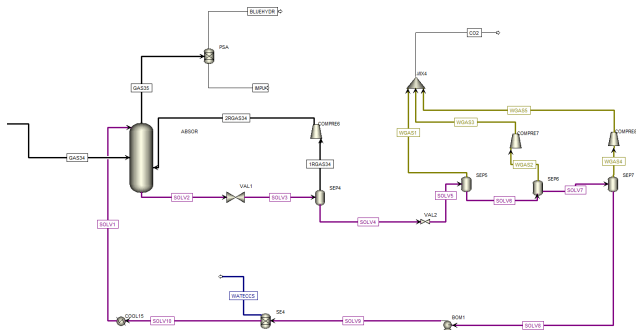
## CARBON CAPTURE

Subsequently, the purified gas, with a high concentration of CO<sub>2</sub> (see Table 6) is directed towards the carbon capture process. In this phase, two alternatives are evaluated based on simulation examples provided by Aspen Plus. Both alternatives employ amines as solvents with the aim of reducing the concentration of

components such as CO<sub>2</sub> and H<sub>2</sub>S present in the gas stream. For both options, a PSA unit must be installed to achieve hydrogen purity of 99.9%. The simulations presented below correspond to examples extracted from the Aspen Plus V12 database, which were modified for the purified stream in this case [23] [24] .

## Physical adsorption

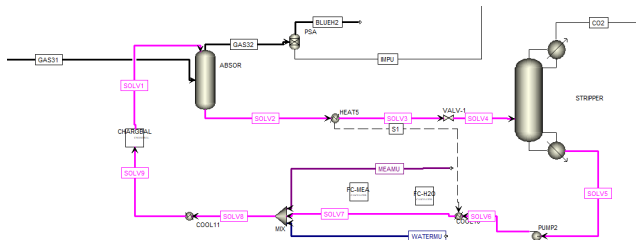
In the first case, DEPG (Dimethyl Ether of Polyethylene Glycol) is employed as the solvent, which eliminates CO<sub>2</sub> through a physisorption mechanism. This system requires high pressure and low temperature. Figure 4 depicts the simulation of the process using DEPG. Table 8 shows the operations conditions.



**Figure 4.** Carbon capture using DEPG as solvent.

### Chemical adsorption

For the second case, MEA is used as the solvent. In this system, the interaction occurs through a chemisorption mechanism. Therefore, it is necessary to specify the equilibrium reaction between the solvent and the gas. Figure 5 illustrates the adsorption with MEA. Table 7 shows the results of adsorption for both cases.



**Figure 5.** Carbon capture using MEA as solvent.

**Table 7:** Volume concentration of gas after capture carbon.

Component	DEPG	MEA
CO	1.1	1.4
CO <sub>2</sub>	2.9	5.2
H <sub>2</sub>	95.6	93.1
N <sub>2</sub>	0.3	0.3

**Table 8:** Operations condition and dimensions of capture processes.

Parameter	DEPG	MEA
Operation temperature (°C)	20	24
Operation pressure (bar)	69	17
Diameter (m)	1.8	1.9
Length (m)	12	3.2
Hydrogen flow rate (T/h)	2.88	3.12
Fluid package	PC-SAFT	ENRTL-RK

It is necessary to mention that although adsorption using MEA results in a higher concentration of hydrogen compared to DEPG (refer to Table 7), there is a loss of

hydrogen of 8.9% by weight when using DEPG, while the loss of hydrogen with MEA adsorption is only 1.24%. Subsequently, the gas proceeds to the PSA unit. For DEPG adsorption, it achieves a throughput of  $58,75 \frac{Tn}{Day}$  whereas for MEA, it is  $63,72 \frac{Tn}{Day}$ . The mass yield of MEA and DEPG processes with PSA corresponds to 0.059 and 0.054 *Kg hydrogen/kg gas*. Additionally, the energy yield is 0.15 and 3.14 *KW /Kg hydrogen* respectively.

### ECONOMICAL ASSESSMENT

According to the simulation results, economic parameters such as CAPEX and OPEX are calculated using different correlations. For the pre-treatment stage, the methodology proposed by Guthrie [25] is employed. The cost and installation of the dual gasifier are calculated using the correlations presented by Gunnarsson [26]. Two-phase separators are assessed using the methodology proposed by Turton [27], while for the combustion reactor in the pre-treatment and the equipment involved (PSA), the methodology presented by Hoffman [28] is followed.

An interest rate of 11% over a duration of 20 years is assumed, with an estimated construction period of 3 years for the plant.

The economic assessment is conducted separately for each stage of the process (pre-treatment, gasification, and carbon capture).

#### Pre-treatment

For economic assessment of pre-treatment. We evaluate the LCOE pellets as parameter in the Table 9 shows the results of assessment.

**Table 9:** Economic assessment of pre-treatment

Parameter	value	Units
CAPEX	\$ 2'490.166	USD
OPEX	\$ 2'304.759	USD
LCOE pellet	77	USD/T pellet
Energy consumption	0.120	MW/T pellet

#### Gasification and gas upgrading

In the gasification process, we utilize the correlation presented in Gunnarsson [26] for cost estimation. Employing the six-rule method, we estimate the equipment cost and operational expenses using the Lower Heating Value (LHV) of syngas. For the stage of gas cleaning and upgrading, the correlations provided by Warren [25] are applied. The results of the assessment are presented in Tables 10 and 11.

The Levelized Cost of Hydrogen (LCOH) includes costs related to gas cleaning and upgrading. According

to GEP company, the LCOH ranges between 2,8 a 3,5 USD/Kg H<sub>2</sub>[29].

**Table 10:** Economic assessment of gasification

Parameter	value	Units
CAPEX	\$ 34'925.863	USD
OPEX	\$ 2'327.240	USD
LCOH	3.05	USD/Kg H <sub>2</sub>

**Table 11:** Economic assessment of gas cleaning and upgrading.

Parameter	value	Units
CAPEX	\$ 97.278.973	USD
OPEX	\$ 12'726.235	USD

## Carbon capture

For the economic assessment, we evaluated both methods of carbon capture. The MEA system was chosen due to its high selectivity to hydrogen compared to the DEPG system. Tables 12 and 13 show the results of the assessment for carbon capture using MEA and DEPG, respectively. In both carbon capture systems, PSA is employed.

The results suggest that the method using DEPG is appropriate, despite its high electricity consumption and low hydrogen production compared to using MEA.

**Table 12:** Economic assessment of carbon capture using MEA with PSA

Parameter	value	Units
CAPEX	\$ 113'685.286	USD
OPEX	\$ 9'682.424	USD
CAPEX/CO <sub>2</sub> T/h	\$ 3'062.331	USD
LCOC	92	USD/T CO <sub>2</sub>
LCOA	183	USD/T CO <sub>2</sub>

**Table 13:** Economic assessment of carbon capture using DEPG with PSA

Parameter	value	Units
CAPEX	\$ 83'075.549	USD
OPEX	\$ 4'781.573	USD
CAPEX/CO <sub>2</sub> T/h	\$ 2'128.205	USD
LCOC	59	USD/T CO <sub>2</sub>
LCOA	119	USD/T CO <sub>2</sub>

## CONCLUSION

Based on the process results, it can be affirmed that the proposed model for torrefaction is valid, as it can

indirectly estimate the higher calorific value of the torrefied EFB.

It is necessary to clarify that the proposed design does not correspond to the optimum. In the cases of gasification, tar cracking, methane cracking, and water gas shift reaction reactors, sensitivity analysis was conducted to estimate the size of equipment for the base case of carbon capture systems. The simulation extracted from the Aspen Plus examples was modified to consider the resulting flow of the plant proposed in this project.

Regarding gasification, the combination of the models proposed by Puig and Bach, along with the energy integration implemented and Neves' model for the decomposition of the EFB, allows for generating a simulation that emulates the experimental results presented by Schmid, including the formation of tar. In the future, it is considered to consider the reactivity of the ashes.

It is observed that the gas upgrading eliminates most of the problems related to the presence of tar, monoxide, and methane. Additionally, increasing the concentration of hydrogen eliminates problems related to subsequent adsorption, as the presence of tar in the stream can affect the efficiency of adsorption.

The cost of this low-carbon hydrogen is close to the market price. The reasons are related to the high quantities of hydrogen produced in gas upgrading processes WGS and CCT.

## ACKNOWLEDGMENTS

Financial support from the Center for Innovation and Technology Colombian Petroleum Institute (Ecopetrol S.A.) and from the Universidad Industrial de Santander (UIS) for this research project is gratefully acknowledged, as well as the technical support and guide from professionals who are part of the agreement 3042921 between UIS and Ecopetrol. The views expressed in this work do not necessarily reflect those of Ecopetrol S.A.

## REFERENCES

1. B. Page and G. Turan, "GLOBAL STATUS OF CCS 2020." Accessed: Dec. 14, 2023. [Online]. Available: <https://www.globalccsinstitute.com/resources/publications-reports-research/global-status-of-ccs-report-2020/#:~:text=The%20Global%20Status%20of%20CCS,over%20the%20past%2012%20months.>
2. M. Noussan, P. P. Raimondi, R. Scita, and M. Hafner, "The role of green and blue hydrogen in the energy transition—a technological and geopolitical perspective," *Sustainability (Switzerland)*, vol. 13, no. 1. MDPI AG, pp. 1–26, Jan. 01, 2021. doi: 10.3390/su13010298.



3. D. Fickling, "Bloomberg. Bloomberg: A Three-Part Series on Hydrogen Energy." Accessed: Nov. 09, 2023. [Online]. Available: <https://www.bloomberg.com/graphics/2020-opinion-hydrogen-green-energy-revolution-challenges-risks-advantages/oil.html>
4. Ministerio de ciencias, "Hoja\_Ruta\_Hidrogeno\_Colombia\_2810," 2021.
5. W. J. Martinez-Burgos et al., "Hydrogen: Current advances and patented technologies of its renewable production," *Journal of Cleaner Production*, vol. 286. Elsevier Ltd, Mar. 01, 2021. doi: 10.1016/j.jclepro.2020.124970.
6. Irena, "RENEWABLE ENERGY TECHNOLOGIES: COST ANALYSIS SERIES," 2012. [Online]. Available: [www.irena.org/Publications](http://www.irena.org/Publications)
7. Y. Jafri, L. Waldheim, W. Consulting, and J. Lundgren, *Emerging Gasification Technologies for Waste & Biomass*. 2020.
8. M. K. Raut, P. Basu, and B. Acharya, "The Effect of Torrefaction Pre-Treatment on the Gasification of Biomass," *International Journal of Renewable Energy & Biofuels*, pp. 1–14, Sep. 2016, doi: 10.5171/2016.823723.
9. B. Erlach, B. Harder, and G. Tsatsaronis, "Combined hydrothermal carbonization and gasification of biomass with carbon capture," *Energy*, vol. 45, no. 1, pp. 329–338, 2012, doi: 10.1016/j.energy.2012.01.057.
10. S. Kerkkaiwan and D. Boonbumrung, "Production of high quality empty fruit bunch pellet by water washing and torrefaction," in *IOP Conference Series: Earth and Environmental Science*, Institute of Physics Publishing, Apr. 2020. doi: 10.1088/1755-1315/463/1/012130.
11. S. Ruksathamcharoen, "EFFECT OF HYDROTHERMAL TREATMENT ON GRINDABILITY AND FUEL CHARACTERISTICS OF EMPTY FRUIT BUNCH DERIVED HYDROCHAR," pp. 1246–1255, 2018.
12. Montealegre, "Evaluación tecnoeconómica de la torrefacción de biomasa residual de la agroindustria de la palma de aceite en un horno rotatorio," 2020.
13. J. Han, Y. Choi, and J. Kim, "Development of the Process Model and Optimal Drying Conditions of Biomass Power Plants," 2020, doi: 10.1021/acsomega.9b03557.
14. M. Azri et al., "A comprehensive study on torrefaction of empty fruit bunches: Characterization of solid, liquid and gas products," *Energy*, vol. 230, p. 120877, 2021, doi: 10.1016/j.energy.2021.120877.
15. Vukmirovic, D., & Kokic, B. Effect of steam conditioning on physical properties of pellets and energy consumption in pelleting process. <https://www.researchgate.net/publication/277198999>
16. M. Puig-gamero, D. T. Pio, L. A. C. Tarelho, and L. Sanchez-silva, "Simulation of biomass gasification in bubbling fluidized bed reactor using aspen plus ®," vol. 235, no. October 2020, 2021, doi: 10.1016/j.enconman.2021.113981.
17. Q. Bach, D. D. Nguyen, and C. Lee, "Effect of Torrefaction on Steam Gasification of Biomass in Dual Fluidized Bed Reactor — a Process Simulation Study," pp. 1042–1051, 2019.
18. D. Neves, H. Thunman, A. Matos, L. Tarelho, and A. Gómez-barea, "Characterization and prediction of biomass pyrolysis products," *Prog Energy Combust Sci*, vol. 37, no. 5, pp. 611–630, 2011, doi: 10.1016/j.peccs.2011.01.001.
19. S. Y. Lee, T. Alam, J. Kim, J. Lee, and S. Park, "Qualitative analysis of tar based on tar sampling conditions for empty fruit bunch gasification," 2021.
20. J. C. Schmid, U. Wolfesberger, S. Koppatz, C. Pfeifer, and H. Hofbauer, "Variation of feedstock in a dual fluidized bed steam gasifier-influence on product gas, tar content, and composition," *Environ Prog Sustain Energy*, vol. 31, no. 2, pp. 205–215, Jul. 2012, doi: 10.1002/ep.11607.
21. Srinivas, S., Field, R. P., & Herzog, H. J. (2013). Modeling tar handling options in biomass gasification. In *Energy and Fuels* (Vol. 27, Issue 6, pp. 2859–2873). <https://doi.org/10.1021/ef400388u>
22. S. R. Janbarari and A. Taheri Najafabadi, "Simulation and optimization of water gas shift process in ammonia plant: Maximizing CO conversion and controlling methanol byproduct," *Int J Hydrogen Energy*, vol. 48, no. 64, pp. 25158–25170, Jul. 2023, doi: 10.1016/j.ijhydene.2022.12.355.
23. Aspen Plus. (2021). Aspen Plus example of CO2 capture from syngas for IGCC using DEPG.
24. Aspen Plus b. (2021). Aspen Plus example of CO2 capture from syngas for IGCC using MEA.
25. Warren D., S., Daniel R., L., J. D., S., Soemantri, W., Rafiqul, G., & Ng, K. M. (2016). *PRODUCT AND PROCESS DESIGN PRINCIPLES Synthesis, Analysis, and Evaluation* (Fourth; WILEY, Ed.). Lightning Source Inc.
26. Gunnarsson, I., Tengberg, F., Thunman, H., Gustavsson, C., & Larsson, A. (2019). Economic assessment of advanced biofuel production via gasification using cost data from the GoBiGas plant. (April 2018), 217–229. <https://doi.org/10.1002/ese3.271>
27. Turton, R., Shaeiwitz, J. A., & Bhattacharyya, D. (2018). *Analysis, Synthesis, and Design of Chemical*

Processes (Fifth; Prentice hall, Ed.). Pearson Education, Inc.

28. Hoffman, Z. (2005). LSU Scholarly Repository Simulation and economic evaluation of coal gasification with SETS reforming process for power production
29. GEP. (2023). OUTLOOK FOR GREEN AND BLUE HYDROGEN MARKET. Retrieved from <https://www.gep.com/blog/strategy/Green-and-blue-hydrogen-current-levelized-cost-of-production-and-outlook#:~:text=The current levelized cost of blue hydrogen production typically lies,USD 6-11 per MMBtu.>

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Designing Better Plastic Management Processes Through a Systems Approach

John D. Chea<sup>a</sup>, Matthew Conway<sup>b</sup>, Gerardo J. Ruiz-Mercado<sup>c</sup>, Pahola Thathiana Benavides<sup>d</sup>, and Kirti M. Yenkie<sup>b\*</sup>

<sup>a</sup> Oak Ridge Institute for Science and Education, hosted by Office of Research & Development, US Environmental Protection Agency, Cincinnati, OH 45268, USA.

<sup>b</sup> Department of Chemical Engineering, Rowan University, Glassboro, NJ, 08028, USA.

<sup>c</sup> Office of Research & Development, US Environmental Protection Agency, Cincinnati, OH, 45268, USA.

<sup>d</sup> Argonne National Laboratory, Lemont, IL, 60439, USA.

\* Corresponding Author: [yenkie@rowan.edu](mailto:yenkie@rowan.edu)

---

## ABSTRACT

Plastics are widely used for their affordability and versatility across many consumer and industrial applications. However, the end-of-life (EoL) management stage can often lead to releasing hazardous chemical additives and degradation products into the environment. The increasing demand for plastics is expected to increase the frequency of material releases throughout the plastic EoL management activities, creating a challenge for policymakers, including ensuring proper material segregation and disposal management and increasing recycling efficiency and material reuse. This research designed a Python-based EoL plastic management tool to support decision-makers in analyzing the holistic impacts of potential plastic waste management policies. The constructed tool was developed to reduce the complexity of material flow analysis calculations, estimating material releases, and environmental impacts. The utility of the tool was tested through the hypothetical nationwide adoption of an extended producer responsibility (EPR) program. The decision-making capability of the tool can facilitate the prediction of long-term outcomes, offering technical knowledge and insight for policymakers seeking to mitigate the environmental and health impacts of plastic pollution.

---

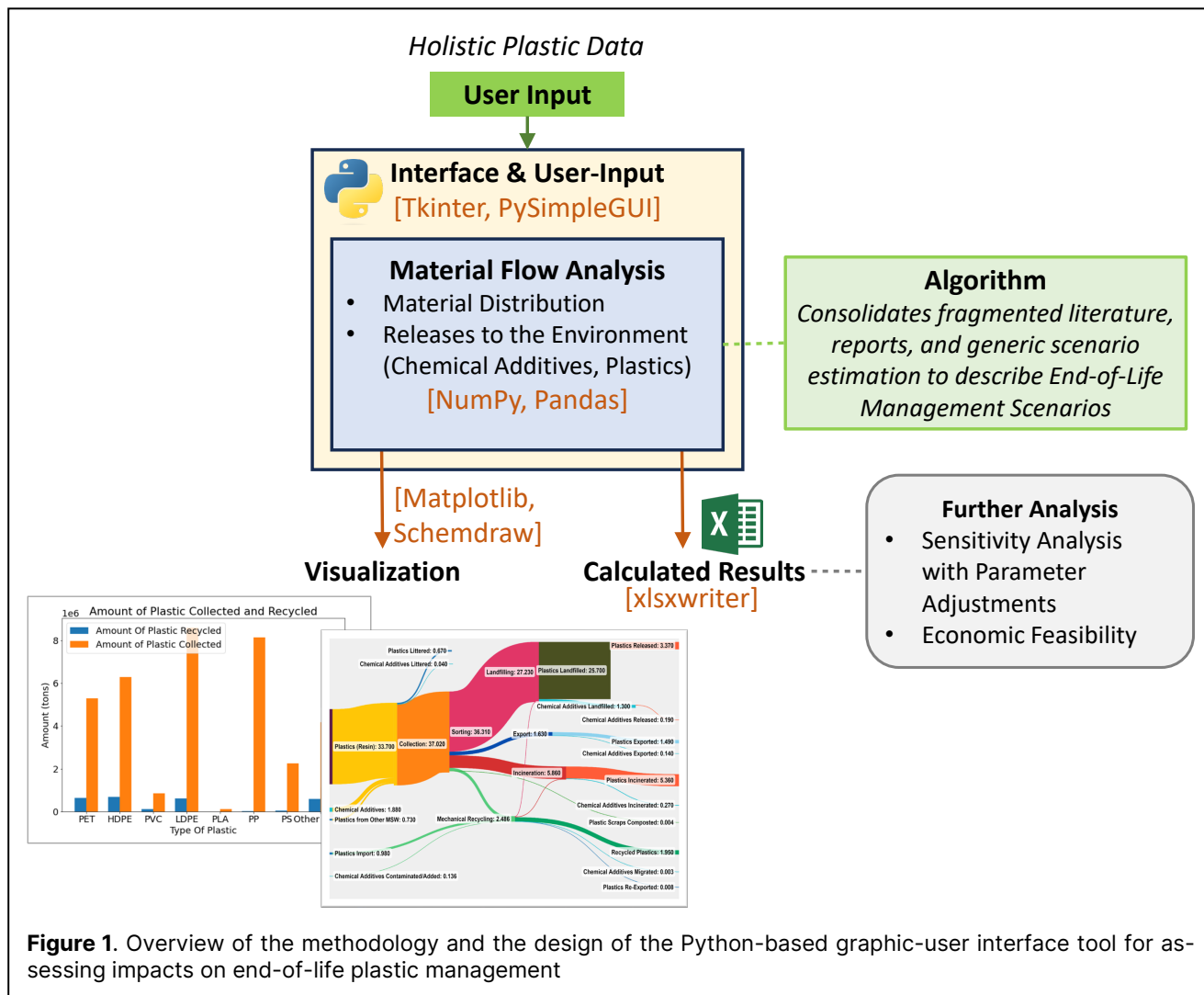
**Keywords:** Polymers, Supply Chain, Modelling

## INTRODUCTION

Plastics are used in many modern applications due to their versatility [1]. Despite their benefits, managing end-of-life (EoL) plastics can release potentially hazardous materials, such as chemical additives, micro- and nano-plastics, and degradation by-products, into the environment [2], [3], [4]. In most cases, chemical additives used during plastic production are not chemically linked to the polymer chain and may migrate to the surrounding environment based on the molecular weight, temperature, compatibility, and solubility in the surrounding medium [5]. These migration mechanisms continue to pose challenges during end-of-life management. The unpredictable nature of plastic waste behavior and the uncertain pathways of chemical additive releases

necessitate a deeper analysis of their impact on the environment, human health, and safety. To minimize the complexities of EoL plastic management analysis, this research introduces a Python-based modeling tool for conducting a holistic analysis of the entire plastic life cycle, simplifying the material flow analysis and scenario-testing process to support decision-making by policymakers and manufacturers. This tool features a holistic overview of plastics and chemical additives released to the environmental compartment (land, water, and air) associated with common EoL management scenarios.

The increasing demand for plastics is expected to double the production rate between 2019 and 2040. Without significant change, the ocean can contain more plastics by weight than fish by 2050 [6]. Reducing plastic pollution requires understanding



the challenges involving all aspects of plastic management before implementing pollution prevention strategies [7].

Chea et al. previously analyzed the existing management of plastics with a material flow analysis to identify potential impacts of releases and exposure scenarios on safety, human health, and the environment [8]. A Python-based modeling tool was initially developed to decrease the complexity of the material flow analysis calculations for decision-making and estimating chemical additive releases and environmental impact [9]. However, the tool was not optimized to model plastics and additives releases under realistic scenarios. This research extended the initial version of the tool by including sensitivity analyses centering on assisting with forecasting the impacts of policies on EoL plastic management efficiency, material releases, and environmental implications. In addition, scenario visualization and graphic user interface (GUI) were improved to streamline the ease of analysis for potential stakeholders. The utility

of the constructed tool was tested in the theoretical implementation of the Extended Producer Responsibility (EPR) program. As a multi-stakeholder problem, the relationships between the potential recycling rate increase and material releases and cost were analyzed. Therefore, the forecasting capability of the tool can assist users with predicting the effects of parameter changes to EoL management practices, ensuring that alterations to the existing designs of EoL management infrastructure are justifiable.

## MATERIALS AND METHODS

### Tool Design and Objectives

The constructed tool is intended to provide a holistic insight into the impacts of policy implementations by using a material flow analysis to estimate relative mass flow intensity and material releases into the environmental compartments (land, water, and air) at various stages throughout EoL management.

The GUI component was constructed in Python using Tkinter and PySimpleGUI modules. Data handling and analysis within this tool utilized NumPy and Pandas. Visualization of the analysis is handled using Matplotlib and Schemdraw. All calculated values from the material flow analysis are transferred through the xlsxwriter module. This method allows for ease of data collection that can be conveniently processed by users unfamiliar with coding. The methodology used to develop a comprehensive assessment tool for evaluating the impacts of policy implementations on EoL plastics management is visually summarized in **Figure 1**.

## Data Collection and Algorithm

The tool user may customize input data as needed to model their community closely. Default values from government reports are available as a starting point without data. The algorithm for completing the material flow analysis is based on a bottom-up approach, which combines and processes the data collected from individual studies to illustrate the complexity of plastic EoL management [8].

**Table 1.** Common Chemical Additives Found in Plastics (Adapted from [2], [10])

Type	Composition Range (%)
Plasticizers	10 – 70
Flame Retardants	3 – 25
Antioxidants	0.05 – 3
UV Stabilizers	0.05 – 3
Heat Stabilizers	0.05 – 3
Slip Agents	0.1 – 3
Lubricants	0.1 – 3
Antistatics	0.1 – 1
Curing agents	0.1 – 2
Blowing agents	0.5 – 20.5
Biocides	0.001 – 1
Colorants	0.25 – 5
Pigments	0.001 – 10
Fillers	0 – 50
Reinforcements	15 – 30

The data used for the algorithm development was sourced primarily through published research data and government reports [2], [11], [12], [13], [14], [15]. The US Environmental Protection Agency (EPA) reported 35.7 million tons of EoL plastic during 2018, which included 14.8% polyethylene terephthalate (PET), 17.6% high-density polyethylene (HDPE), 2.4% polyvinyl chloride (PVC), 24.1% low-density polyethylene (LDPE), 0.3% polylactic acid (PLA), 22.8% polypropylene (PP), 6.3% polystyrene (PS), and 11.7% categorized as other plastics [14]. The chemical additive content of these plastics may vary based on

the intended applications. **Table 1** provides common chemical additives found in plastics and the associated range of usage concentration, demonstrating the uncertainties inherently caused by the knowledge gap between manufacturing and EoL management. These data were converted into usable formats, serving as the basis for conducting the material flow analysis and scenario testing within the tool.

## Case Study Formulation

The utility of the constructed tool was demonstrated through a case study involving adopting a nationwide Extended Producer Responsibility (EPR) program. This program has been widely adopted by the European Union (EU) to enhance plastic waste management strategies, in which producers are expected to design products with the eventual intention of recovery in the EoL management stage and finance the recovery efforts through fees. Plastic packaging waste, in particular, constitutes up to 40.7% of the total plastic waste generated in 2018, 13.6% of which was successfully recycled [14].

The variation in packaging types, collection methods, and existing recycling infrastructure throughout the US attests to the low recycling rate observed. Decision-makers may use the constructed tool to test the effects of implementing the EPR program, which is mandatory to meet specific recycling goals before adopting the regulations.

As of April 2023, six US states (CA, CO, ME, NJ, OR, and WA) have implemented packaging EPR law [16]. Although the focus varies between states, these laws require producers to make packaging recyclable or compostable, achieve higher recycling rates for single-use plastics, reduce plastic packaging volumes, and incorporate post-consumer recycled content into various products. The existing EPR infrastructure within the US includes (1) having a producer responsibility organization collect waste, process, and move waste as needed and (2) allowing municipalities to manage waste then be reimbursed by the producer responsibility organization [17].

These efforts aim to promote a circular economy, reduce landfill waste, encourage sustainable packaging practices, and upgrade recycling systems within each state.

This assessment primarily targeted the enhancement of plastic packaging recycling, the projected increase in recyclable volumes, and estimated potential reductions in releases to environmental compartments (land, water, and air). These releases include EoL plastics, chemical additives, and other byproducts created because of conventional plastic EoL management. An estimation of the economic



recovery from material reclamation from nationwide EPR implementation was also completed. It should be noted that the calculated values are based on plastic EoL management performance from 2018 and do not necessarily represent management trends from other years.

## RESULTS & DISCUSSIONS

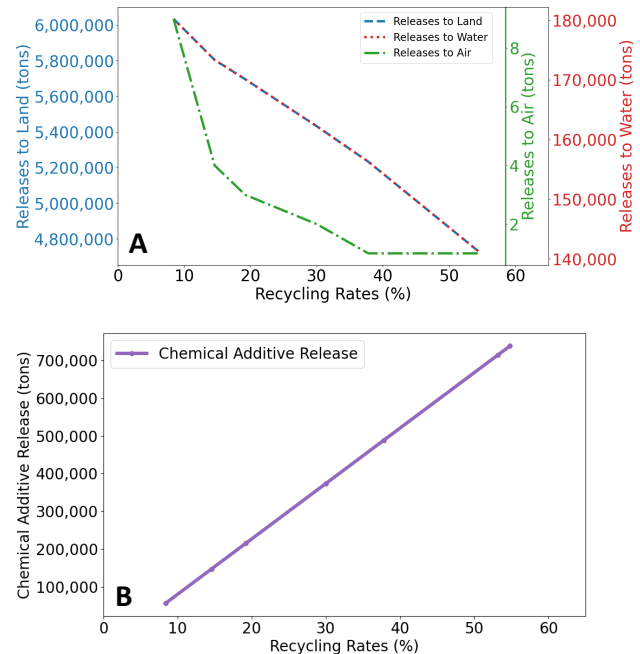
### Recycling Rate Projections and Environmental Implications

The previous study on EoL plastic management in the United States, based on 2018 MSW data, identified that over 35.7 million tons of plastic waste were generated, with only approximately 8.4% successfully recycled [8]. Their material flow analysis highlighted primary plastic types and their allocation in recycling, landfilling, and incineration, emphasizing the limited recycling efforts and substantial landfilling of EoL plastics. Moreover, the potential range of chemical additives in plastic products and the potential environmental implications were reported, emphasizing the need for enhanced processing infrastructure to mitigate environmental releases.

This analysis sets the stage for a comparative analysis between existing EoL plastic management practices and the potential effects of EPR programs, specifically focusing on the EoL plastic packaging component. There is a positive correlation between the recycling rate of packaging plastics and the effective adoption of EPR programs nationwide. This program can effectively reduce the number of plastics sent to the traditional EoL management by requiring manufacturers to take responsibility for their products throughout the material life cycle. Therefore, increased plastic collection and recovery rates and decreased plastic mass in landfills or incineration can be observed.

The effectiveness of EPR programs worldwide has been reported to achieve a packaging recycling rate as high as 75%, while the statistics within the US may theoretically reach a maximum of 48% due to the lack of sustainable funding in various communities throughout the country [18]. Based on the material flow analysis in 2018, over 17 million tons (54.8% of generated EoL plastics) can be removed from the conventional management process. Consequently, material releases to land, water, and air decreases with increasing recovery rate, as illustrated in **Figure 2A**. The recycling rates (%) include practices such as plastic exports and conventional management of non-packaging plastics, which are unaffected by the EPR programs. The inversely proportional relationship observed is expected because the conventional EoL plastic management process can subject the collected materials to many generic steps (reprocessing, incineration, landfilling, composting, transport) because of the knowledge gap between the manufacturer and EoL plastic

management workers. Additional processing steps can lead to the propagation of efficiency issues and material releases into the environment. The EPR program minimizes processing steps by involving the producers with EoL management, reducing the frequency of material releases into the environment.



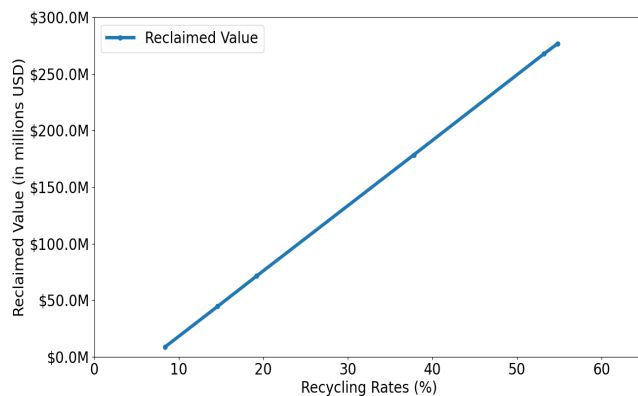
**Figure 2.** Extended producer responsibility scenario-specific plot generated from the python-based modeling tool. (A) The potential reduction in environmental releases based on the effectiveness of Extended Producer Responsibility (EPR) program. (B) The potential chemical additive releases to the environment as a function of material reprocessing.

The producers are expected to possess the most information regarding the composition and processability of the EoL plastics they created. However, concerns may be raised regarding the presence of chemical additives used. If subjected to physical processing, such as extrusion, chemical additives may migrate from the polymer matrix and be released into the surrounding environment. The extrusion process also can release VOCs, monomers, and other degradation products because of the heat generated from friction [2]. **Figure 2B** illustrates the direct relationship between the plastic recycling rate and the release of chemical additives to the environmental compartments. The recycling rate (%) was varied between the 2018 plastic recycling rate (8.4% wt.) [13] to the maximum recycling rate anticipated by EPR programs in the US (54.8% wt.) [18] to demonstrate the differences between maintaining the existing material management paradigm or implementing EPR programs nationwide. Although EoL management steps have been reduced through the EPR program, reprocessing EoL plastics can

cause unintentional releases and contamination, leading to hazardous workplace exposure [8]. The chemical additive releases occur as a result of the reliance on additives as processing aids in the physical recycling practices. This process subjects plastics to high friction-induced heat, increasing the likelihood of chemical migration to the surrounding environment, especially because chemical additives are generally not linked chemically with the plastic products [2]. The release of chemical additives may be minimized if the EoL plastic structural integrity remains intact and reuse is possible with minimal cleaning. However, additional treatment steps may be required to remove potentially toxic or unknown contaminants. Such an endeavor can be cost-intensive and require case-specific economic analysis.

### Designing Economical Plastic Recycling Infrastructure

Material value reclamation is one of the most crucial economic factors to consider before adopting the EPR program as part of the conventional plastic management processes. **Figure 3** illustrates a proportional relationship between the effective recycling rate of packaging waste and the reclaimed value. Similar to **Figure 2**, the recycling rate (%) was varied between the 2018 recycling rate and the maximum recycling potentials due to implementing the EPR programs in the US (54.8% wt.) The high recovery rate can reduce the raw material requirement for creating new plastic packaging products. This event can incentivize producers to design more recyclable products, promoting the circular economy concept [1].



**Figure 3.** The effects of increasing EoL plastic recycling rate on the potential reclaimed material values following Extended Producer Responsibility (EPR) program

A comprehensive economic analysis at the national level is presently beyond the scope of the holistic assessment tool. Additionally, market variations throughout the US can make a comprehensive economic analysis challenging. Manufacturers may encounter different investment costs for EPR programs ranging between 0 – 1% of their gross revenue, depending on the specific demands

of their markets [19]. Notably, the federal government assumes no financial burden in adopting EPR programs, with all associated costs shifted onto producers and consumers [20]. It is also crucial to recognize that the success of the EPR framework is dependent on effective collaboration between stakeholders, robust regulatory frameworks, and continuous monitoring and adaptation to evolving market conditions. Such measures are essential to ensure that the program remains economically viable and sustainable and contributes meaningfully to the circular economy.

Since EPR inherently subtracts gross revenue from producers, incentives should be offered to encourage participation. Maine has taken the initiative to pass the EPR program for packaging in July 2021, providing reimbursement to municipalities that choose to participate in improving the recycling infrastructure. Exemption to the law was offered for producers that made less than \$2 million in gross revenue threshold or used less than 1 ton of packaging materials. Oregon shortly passed the same law, which gave rise to the Producer Responsibility Organization (PRO), designed to improve and expand recycling service, including funding waste prevention grants and studies related to recycling systems. Other states that have adopted EPR programs have adopted distinct approaches to determining producer fees and recycling goals [21]. However, it can generally be said that incentives and exemptions have demonstrated their usefulness in pioneering EPR programs by encouraging participation from producers, driving behavioral change throughout the plastic life cycle, and offering a competitive advantage regarding sustainability claims.

### Data Limitations

It should be noted that the current version of the tool utilizes public plastic management data in 2018 [13]. Therefore, the calculated values from the analysis are not entirely representative of data from the previous year. However, this approach remains valid if the tool user aims to discover the potential trends from potential policy implementation. Future improvements to the tool may include data from other years to increase the model accuracy and account for variation over time. Learning from the data approach can also be included to determine potential correlations for future projections. Policymakers should consider the complexities of plastic management, releases, and costs while drafting strategies for enhancing plastic waste management.

## CONCLUSIONS

Addressing EoL plastic management challenges requires strategic approaches to ensure maximum material recovery and minimize environmental

releases and reduce cost. The Python-based tool developed herein automates the material flow analysis process, describing plastic flow allocation throughout the life cycle and possible releases to the environmental compartments to aid decision-making.

This tool was used to simulate the implementation of EPR programs within the US from a holistic perspective, which correlated increasing recycling rates to reducing EoL plastics amount within traditional EoL management processes. Therefore, the US recycling infrastructure can shift toward the circular economy state, emphasizing the importance of reclaiming material value to encourage producers to design products with improved inherent recyclability. Ultimately, raw material costs during manufacturing can be reduced. However, there is an inversely proportional relationship between the recycling rate (8.4% base case to 54.8% maximum) and environmental releases because the chemical releases that occur during conventional plastic recycling are unmitigated. Alternative strategies for reducing chemical additive releases and exposure risks are required before the recycled plastics can safely be reused in the subsequent life cycle.

Improving the EoL plastic management is a multi-stakeholder problem. Unexplored factors, such as market variations and diverse manufacturer investment costs between regions, could impede the nationwide implementation of the EPR program. It should be noted that the EPR program case study serves as an example of the application of the tool to improve the management of EoL plastic packaging. Policies affecting the management of other plastic groups and types may also be tested to ensure a feasible EoL plastic management infrastructure design and foster a circular economy.

## ACKNOWLEDGEMENTS

This research was supported in part by an appointment to the US Environmental Protection Agency (EPA) Research Participation Program administered by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between the US Department of Energy (DOE) and the US EPA. ORISE is managed by ORAU under DOE contract number DE-SC0014664.

## DISCLAIMER

The views expressed in this article are those of the authors and do not necessarily represent the views or policies of the EPA. Any mention of trade names, products, or services does not imply an endorsement by ORAU/ORISE, the US Government, or the EPA. The EPA

does not endorse any commercial products, services, or enterprises.

## REFERENCES

1. R. Barra and S. A. Leonard, "Plastics and the circular economy," Scientific and Technical Advisory Panel to the Global Environment Facility, 2018.
2. J. N. Hahladakis, C. A. Velis, R. Weber, E. Iacovidou, and P. Purnell, "An overview of chemical additives present in plastics: Migration, release, fate and environmental impact during their use, disposal and recycling," *Journal of Hazardous Materials*, vol. 344, pp. 179–199, Feb. 2018, doi: 10.1016/j.jhazmat.2017.10.014.
3. A. Herrera *et al.*, "Bioaccumulation of additives and chemical contaminants from environmental microplastics in European seabass (*Dicentrarchus labrax*)," *Science of The Total Environment*, vol. 822, p. 153396, May 2022, doi: 10.1016/j.scitotenv.2022.153396.
4. I. Wojnowska-Baryła, K. Bernat, and M. Zaborowska, "Plastic Waste Degradation in Landfill Conditions: The Problem with Microplastics, and Their Direct and Indirect Environmental Effects," *IJERPH*, vol. 19, no. 20, p. 13223, Oct. 2022, doi: 10.3390/ijerph192013223.
5. T. R. Crompton, *Additive migration from plastics into food: a guide for analytical chemists*, New ed. Shawbury: Smithers Rapra Technology, 2007.
6. World Economic Forum, "The New Plastics Economy - Rethinking the Future of Plastics," Ellen Macarthur Foundation and McKinsey & Company, 2016. [Online]. Available: <http://www.ellenmacarthurfoundation.org/publications>
7. EPA Office of Resource Conservation and Recovery, "Draft National Strategy to Prevent Plastic Pollution," US Environmental Protection Agency, 2023.
8. J. D. Chea, K. M. Yenkie, J. F. Stanzione, III, and G. J. Ruiz-Mercado, "A Generic Scenario Analysis of End-of-Life Plastic Management: Chemical Additives," *Journal of Hazardous Materials*, p. 129902, Sep. 2022, doi: 10.1016/j.jhazmat.2022.129902.
9. J. D. Chea, M. Conway, A. Lehr, G. J. Ruiz-Mercado, and K. M. Yenkie, "TRACKING CHEMICAL ADDITIVE RELEASES IN THE PLASTICS END-OF-LIFE MANAGEMENT STAGE TO CLOSE THE LOOP," presented at the Foundations of Computer Aided Process Operations / Chemical Process Control, 2023.
10. E. Hansen, "Hazardous substances in plastic

materials," COWI-Denmark and Danish Technological Institute, 2013.

11. O. Horodytska, A. Cabanes, and A. Fullana, "Non-intentionally added substances (NIAS) in recycled plastics," *Chemosphere*, vol. 251, p. 126373, Jul. 2020, doi: 10.1016/j.chemosphere.2020.126373.
12. J. R. Jambeck *et al.*, "Plastic waste inputs from land into the ocean," *Science*, vol. 347, no. 6223, pp. 768–771, Feb. 2015, doi: 10.1126/science.1260352.
13. US EPA, "National Overview: Facts and Figures on Materials, Wastes and Recycling," United States Environmental Protection Agency. [Online]. Available: <https://www.epa.gov/facts-and-figures-about-materials-waste-and-recycling/national-overview-facts-and-figures-materials>
14. US EPA, "Advancing Sustainable Materials Management: 2018 Tables and Figures - Assessing Trends in Materials Generation and Management in the United States," United States Environmental Protection Agency, 2020.
15. E. U. T. van Velzen, M. Jansen, M. T. Brouwer, A. Feil, K. Molenveld, and Th. Pretz, "Efficiency of recycling post-consumer plastic packages," presented at the PROCEEDINGS OF PPS-32: The 32nd International Conference of the Polymer Processing Society - Conference Papers, Lyon, France, 2017, p. 170002. doi: 10.1063/1.5016785.
16. Source Intelligence, "Packaging EPR Laws in the US," Source Intelligence. [Online]. Available: <https://blog.sourceintelligence.com/packaging-epr-laws-in-the-us>
17. National Academies of Sciences, Engineering, and Medicine, *Recycled Plastics in Infrastructure: Current Practices, Understanding, and Opportunities*. Washington, DC: The National Academies Press, 2023, p. 27172. doi: 10.17226/27172.
18. The Recycling Partnership, "Study of Extended Producer Responsibility Policy Across 7 Jurisdictions Worldwide Shows it Dramatically Increases Recycling Rates," The Recycling Partnership. [Online]. Available: <https://recyclingpartnership.org/study-of-extended-producer-responsibility-policy-across-7-jurisdictions-worldwide-shows-it-dramatically-increases-recycling-rates/>
19. Natural Resources Council of Maine, "Real World Data Shows That There is No Correlation between EPR for Packaging and Consumer Prices." 2023. [Online]. Available: <https://www.nrcm.org/wp-content/uploads/2022/06/talking-pts-consumer-prices-epr.pdf>
20. Recycling Reinvented, "Extended Producer Responsibility Cost-Benefit Study." Recycling Reinvented, 2014. [Online]. Available: <https://www.globalpsc.net/wp-content/uploads/2014/03/RR-EPR-MN-Study-Working-Paper-3-EPR-Cost-Benefit-Study-0314.pdf>
21. National Conference of State Legislatures, "Extended Producer Responsibility," National Conference of State Legislatures, 2023. [Online]. Available: <https://documents.ncsl.org/wwwncsl/Environment/Extended-Producer-Responsibility-f01.pdf>

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Resilient-aware Design for Sustainable Energy Systems

Natasha J. Chrisandina<sup>a,b</sup>, Shivam Vedant<sup>b,c</sup>, Catherine Nkoutche<sup>a,b</sup>, Eleftherios Iakovou<sup>b,d</sup>, Efstratios N. Pistikopoulos<sup>a,b</sup>, Mahmoud M. El-Halwagi<sup>a,b\*</sup>

<sup>a</sup> Texas A&M University, Artie McFerrin Department of Chemical Engineering, College Station, Texas, USA

<sup>b</sup> Texas A&M University, Texas A&M Energy Institute, College Station, Texas, USA

<sup>c</sup> Texas A&M University, Department of Multidisciplinary Engineering, College Station, Texas, USA

<sup>d</sup> Texas A&M University, Department of Engineering Technology and Industrial Distribution, College Station, Texas, USA

\* Corresponding Author: [el-halwagi@tamu.edu](mailto:el-halwagi@tamu.edu).

## ABSTRACT

To mitigate the effects of catastrophic failure while maintaining resource and production efficiencies, energy systems need to be designed for resilience and sustainability. Conventional approaches such as redundancies through backup processes or inventory stockpiles demand high capital investment and resource allocation. In addition, responding to unexpected “black swan” events requires that systems have the agility to transform and adapt rapidly. To develop targeted solutions that protect the system efficiently, the supply chain network needs to be considered as an integrated multi-scale system incorporating every component from individual process units all the way to the whole network. This approach can be readily integrated with analogous multiscale approaches for sustainability, safety, and intensification. In this work, we bring together classical supply chain resilience with process systems engineering to leverage the multi-scale nature of energy systems for developing resilience enhancement strategies that are resource-efficient. In particular, we adapt qualitative risk analysis methods to uncover critical system components and major vulnerabilities to guide resource allocation decisions. To account for these vulnerabilities, we explore the feasible region of operation around each node of the supply chain. An optimization formulation is devised to generate multiscale alternative. The approach is demonstrated through a case study involving the production of biofuels, demonstrating the range of adaptation strategies possible when process-level strategies are incorporated into overall supply chain design.

**Keywords:** Supply Chain, Multiscale Modelling, Planning & Scheduling, Renewable and Sustainable Energy, Energy Systems

## INTRODUCTION

The study of energy system resilience, which is typically defined as the ability of systems to manage possible disturbances and recover from them, has increased in importance over the last decade as energy systems become more complex and disturbances become more frequent and severe [1, 2]. The integration of multiple energy sources, including intermittent and geographically-dispersed renewables, facilitates the transition to a more sustainable energy system, but also introduces complexity and vulnerability [2]. Therefore, energy systems need to be designed and operated in such a way that they are prepared to adapt and recover from future disturbances whether or not these disturbances are expected.

In minimizing the impact that a particular

disturbance has on energy system infrastructure, resilience contributes to the sustainability of the energy system in that less waste in the form of broken equipment and unused raw material or product is generated [3, 4, 5]. Additionally, systems that are better able to withstand disturbances are also able to continue operating for a longer period of time, maximizing their useful lifetimes. However, commonly-deployed resilience strategies center around reserving resources (either raw materials or finished products) as safety inventory to maintain high production output for as long as possible [6]. This strategy is useful, but relying solely on storing large amounts of unused inventory also raises its own sustainability and safety concerns [3]. For example, large-scale storage of hydrogen for energy currently relies on artificially-built salt caverns or depleted natural gas reservoirs, but there



may be possible reactions between the stored hydrogen with the microorganisms and mineral constituents of the reservoir, leading to deterioration of the hydrogen storage or unwanted deposits of reaction products [7]. Furthermore, long-term backup inventory storage can represent a financial drain on companies, which may make it economically unattractive. Therefore, cost-effective targeted resilience enhancement strategies need to be considered to balance the issues of resilience, sustainability, and economic cost.

Consideration of energy system resilience (and adjacent concepts) independently and alongside sustainability has been demonstrated in the open literature. Panteli et al. [8] proposed that power system resilience consists of two separate characteristics: operational resilience, which corresponds to the ability to ensure uninterrupted power supply, and infrastructure resilience, which refers to the physical strength of the power system to mitigate faulty portions of the system. Moreno-Sader et al. [9] proposed the use of a modified return-on-investment (ROI) metric that includes safety, sustainability, resilience, and reliability weights to screen process alternatives early in the design process. The weights are calculated based on whether a proposed design reaches a desired target value for each of the four desired objectives. Hosseini-Motlagh et al. [10] designed a power supply chain to minimize unmet electricity demand as well as pollution emission under uncertainty.

The complexities within an energy system not only come from supply chain dynamics such as material flow, customer energy demand, and transportation linkages, but also in the chemical reactions that occur within manufacturing facilities. Harnessing the physical and chemical synergies in energy systems through multi-scale systems engineering could be key to unlocking a variety of targeted resilience strategies that do not require significant resource or capital investments [11]. With a multi-scale approach, tools and methods built for supply chain, process, unit operations, and reaction-scale optimization are integrated to provide an accurate representation of the interactions across spatio-temporal scales [11, 12]. The approach allows bespoke models to be constructed that contains appropriate levels of detail at relevant spatio-temporal scales to open up the possibility for targeted resilience strategies.

Multi-scale approaches have been utilized in the design and optimization of energy systems to achieve economic and/or sustainability goals. Demirhan et al. [12] developed a multi-scale model for an energy system which uses solar and wind resources to supply electricity via various storage technologies at the lowest possible cost. The model accounts for renewable resource availability, storage technology constraints, as well as demand fluctuations. Shao et al. [13] proposed a multi-scale model for a hydrogen-based off-grid microgrid to generate both

power and heat for rural areas. A two-stage stochastic formulation was used to derive optimal capacity sizing and scheduling for both normal and on-emergency scenarios. Lin [14] explored the life-cycle impact of different technology pathways to develop future energy systems that integrate renewable resources, battery storage, and dense energy carrier production.

For a holistic consideration of energy system resilience, a multi-scale modeling approach can be beneficial to integrate risk factors and resilience enhancement strategies on the supply chain level and on the component level. This paper, therefore, aims to illustrate how process-level considerations contribute to overall supply chain resilience in a cost-competitive and sustainable manner. A qualitative resilience analysis is conducted to identify critical model variables that directly affect overall system performance. Next, an integrated system model is constructed around key variables and optimized. What-if scenarios are applied on the optimized network to demonstrate its resilience against several supply chain disturbances. The proposed methodology is demonstrated through a small regional biofuel supply chain.

## METHODOLOGY

### Consideration of Process Feasible Regions

In a typical supply chain optimization formulation, the manufacturing facilities that produce goods to sell are modeled as nodes with a fixed production rate; that is, given some quantity of raw materials, the quantity of products made is known [3, 6]. Responding to disruptions involves adding redundancies into the supply chain superstructure through alternate suppliers or transport routes, excess inventory, and overdesign of manufacturing capacity. Conversely, in process systems literature, a manufacturing facility is modeled as a set of processes (chemical or physical) that can be described by their operating parameters. Given a process design, it is well-known that there exists a region of feasible operation where different operating parameters will yield a different amount of product [15].

In our previous work, we demonstrated the benefits of designing processes with high reliability on enhancing overall supply chain resilience under disruptions [16]. Reliability is achieved through choosing to install process equipment with low failure rates and high repair rates. In this work, we expand on the concept of implementing process-level strategies to enhance supply chain resilience by considering the process feasible region within the supply chain optimization problem [17]. In effect, each manufacturing facility will be represented as a node with operational parameters that can be optimized for different scenarios. This necessitates a mathematical model for the process node that is then integrated into the supply chain optimization formulation.

To formulate a mathematical model of each process, computer-aided simulation software (e.g. Aspen Plus) and functional equations of key units within the module are first used to generate a computer model of the module. Next, the operating parameters that are most critical to process output are identified. A surrogate model for each critical unit is then generated to represent the relationship between the critical parameters with process output. These surrogate models are then integrated as additional constraints in the supply chain formulation. A generalized form of the surrogate model is shown in Eq. 1, where the unit-level parameter realizations  $x_{u,f,t}$  determine the unit efficiency  $Q_{u,f,t}$ .

$$Q_{u,f,t} = \Phi(x_{u,f,t}) \quad (1)$$

## Integrated Model

In this work, the energy system is assumed to be composed of a set of supplier cities that contain a set amount of raw material and could also contain modular refineries of a fixed nameplate capacity for production of a dense energy carrier (DEC). Finally, the product is transported to market cities to fulfill the required demand. Mass balance constraints for a similar system is available in Chrisandina et al. [11]. Additional constraints to include process information are outlined in this section.

## Nomenclature

The symbols used for the rest of the paper are defined in Table 1 below.

**Table 1:** Symbols and definitions for sets, parameters, and variables

Notation	Description
<u>Sets</u>	
$\mathcal{F}$	Set of manufacturing sites
$S$	Set of supplier sites
$M$	Set of market sites
$T$	Set of scheduling time periods
$U$	Set of process units
<u>Parameters</u>	
$yield_f$	Yield of product in site $f$
$G_{max_{m,t}}$	Demand for product in market $m$ at time $t$
$d_{i,j}$	Distance between two locations $i$ and $j$
$service_m$	Minimum service rate for market $m$
$x_u^{Nom}$	Nominal parameter value for unit $u$
$C_t$	Cost of shipping feedstock [\$/mile-ton]
$C_b$	Cost of feedstock [\$/ton]
$C_p$	Cost of shipping product [\$/mile-ton]
<u>Variables</u>	
$G_{f,t}$	Production from site $f$ at time $t$
$G'_{f,m,t}$	Product delivered from site $f$ to market $m$ at time $t$
$B_{s,f,t}$	Raw material from supplier $s$ consumed

	by site $f$ at time $t$
$Q_{f,t}$	Efficiency of process modules in site $f$ at time $t$
$x_{u,f,t}$	Realized parameter value for unit $u$ in site $f$ at time $t$
$Q_{u,f,t}$	Efficiency for unit $u$ in site $f$ at time $t$
$mod_{f,t}$	Number of process modules in site $f$ at time $t$
<u>Generalized functions</u>	
$\Phi(x_{u,f,t})$	Correlation between unit-level parameters and unit efficiency
$\Lambda_u(Q_{u,f,t})$	Correlation between unit efficiency and overall process module efficiency
$\Psi_f(mod_{f,t} T )$	Correlation between number of process modules and capital expense

## Constraints

These constraints are modified from the original mass balance constraints to include process efficiency information.

The product yield at each manufacturing site is governed by the amount of raw material supplied to the site and the efficiency of the process modules placed on site.

$$G_{f,t} = yield_f \times \sum_s B_{s,f,t} \times Q_{f,t} \quad (2)$$

The efficiency of a process module is defined as a multiplier on the nominal process output which depends on the operating conditions to which the process module is set. The process module efficiency is governed by the efficiencies of the critical units within the process module. The specific function that relates unit efficiency with process module efficiency depends on the exact configuration of the module.

$$Q_{f,t} = \Lambda_{u \in U}(Q_{u,f,t}) \quad (3)$$

## Objectives

The main objective of the optimization problem is to minimize total annual cost. This includes the annualized fixed cost in the purchase of process modules, as well as the operating cost (raw material purchase and transportation). The purchase cost is a function of the number of process modules deployed in the supply chain, and the specific function depends on how capacity scales with cost. The annualized fixed cost (AFC) is calculated by dividing the total capital cost across the lifetime of the process modules.

$$Cost = AFC + OPEX \quad (4a)$$

$$CAPEX = \Psi_{f \in \mathcal{F}}(mod_{f,t=|T|}) \quad (4b)$$

$$AFC = \frac{CAPEX}{years} \quad (4c)$$

$$OPEX_t = \sum_f^F \sum_s^S B_{s,f,t} \times Cb + \sum_f^F \sum_s^S (B_{s,f,t} \times d_{s,f} \times Ct) + \sum_f^F \sum_m^M (G'_{f,m,t} \times d_{f,m} \times Cp) \quad (4d)$$

To represent system resilience, we include a constraint on the target demand fill rate required at every scheduling time step.

$$\sum_f^F G'_{f,m,t} \geq G_{max_{m,t}} \times service_m \quad (5)$$

## CASE STUDY

In this work, we adapt the case study introduced by Lopez-Molina et al. [18] where refuse-derived fuel (RDF) is converted to methanol and then sold as fuel. Gasification-based technology is used, and a modular biorefinery is assumed for each manufacturing site. The goal of this case study is to compare the resilience of the supply chain to various external disruptions with and without the incorporation of process surrogate models in the supply chain formulation.

### Background Information

The RDF-to-methanol conversion process occurs via gasification to produce syngas as an intermediary, purification of the syngas, and methanol synthesis as a final step. Technical and operational details of the process are given in the literature [19, 20], and key process inputs

and outputs are shown in Fig. 1. A simplified process flow diagram is shown in Fig. 2. In this work, linear surrogate models (see Eq. 6) are generated for each critical parameter of the process, as listed in Table 2. Linear surrogates are chosen due to the small range of possible values for each critical parameter.

$$Q_{u,f,t} = m_u \frac{x_{u,f,t}}{x_{u,0}} + b_u \quad (6)$$

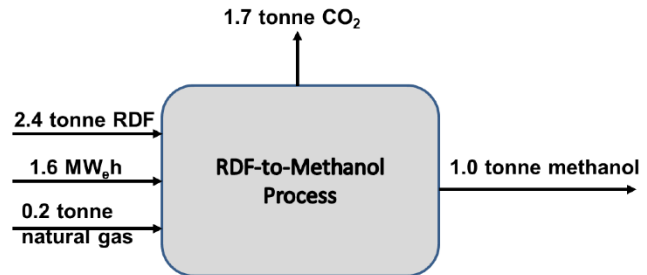


Figure 1: Key process inputs and outputs [18]

We consider the case of six cities: three producer cities which act as suppliers and manufacturers (A, B, C), and three market cities which purchase methanol (D, E, F). The specific supply availability and demand requirements are given in Fig. 3, and as a base-case scenario a 90% target demand fill rate is assumed. A planning horizon of one year with monthly schedules is assumed. The strategic decision to be made is where to locate process modules, and the operational decisions to be made are how many process modules to deploy, tuning of process

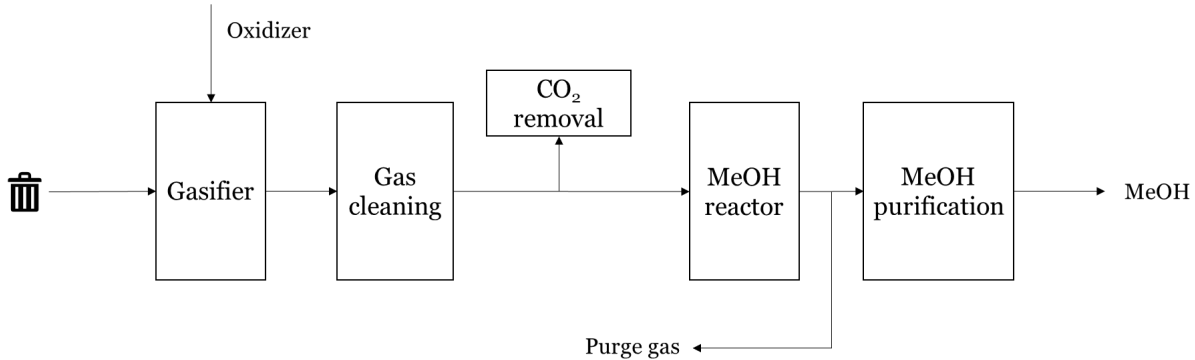
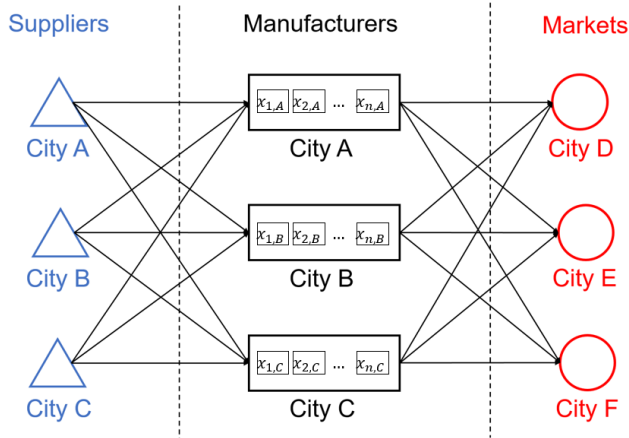


Figure 2: A simplified process flow diagram of the methanol synthesis process

Table 2: Critical parameters for methanol synthesis and their feasible operating range [20]

	Nominal operating condition ( $x_{u,0}$ )	Minimum	Maximum	Gradient ( $m$ )	Intercept ( $b$ )
MSW moisture content [%]	25	15	40	-0.24	1.23
Gasifier temperature [°C]	850	700	900	1.72	-0.72
Gasifier equivalence ratio	0.42	0.35	0.6	-1.31	2.25
Reactor pressure [bar]	60	40	90	0.68	0.28
Recycle ratio	4.75	4.05	5.5	0.26	0.73

parameters in a location (listed in Table 2), and the connection between supplier and market cities.



**Figure 3:** The supply chain superstructure

## RESULTS AND DISCUSSION

### Qualitative Resilience Analysis Results

In order to identify major risks to the supply chain, a table of potential failure modes and mitigation strategies has been constructed (see Table 3). Failure modes on supply chain entities (suppliers, logistic providers, manufacturing sites, and markets) as well as on process components (gasifier, syngas purifier, methanol reactor) are

considered. It can be seen that there are multiple potential mitigation strategies for each entity failure mode and that in many cases operational changes on the process module level (a sub-entity of a manufacturing site) can help in failure mitigation. It is expected, therefore, that simultaneous optimization of process and supply chain parameters will yield a supply chain schedule that is more resilient to these failure modes. The next sub-section will aim to quantify these benefits.

### Impact of Process Tuning on Supply Chain Performance

In the base-case scenario, processes are operated at their nominal parameter values ( $Q_{f,t} = 1$ ) with a total cost of \$317MM. Several failure scenarios are tested to see if the supply chain is able to adjust. Table 4 shows the feasibility of different failure scenarios that occur in the beginning of the planning horizon. It can be seen that reacting to deviations away from the base-case scenario without process tuning increases the total cost of the supply chain, as more process modules are needed to compensate for additional production. However, process-level tuning allows for existing modules to increase production within their feasible regions as needed, which significantly reduces the additional capital investment needed to meet the new demand. In these cases, raw material and utility costs contribute to the increase in total economic cost but no new process modules were

**Table 3:** Potential failure modes and mitigation strategies

Entity	Sub-entity	Potential failure mode	Mitigation strategies
Supplier		Decline in supply availability	Multiple suppliers Tune process module to increase yield Forecasting – storage for future
Logistics provider		Technical problems with vehicle	Vehicle maintenance Investment in updated vehicles Reroute vehicles
Manufacturing site	Geographical location	Location experiences natural disaster event	Relocate process modules to other manufacturing locations
	Process module	Module temporarily unavailable	Surge production in other manufacturing sites Repair/replace module
	Gasifier	Temperature Moisture content	
	Syngas scrubber	Absorption failure	
	Methanol reactor	Low yield	
Market		Increase in demand	Acquire new modules Surge production in manufacturing sites
		Demand concentrated in one location	Relocate process modules to manufacturing sites in proximity Ship raw material to manufacturing sites in proximity

purchased. Table 5 shows the process parameter realizations that correspond to various service fill rates, which demonstrates how different process units are operated to meet specific market demands. Furthermore, the deployment of additional process modules represents an increase in material consumption that is used to construct the process modules, meaning that process-level tuning also opens up the possibility for lower emission footprint across the supply chain while delivering the same level of service throughout the year.

**Table 4:** Total economic cost of different supply chain scenarios – comparison between supply chain-only design and supply chain design with process surrogate model included

Scenario	Supply chain only	Supply chain + process
Target fill rate increases to 95%	\$357MM (+13%)	\$330MM (+4%)
Target fill rate increases to 100%	\$367MM (+16%)	\$342MM (+8%)
City E demand increases to 110%, start of year	\$357MM (+13%)	\$332MM (+4%)
City E demand increases to 110% for one month	\$343MM (+8%)	\$318MM (+0%)
City B is partially disrupted	\$344MM (+8%)	\$318MM (+0%)

**Table 5:** Process parameter realization for various service fill rates

Process parameter	Base case	Fill rate 95%	Fill rate 100%
MSW moisture content [%]	25	25	25
Gasifier temperature [°C]	850	850	900
Gasifier equivalence ratio	0.42	0.42	0.42
Reactor pressure [bar]	60	66	66
Recycle ratio	4.75	5	5.2

Some scenarios are also tested where failure modes happen in the middle of the planning horizon. Without process tuning, supply chain operational decisions are updated to address failure. With process tuning, however, process module operational decisions are also updated which provides additional options for failure mitigation. Taking Scenario #4 from Table 4 as an example, if the demand increase occurs in the middle of the year

the process module in City C (the closest manufacturing site) slightly increases its capacity temporarily to meet the additional demand without affecting the operations of the other manufacturing sites or investing in additional process modules. This temporary capacity increase is achieved by increasing the reactor pressure by 10% and recycle ratio by 5%. When the demand level lowers, the process module in City C adjusts its operation again to its baseline level.

A scenario where process-level failure occurs is also tested, where the process module in City B experiences a failure that lowers its production capacity until recovery efforts have concluded. In this scenario, City B experiences a partial failure that affects 50% of its capacity and recovery efforts are concluded at the end of the planning horizon. The other manufacturing sites respond to this failure by increasing their production so that demand can still be met, alleviating the need to invest in a replacement process module.

## CONCLUSION

To achieve a resilient and sustainable energy system, incorporating targeted strategies to mitigate potential failures while maintaining efficient resource utilization and low capital investment is critical. A multi-scale approach to design and optimization of energy systems enables a holistic assessment of potential failure modes and mitigation strategies that operate on multiple spatio-temporal scales. In this work, we have demonstrated a first attempt at simultaneous optimization of supply chain and process design by including critical process parameters as additional constraints within a supply chain model formulation. We showed that the consideration of the feasible region around each process allows the energy system design to operate in various disruption scenarios with low additional capital investment and no additional process module constructions, limiting the economic and environmental cost of these mitigation strategies. Furthermore, we also demonstrated that targeted changes can also be implemented in the middle of the planning horizon on specific nodes without affecting the overall system design or other nodes' operations. To further develop this methodology, more bespoke process surrogate models that account for the relationship between multiple critical parameters will be beneficial. Additionally, the emissions footprint of establishing or transporting additional process modules to a different manufacturing site can be incorporated to further elucidate the environmental cost of these resilience strategies. Finally, the effects of running processes at the boundary of their feasible regions on the reliability of the modules themselves need to be studied, as this may lead to an increased need for maintenance.

## ACKNOWLEDGEMENTS



The authors gratefully acknowledge support from Texas A&M University, USA, and the Texas A&M Energy Institute, USA.

## REFERENCES

1. Jasiūnas J, Lund PD, Mikkola J. Energy system resilience – a review. *Renewable Sustainable Energy Rev* 150: 111476 (2021)
2. Ridha E, Nolting L, Praktiknjo A. Complexity profiles: A large-scale review of energy system models in terms of complexity. *Energy Strategy Rev* 30: 100515 (2020)
3. Zavala-Alcivar A, Verdecho MJ, Alfaro-Saiz JJ. A conceptual framework to manage resilience and increase sustainability in the supply chain. *Sustainability* 12:6300 (2020)
4. Pasman H, Kottawar K, Jain P. Resilience of process plants: What, why, and how resilience can improve safety and sustainability. *Sustainability* 12: 6152 (2020)
5. Negri M, Cagno E, Colicchia C, Sarkis J. Integrating sustainability and resilience in the supply chain: A systematic literature review and a research agenda. *Bus Strategy Environ* 30:2858-2886 (2021)
6. Kamalahmadi M, Shekarian M, Parast MM. The impact of flexibility and redundancy on improving supply chain resilience to disruptions. *Int J Prod Res* 60: 1992-2020 (2022)
7. Crotogino F, Donadei S, Bünger Ulrich, Landinger H. Large-scale hydrogen underground storage for securing future energy supplies. In: 18<sup>th</sup> World Hydrogen Energy Conference. Ed: Stolten D, Grube T. Institute of Energy Research - Fuel Cells (2010)
8. Panteli M, Mancarella P, Trakas DN, Kyriakides E, Hatzigiorgioudis ND. Metrics and quantification of operational and infrastructure resilience in power systems. *IEEE Trans Power Syst* 32: 4732-4742 (2017)
9. Moreno-Sader K, Jain P, Ballestas Tenorio LC, Mannan MS, El-Halwagi MM. Integrated Approach of safety, sustainability, reliability, and resilience analysis via a return on investment metric. *ACS Sustain Chem Eng* 7: 19522-19536 (2019)
10. Hosseini-Motlagh SM, Samani MRG, Shahbazbegian V. Innovative strategy to design a mixed resilient-sustainable electricity supply chain network under uncertainty. *Appl Energy* 280: 115921 (2020)
11. Chrisandina NJ, Vedant S, Iakovou E, Pistikopoulos EN, El-Halwagi MM. Multi-scale integration for enhanced resilience of sustainable energy supply chains: Perspectives and challenges. *Comput Chem Eng* 164: 107891 (2022)
12. Demirhan CD, Tso WW, Powell JB, Heuberger CF, Pistikopoulos EN. A multiscale energy systems engineering approach for renewable power generation and storage optimization. *Ind Eng Chem Res* 59: 7706-7721 (2020)
13. Shao ZT, Cao XY, Zhai QZ, Guan XH. Risk-constrained planning of rural-area hydrogen-based microgrid considering multiscale and multi-energy storage systems. *Appl Energy* 334: 120682 (2023)
14. Lin YL. A multiscale modeling and optimization framework for the design of energy systems with embedded life cycle considerations. Master's thesis, Texas A&M University (2023)
15. Grossmann IE, Calfa BA, Garcia-Herreros P. Evolution of concepts and models for quantifying resiliency and flexibility of chemical processes. *Comput Chem Eng* 70: 22-34 (2014)
16. Chrisandina NJ, Vedant S, Iakovou E, Pistikopoulos EN, El-Halwagi MM. Resilience-aware multi-scale integration of distributed energy systems. *Comput Aided Chem Eng* 52: 2953-2958 (2023)
17. Chrisandina NJ, Deshpande AA, Al-Fadhli FM, Iakovou E, El-Halwagi MM, Pistikopoulos EN. Integration of risk, flexibility, and resilience in the optimization of water-energy nexus. Manuscript in review (2024)
18. Lopez-Molina A, Sengupta D, Shi C, Aldamigh E, Alandejani M, El-Halwagi MM. An integrated approach to the design of centralized and decentralized biorefineries with environmental, safety, and economic objectives. *Processes* 8: 1682 (2020)
19. Shehzad A, Bashir MJK, Sethupathi S. System analysis for synthesis gas (syngas) production in Pakistan from municipal solid waste gasification using a circulating fluidized bed gasifier. *Renewable Sustainable Energy Rev* 60: 1302-1311 (2016)
20. Lücking LE. Methanol Production from Syngas Process modelling and design utilising biomass gasification and integrating hydrogen supply. Master's thesis, Delft University of Technology (2017)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Design and Optimization of a Multipurpose Zero Liquid Discharge Desalination Plant

Dev Barochia<sup>a,b</sup>, Hasan Nikkhah<sup>a,b</sup>, and Burcu Beykal<sup>a,b\*</sup>

<sup>a</sup> Department of Chemical & Biomolecular Engineering, University of Connecticut, Storrs, CT, USA

<sup>b</sup> Center for Clean Energy Engineering, University of Connecticut, Storrs, CT, USA

\* Corresponding Author: [beykal@uconn.edu](mailto:beykal@uconn.edu).

---

## ABSTRACT

We study the design and optimization of a multicomponent seawater desalination process with zero liquid discharge (ZLD). The designed process is highly integrated with multiple sub processing units that include humidification-dehumidification, Lithium Bromide absorption chiller, multi-effect evaporators, mechanical vapor compression, and crystallization. Aspen Plus software with E-NRTL and SOLIDS thermodynamic packages are used for modeling and simulation of desalination and crystallization units, respectively. In addition to this, we use data-driven optimization to find the best operating condition (i.e., the temperature of the last effect evaporator) that minimizes the overall energy consumption of the designed plant with an output constraint imposed on the mass fraction of salts going to the ZLD system should be greater than 20 wt.% to achieve the ZLD goal. We use a local sample-based data-driven optimizer, Nonlinear Optimization with the Mesh Adaptive Direct Search (NOMAD) algorithm, to perform constrained simulation-based optimization. The results show that at the optimized temperature (71.58 °C), our design produces 1777 kg/hr drinking water with an energy consumption of 536 kW in comparison to 580 kW of energy consumption for the same plant output in the base case design (not optimized). Thus, data-driven optimization of the evaporator temperature improves the overall energy consumption by 7.5% and achieves higher desalination efficiency. Further, the integration of the crystallizer unit into the overall desalination process allows us to produce about 43 kg/h of NaCl and achieve ZLD.

---

**Keywords:** Zero Liquid Discharge, Desalination, Data-Driven Optimization, Aspen Plus, Multicomponent Analysis

## INTRODUCTION

The global desalination capacity has experienced significant growth in recent decades due to the rising demand for freshwater. Accordingly, the desalination processes have sustained their position as a dependable means of obtaining potable water. Different desalination technologies including reverse osmosis, multi-stage flash desalination, and multi-effect distillation (MED) are used to produce clean water from seawater [1-2]. However, there are some critical factors to consider when designing and implementing desalination technologies.

The primary area of concern is the environmental impact of desalination. Energy consumption and the challenges associated with brine treatment are the leading causes of environmental issues that surround this technology [3]. From the brine management perspective, the

aquatic ecosystem, natural hydrologic cycles, and public health are all negatively impacted by the large quantities of concentrated brine discharge produced by conventional desalination systems. On the contrary, zero liquid discharge (ZLD) systems aim to recover the vast majority of liquid wastewater, thereby mitigating the release of harmful pollutants and reducing the expenses associated with brine waste disposal [4]. However, ZLD as a brine treatment method is notably high in energy consumption. For example, a predominant approach for achieving ZLD is via mechanical vapor compression (MVC). This technology is utilized by compressing hot water vapor from units, such as evaporators, and recycling the heat from that steam throughout the desalination process with an energy consumption of approximately 20-25 kWh per m<sup>3</sup> of desalinated water produced [5]. Tahir and Al-Ghamdi [6] also proposed a ZLD system that consists of a MED

unit, a humidification-dehumidification (HDH) unit, and an evaporative crystallizer. According to their results, the highest performance ratio (PR) value of 12 is achieved when the brine temperature is 85°C and the temperature differential is 2°C. The HDH system further recovered potable water from the saline waste stream of the MED unit with a recovery ratio of 30% to 40%. A crystallizer unit is then used to produce salts derived from the highly concentrated brine effluent. Overall, their results showed that it was possible to attain a minimum specific energy usage of 720 kJ/kg.

From the energy consumption perspective, many desalination processes are fossil fuel-powered, which comes with a high carbon footprint [7]. There have been many efforts to reduce the energy consumption of seawater desalination by incorporating different technologies such as MVC or using a Lithium Bromide (LiBr) absorption chiller [8,9]. For example, Abdulrahim and Darwish [9] presented a novel arrangement of an absorption cycle that utilizes solar radiation as the driving energy to produce chilled water for air conditioning purposes and feed saturated steam to a thermal desalination process. The proposed process had a gain output ratio of 5.7 with a cooling capacity of 1500 kW for 32 kg/s of seawater feed flow.

Previously, we have also presented a new desalination plant design that integrates multi-effect evaporators (MEEs), a LiBr absorption chiller, an MVC system, and an HDH unit to decrease the energy consumption of seawater desalination [3]. The LiBr absorption chiller and the MVC systems were employed to reclaim thermal energy from the evaporators, while the humidifiers were utilized to intensify the concentration of the brine effluent to a level appropriate for ZLD processing. Our results showed that this highly integrated system exhibited superior performance in terms of gain-output ratio of 11 and recovery ratio of 81% when compared to a conventional HDH system. Additionally, our design offered a cooling capability of 6 kW, which may be utilized for air conditioning purposes. While this proposed process was successful with seawater desalination, it did have room for improvement through mathematical optimization and the addition of a crystallization unit. Especially, data-driven and simulation-based optimization strategies have proven to be a highly successful approach for identifying the optimal process conditions across various applications [10-13].

Motivated by this, in this study, we further expand our process by adding a flash separator, crystallizer, and solid separator units to produce NaCl salts and reduce the brine discharge to almost zero. We also transitioned from two-component analysis to multicomponent analysis, as initially the feedstock was only simulated with water and NaCl. Moreover, we use data-driven optimization to find the optimized operation conditions for the MEEs to reduce the energy consumption of the overall process.

To this end, we use the nonlinear optimization with mesh adaptive direct search algorithm (NOMAD) [14] and impose output constraints on the discharged NaCl concentration in the effluent stream to ensure ZLD.

## METHODOLOGY

In this work, desalination is carried out using seawater as the source water. Previously, we used a two-component system to define the seawater composition (3.5 wt.% NaCl with the remaining being water) [3]. Here, we use a multicomponent analysis for the seawater components with their respective composition values shown in Table 1.

**Table 1:** Seawater feed composition [15].

Component	Mass Fraction
Water	0.966346
NaCl	0.0261789
MgCl <sub>2</sub>	0.00226156
MgSO <sub>4</sub>	0.00323751
KCl	0.000706899
NaHCO <sub>3</sub>	0.000188049
CaCl <sub>2</sub>	0.00108062
LiCl	8.928e-07

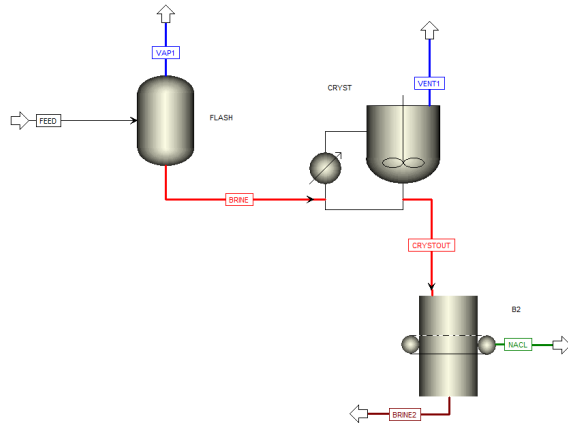
## Modeling & Simulation of the Base Case Desalination Units

The base case desalination process consists of the integration of four different technologies: HDH, LiBr absorption chiller, MEE, and MVC. Seawater with a flowrate of 1950 kg/hr first enters the LiBr absorption chiller cycle where it is preheated. Afterward, the preheated seawater stream enters three multi-effect evaporation units where water is evaporated to steam. The effluent steam from the last effect is then compressed in the MVC subprocess, where we compress the incoming vapor up to 1.5 bar. This is then used to activate the generator in the LiBr cycle. The steam exiting the generator is recycled to the evaporators and as energy is released, the steam condenses and is collected as desalinated water. The brine discharged from the MEE units goes to the humidifiers to be further concentrated. This process is modeled in Aspen Plus (V12.0) with the E-NRTL thermodynamic package using the default parameters. The details on model assumptions and simulation setup for this process can be found in [3].

## Path to Zero Liquid Discharge

To reach ZLD and minimize brine discharge, we add a flash separator, crystallizer, and solid filter units to the base case design (Figure 1). The concentrated brine

effluent from the humidifiers goes to the flash separator first, where water vapor is recovered as desalinated water, and the brine is further sent to the crystallizer unit to be precipitated and crystallized.



**Figure 1:** Crystallizer flowsheet with a flash separator, crystallizer unit, and solids separator.

**Table 2:** Solubility data of NaCl in water [15].

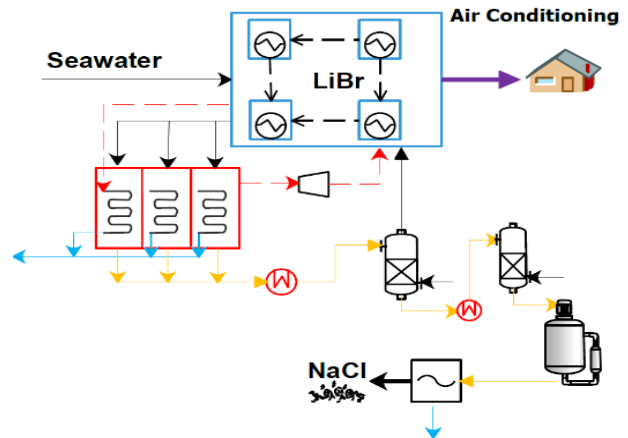
Temperature (°C)	Solubility (g/L)
0	355.2311
10	355.8293
20	357.1254
25	358.1224
30	359.2191
40	362.1104001
50	365.6
60	369.4882

**Table 3:** Operating parameters for the ZLD units.

Unit	Temperature (Celsius)	Pressure (Bar)
Crystallizer	25	0.3
Flash Separator	85	0.4
Evaporator (optimized)	71.58	0.5

To have accurate results for the crystallizer, we use experimentally determined solubility data of NaCl in water [15], as shown in Table 2. This solubility data is utilized by the SOLIDS thermodynamic package in Aspen Plus to determine how much NaCl salt will crystallize. Finally, the output stream of the crystallizer is processed in a solid separator and NaCl salts are recovered. The other salts present within the brine are treated as inert salts, where

we assume that their crystallization is negligible due to their relatively small concentration compared to NaCl. The solid separation efficiency is also assumed to be 100%. The key operating parameters used for simulating the ZLD units are provided in Table 3, with an overview of the process shown in Figure 2.



**Figure 2.** Process overview of the designed multipurpose desalination system.

### Data-Driven Optimization

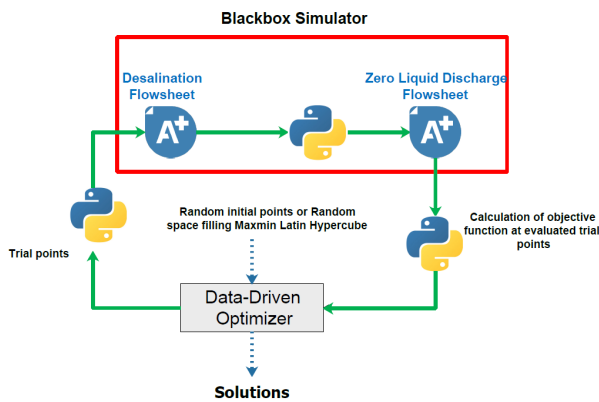
We used data-driven black-box optimization to optimize our operating conditions and minimize the energy consumption of the plant while maintaining ZLD. We formulate the following optimization problem:

$$\begin{aligned}
 & \min_x f(x): \text{Energy consumption} \\
 & \text{s.t.} \\
 & \quad g_i(x) \leq 0 \quad \forall i \\
 & \quad h_j(x) = 0 \quad \forall j \\
 & \quad x \in [x^L, x^U]
 \end{aligned} \tag{1}$$

where  $x$  is evaporator temperature for the last effect (*i.e.*, the decision variable) bounded between 60°C to 80°C,  $g(x)$  is the inequality constraint regarding the mass fraction of salt in the discharge brine which has to be greater than or equal to 21% wt. (*i.e.*,  $g(x) = 0.21 - \text{mass fraction of NaCl in brine} \leq 0$ ), and  $h(x)$  are the equality constraints pertaining mass and energy balances that are satisfied in Aspen Plus. We also performed sensitivity analysis on the base case simulation to identify the operating variables (*i.e.*, decision variables) with the greatest effect on energy consumption. We found that the evaporator temperature had the greatest impact on energy consumption. Here, we do not consider all evaporator temperatures as decision variables because the 1<sup>st</sup> and 2<sup>nd</sup> effect evaporator temperatures depend on the temperature of the last effect. Hence, only the temperature of the last effect evaporator is a true degree of freedom for optimization.

Our strategy for mathematically optimizing this process is considered a black-box optimization, as the

mechanistic expressions are accounted in Aspen Plus and not explicitly available to us. As a result, we collect input-output data from the simulator to identify the best operating condition for the evaporator temperature that minimizes the total energy consumption of the designed plant while achieving ZLD. To this end, we utilize the NOMAD and its progressive barrier constraint handling approach to solve Equation 1, as NOMAD is a local data-driven optimization algorithm. This algorithm is easy to implement and has shown good optimization performance for black-box problems with less than 50 variables. NOMAD can also handle nonlinear inequality constraints with a relatively low computational time [14].



**Figure 3.** Schematic of the connection between two Aspen Plus flowsheets and the flow of information between a generic data-driven optimizer and Aspen Plus.

To perform data-driven optimization, we first randomly initialize the algorithm within the lower and upper bounds of the decision variables. Following that, we solve the Aspen Plus simulation at this sampling point to collect the corresponding output values for the objective function and constraints. It is important to note that we execute Aspen Plus in two separate flowsheets due to having two different thermodynamic packages for the desalination and ZLD units. Once the first simulation (desalination) converges at a given sampling point, the data for the brine output in the first flowsheet is inputted as the feed conditions for the second flowsheet (ZLD). Final energy consumption from both processes is calculated by summing the energy consumption from individual flowsheets, and this serves as the output data along with the discharge brine concentration constraint. This data exchange is facilitated by a Python script (V.3.11) which relays the final input-output data to the NOMAD. After that, the algorithm generates new evaluation points to iteratively improve the incumbent solution, and the loop continues until a stopping criterion is met, such as reaching the maximum number of evaluations, or a tolerance for the mesh size [14]. Figure 3 illustrates how we performed the optimization over two flowsheets for a generic data-driven optimization process. We perform the process

simulation and the data-driven optimization on a Dell Precision Small Form Factor with 11th generation Intel Core i7- 11700, 4.9 GHz processor, and 16GB DDR4 installed memory running Microsoft Windows.

## RESULTS

The simulation results comparing the base case desalination and the new ZLD design are shown in Table 4. The key process parameters such as relative humidity and seawater temperature are set to be 80% and 20 °C, respectively. A full set of the base case parameters can be found in [3]. We observe that the optimization of the evaporator temperature reduced the desalination energy consumption by 7.5% (580 kW energy in the base case to 536 kW energy in the ZLD) for the same feed seawater and freshwater production flowrates. The energy consumption of the crystallizer flowsheet is 164 kW, which comes from the flash separator that is used to further concentrate the brine to a concentration that produces significant crystallization, salt production, and water vapor that can be recycled for heating. We observe that the water vapor is discharged at 217 kg/hr at 85 °C and 0.4 bar from the flash separator. This water vapor can also be condensed which will increase the freshwater production by 217 kg/hr which is another benefit of adding crystallization. The marginally higher total energy consumption from the two flowsheets is also an expected result for the new design, as we added three new units to achieve ZLD.

Moreover, the results show that with the addition of a crystallizer unit, we were able to produce approximately 43 kg/h of NaCl and nearly doubled the total dissolved solids (TDS) concentration of the discharged brine. Initially, according to the composition of seawater, there is about  $1950 \times 0.026 = 50.7 \text{ kg/h}$  of NaCl in our process. This means that our process achieves approximately 85% NaCl removal efficiency.

Furthermore, we show the compositions of the final brine to be discharged in Table 4, and it can be seen that the final brine is highly concentrated (TDS > 530,000 mg/L) with a mass flow rate of 43.88 kg/h. Compared to the process without ZLD units, we reduced the brine discharge from 302 to 43.88, approximately an 85% reduction. This would be valuable in reducing the environmental footprint of desalination. This brine reduction is due to the higher water recovery that we obtained by optimizing the process as well as the flash separator. Before adding the ZLD units to the base case simulation, the water recovery was about 81%. After adding the ZLD units, it further increased by 10 %. In our new design, we can recover 1777 kg/hr of freshwater from 1950 kg/hr of seawater



**Table 4.** Base case and ZLD results.

	Base Case	ZLD
Air temperature [°C]	30	30
Feed seawater Temperature [°C]	20	20
Seawater flowrate [ $\frac{m^3}{hr}$ ]	1.95	1.95
Relative humidity [%]	80	80
Humidifier pressure [bar]	0.5	0.5
Cooling Capacity (kW)	6	6
Crystallizer Salt Production [ $\frac{kg}{hr}$ ]	-	43.27
Freshwater Production [ $\frac{kg}{hr}$ ]	1560	1560+(217)
Energy Consumption of Desalination Flowsheet [kW]	580	536
Crystallizer Unit Energy Consumption [kW]	-	164
	<b>Base Case Brine</b>	<b>ZLD Brine (Filter Output)</b>
NaCl (%)	16.7	16.4
MgCl <sub>2</sub> (%)	2.75	18.92
MgSO <sub>4</sub> (%)	1.17	8.02
KCl (%)	0.75	5.1
NaHCO <sub>3</sub> (%)	0.05	0.2
CaCl <sub>2</sub> (%)	0.65	4.5
LiCl (%)	0.003	0.6
TDS (mg/L)	221,700	537,400
Mass flow (kg/hr)	302	43.88

fed. In this study, we also did a simple calculation of revenue that can be gained by adding the ZLD components to the desalination units. Our base case had freshwater production as a revenue stream; however, with the optimized ZLD case, the amount of freshwater produced increased due to using a flash separator in the crystallizer flowsheet. The ZLD discharge case also produces salt through crystallization, providing another significant revenue stream while desalinating the brine and reducing the environmental impact of the discharge. Table 5 shows a simple calculation of the revenue that we can gain from water and salt production. Our analysis assumes that the average selling price for NaCl salt is 0.2 \$/kg [3] and the water selling cost is \$0.12/kg. The integration of ZLD units will result in generation of 217 kg/h of additional water in compared to process without ZLD units. This will result in \$206,237 /yr revenue from selling water. We also assume that the plant will be operating for 330 days annually.

Our calculations show that adding the crystallization units will result in \$274,776/yr revenue compared to a process without crystallization units. However, a detailed technoeconomic analysis is required to understand the economic viability of the process.

**Table 5:** Revenue streams of the ZLD process.

Revenue source	Assumptions	Amount
Salt Production	0.20 $\frac{\$}{kg}$	68,539 $\frac{\$}{yr}$
Fresh Water Production	0.12 $\frac{\$}{kg}$	206,237 $\frac{\$}{yr}$
Total revenue		274,776 $\frac{\$}{yr}$

## CONCLUSIONS

In this work, we modeled and simulated a zero liquid discharge desalination process using a Lithium Bromide absorption chiller, multi-effect evaporators, and mechanical vapor compression with crystallization. This process enabled us to recover more than 90% of seawater while producing NaCl salts. The results showed that we can have more than 43 kg/hr NaCl out of 1950 kg/hr seawater with approximately 85% NaCl removal efficiency. Our preliminary revenue analysis showed that we can gain \$274,776/yr with fresh water and salt production serving as the main products. In addition to this, we performed data-driven optimization to minimize energy

consumption in the process. Although the optimization in our study only considered the temperature of the last evaporator effect, the results showed that we can reduce the energy consumption by 7.5%. In the future, we will explore the optimization of other operational parameters and characterize their effects on the overall energy consumption. Some additional parameters that we are considering for optimization include the flash separator temperature and pressure as well as the operating pressure of the evaporators. These parameters showed significant influence on energy consumption during our initial sensitivity analysis, but further work must be done to see if they are viable candidates for optimization parameters. Furthermore, we will perform a comprehensive techno-economic analysis and life cycle assessment to quantify the economic and environmental impacts of the process, respectively.

## ACKNOWLEDGEMENTS

This work is conducted with the computing resources provided by the University of Connecticut.

## REFERENCES

1. Nakoa K, Rahaoui K, Date A, Akbarzadeh A. Sustainable zero liquid discharge desalination (SZLDD). *Sol Energy* 135:337-347 (2016)
2. Chen Q, Akhtar FH, Burhan M, Kumja M, Ng KC. A novel zero-liquid discharge desalination system based on the humidification-dehumidification process: A preliminary study. *Water Res* 207:117794 (2021)
3. Nikkhah H, Beykal B. Process design and techno-economic analysis for zero liquid discharge desalination via LiBr absorption chiller integrated HDH-MEE-MVR system. *Desalination* 558:116643 (2023)
4. Pinna S, Bigham S. Multiple-effect desiccant-based zero liquid discharge desalination systems. *Desalination* 502:114942 (2021)
5. Finnerty C, Zhang L, Sedlak DL, Nelson KL, Mi B. Synthetic graphene oxide leaf for solar desalination with zero liquid discharge. *Environ Sci Technol* 51(20):11701-11709. (2017)
6. Tahir F and Al-Ghamdi SG. Integrated MED and HDH desalination systems for an energy-efficient zero liquid discharge (ZLD) system. *Energy Rep* 8:29-34 (2022)
7. Lawal DU, Antar MA, Khalifa AE. Integration of a MSF desalination system with a HDH system for brine recovery. *Sustainability* 13(6):3506 (2021)
8. Shaaban AM, Antar MA, Khalifa AE, El-Shaarawi MA. Analysis of Integrated H<sub>2</sub>O-LiBr Absorption Cooling and Single-Effect Evaporation Desalination System. *Arab J Sci Eng* 45:5273-5284 (2020)
9. Abdulrahim HK, Darwish MA. Thermal desalination and air conditioning using absorption cycle. *Desalin Water Treat* 55.12:3310-3329 (2015)
10. Beykal B, Boukouvala F, Floudas CA, Pistikopoulos EN. Optimal design of energy systems using constrained grey-box multi-objective optimization. *Comput Chem Eng* 116:488-502 (2018)
11. Beykal B, Boukouvala F, Floudas CA, Sorek N, Zalavadia H, Gildin E. Global optimization of grey-box computational systems using surrogate functions and application to highly constrained oil-field operations. *Comput Chem Eng* 114:99-110 (2018)
12. Nikkhah A, Nikkhah H, Shahbazi A, Zarin MKZ, Iz DB, Ebadi MT, Fakhroleslam M, Beykal B. Cumini and eucalyptus essential oil standardization using fractional distillation: Data-driven optimization and techno-economic analysis. *Food Bioprod Process* 143:90-101 (2024)
13. Beykal B, Avraamidou S, Pistikopoulos IPE, Onel M, Pistikopoulos EN. Domino: Data-driven optimization of bi-level mixed-integer nonlinear problems. *J Glob Optim* 78:1-36 (2020)
14. Le Digabel S. Algorithm 909: NOMAD: Non-linear Optimization with the MADS algorithm *ACM Transactions on Mathematical Software* 37(4):44:1-44:15 (2011)
15. Atherton S, Solubilities of Inorganic and Organic Compounds: A Compilation of Quantitative Solubility Data from the Periodical Literature, Vol-ume 1. D. Van Nostrand Company (1919).

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Opportunities for Process Intensification with Membranes to Promote Circular Economy Development for Critical Minerals

Molly Dougher<sup>a</sup>, Laurianne Lair<sup>a</sup>, Jonathan Aubuchon Ouimet<sup>a</sup>, William A. Phillip<sup>a</sup>, Thomas J. Tarka<sup>b</sup>, and Alexander W. Dowling<sup>a\*</sup>

<sup>a</sup> University of Notre Dame, Department of Chemical and Biomolecular Engineering, Notre Dame, IN 46556, United States

<sup>b</sup> National Energy Technology Laboratory, US Department of Energy, Pittsburgh, PA 15236, United States

\* Corresponding Author: [adowling@nd.edu](mailto:adowling@nd.edu).

## ABSTRACT

Critical minerals are essential to the future of clean energy, especially energy storage, electric vehicles, and advanced electronics. In this paper, we argue that process systems engineering (PSE) paradigms provide essential frameworks for enhancing the sustainability and efficiency of critical mineral processing pathways. As a concrete example, we review challenges and opportunities across material-to-infrastructure scales for process intensification (PI) with membranes. Within critical mineral processing, there is a need to reduce environmental impact, especially concerning chemical reagent usage. Feed concentrations and product demand variability require flexible, intensified processes. Further, unique feedstocks require unique processes (i.e., no one-size-fits-all recycling or refining system exists). Membrane materials span a vast design space that allows significant optimization. Therefore, there is a need to rapidly identify the best opportunities for membrane implementation, thus informing materials optimization with process and infrastructure scale performance targets. Finally, scale-up must be accelerated and de-risked across the materials-to-process levels to fully realize the opportunity presented by membranes, thereby fostering the development of a circular economy for critical minerals. Tackling these challenges requires integrating efforts across diverse disciplines. We advocate for a holistic molecular-to-systems perspective for fully realizing PI with membranes to address sustainability challenges in critical mineral processing. The opportunities for PI with membranes are excellent applications for emerging research in machine learning, data science, automation, and optimization.

**Keywords:** Renewable and Sustainable Energy, Supply Chain, Multiscale Modelling, Process Intensification, Membranes, Machine Learning

## INTRODUCTION

As of 2022, there are 50 critical minerals [1] essential to renewable energy, such as lithium and rare earth elements (REEs). Lithium is vital to the function of several different types of batteries. REEs, particularly dysprosium, neodymium, praseodymium, and terbium, have unique properties (e.g., magnetic, fluorescence), which make them essential components of many electronics and clean energy technologies (e.g., displays, hard disk drives, wind turbines, electric vehicles). [2] Further, the term “rare” denotes that REEs are economically difficult

to produce or concentrate [3]; although relatively abundant, REEs are generally present in very low concentrations.

Overall, there is a need to diversify the supply of critical minerals, as primary sources (e.g., minerals, clays, and brines) have been declining in quality (i.e., the concentration of critical minerals) [4,5] and the global energy transition (amongst other factors) has increased the demand for critical minerals. The combination of these issues drives a need to develop enhanced recovery processes, including for end-of-life materials, to obtain critical mineral supplies, improve sustainability, and progress

toward circularity goals. [6] These sustainability-driven efforts include moving attention towards alternative feedstocks such as recycled electronic waste and unconventional sources like coal fly ash. Additionally, the concentrations of critical minerals in some alternative and unconventional sources are beginning to surpass those in primary sources. [2,7,8] Regardless of the source, critical mineral processing generally relies on techniques that can use hazardous materials (e.g., acid leaching). [9] One pathway to enhance sustainability and improve the process performance of critical mineral systems may be to integrate membrane technology, which has a track record of success in other industries (e.g., desalination, water treatment, chemical processing, and biomanufacturing). [10]

This brief literature review focuses on opportunities for membrane technologies to optimize and intensify critical mineral processing systems. We emphasize the importance of a molecular-to-systems perspective with close collaborations between process systems engineering (PSE) and other fields (e.g., membrane, data, and computational sciences) to accelerate the development of a circular economy.

## CRITICAL MINERAL PROCESSING

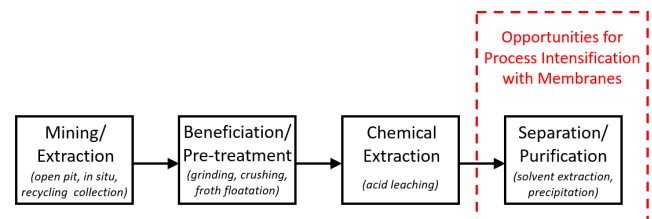
### Evolving Supply Chains

The continued development of domestic critical mineral infrastructure is vital to decreasing reliance on foreign supply chains. Taking the United States as an example, in 2022, the country produced only 14 critical minerals domestically, relying entirely on net imports for 12 critical minerals and more than 50% on net imports for another 31 critical minerals. [11] Developing domestic supply chains will safeguard against unforeseen insecurities in the global market. For example, the COVID-19 pandemic caused disruptions across all aspects of supply chains, where production of certain critical minerals (e.g., manganese) decreased, and consumption of others (e.g., platinum) increased. [11] Therefore, securing reliable supply chains of critical minerals will be essential to national security and the clean energy sector.

The growing dependence on technology and the decreasing quality of primary ore necessitates a more concerted shift towards harvesting critical minerals from secondary and unconventional sources, including recycled materials. In particular, integrating recycled waste into current infrastructure can promote a circular economy for critical minerals; therefore, developing new and existing processes is essential to achieving global climate goals. There has been some progress in recycling practices; for example, the domestic supply of 7 critical minerals in the United States consisted exclusively of recycled materials [11].

## Technology and Challenges

Critical mineral processing systems often contain the four steps presented in Figure 1: mining/extraction, beneficiation/pre-treatment, chemical extraction, and separation/purification. [12] The technology used within these steps depends on the feed stream. For example, mineral ore requires mining techniques (e.g., open-pit, underground, *in situ* extraction) to extract host rocks or the critical minerals [13]; brines frequently need to be pumped from underground [14]; and end-of-life consumer waste (e.g., cell phones) requires diversion from landfills via collecting the products to be recycled (i.e., urban mining). Beneficiation (or pre-treatment) encapsulates the techniques that concentrate the critical mineral feeds by removing unwanted parts (e.g., nonvaluable rock and casings on electronic devices) or physically pre-concentrate liquid sources. Typically, both chemical extraction and separation/purification employ hydrometallurgical (i.e., leaching, solvent extraction, and precipitation) or pyrometallurgical techniques (i.e., high temperatures) [7], which obtain critical minerals in usable quantities and purities.



**Figure 1.** Generalized block-flow diagram for critical mineral processing route.

Overall, there is motivation to revolutionize the processing structure of critical minerals to reduce the environmental impact, physical footprint, or cost of purification. Conventional processing can produce large waste streams, which can be detrimental to the environment, or may have large physical footprints that disturb habitats. For example, acid leaching (hydrometallurgy) dissolves solid critical minerals into a liquid phase, creating acidic waste that requires treatment before disposal. On the other hand, lithium recovery from salt-lake brine has an extensive physical footprint, using large evaporation lakes/ponds to precipitate contaminants out of the solution successively. [15] Separation processes such as solvent extraction use long sequences of settling tanks and separators and have slow process dynamics; thus, they cannot quickly adapt to changes such as new feed concentrations or product specifications. In many of these areas, the focus has shifted to designing processes that reduce the environmental footprint and can address variability in the feed concentrations. Some sustainable solutions address specific environmental concerns (e.g., an

**Table 1.** A breakdown of different critical mineral sources and 30 references that highlight various membrane technology implementations for each respective feedstock (lithium- or rare earth elements-based).

LIB: lithium-ion battery; SLM: supported liquid membrane; LLM: liquid-liquid membrane; LEM: liquid emulsion membrane; PIM: polymer inclusion membrane

\* Industrial waste, mine tailings, and coal leachate

+ Nanofiltration, ion exchange, electrodialysis, electrolysis, capacitive deionization, diffusion dialysis

Feed	Membrane Process	Charged Separations*	SLM, LLM, LEM, PIM	Adsorption	Membrane Distillation
Li	Ore		[51]		
	Brine	[16–18]	[20,52–58]	[22,31]	[24,60]
	LIB	[19]	[47,61]	[62]	
REEs	Ore		[63]	[23]	
	Magnet		[64]	[65–67]	
	Waste*		[68,69]	[21,70]	

electronic waste recycling process that limits the use of mineral acid leachates [9]); however, the processes are typically bespoke. Replacing certain hydrometallurgical techniques with mass separating agents (e.g., membrane systems) can offer environmental benefits for all critical mineral feed streams and facilitate process intensification (PI).

## CRITICAL MINERAL PROCESS INTENSIFICATION

### Membrane Opportunities

Membranes are uniquely equipped to aid in PI, as they can potentially improve efficiency, energy consumption, and cost in many chemical engineering systems. [10] Table 1 organizes a non-exhaustive set of references by the lithium- or REE-based feed stream and the corresponding applications of membrane technologies.

Recently, there has been significant work in developing membrane units for critical mineral separations. For example, comprehensive lithium recovery processes from salt-lake brine, primarily relying on nanofiltration (NF), have produced high-purity lithium products. [16–18] Kumar *et al.* [19] developed a process for recycling lithium-ion batteries (LIB) that leverages membranes to obtain battery-grade lithium carbonate. However promising, these processes are limited by their bench-scale design for singular feed streams. Membranes with charge-based separation mechanisms have been extensively applied to lithium extraction. For example, NF membranes have successfully separated lithium and magnesium ions in salt-lake brines by leveraging electrostatic interactions. [20] Similarly, liquid membranes have shown similar separation capabilities with REEs from coal fly ash leachate while utilizing less hazardous solvents than conventional solvent extraction. [21] Adsorption-based membrane techniques have demonstrated enhanced separations for both lithium and REEs. [22,23]

Additionally, membrane distillation has been implemented to achieve lithium recovery from brine four times faster than conventional evaporation ponds, with 20 times less surface area, and a corresponding reduction in physical footprint. [24]

PSE can guide future membrane studies like those presented in Table 1. Specifically, membrane cascades, which enable fractionation (i.e., utilize staged separations to isolate single-ion product streams), will be vital within critical mineral processing due to the complex feed streams typical of the industry. For example, batteries in electric vehicles can contain the following critical materials: lithium, cobalt, manganese, nickel, and graphite. [25] While the components are relatively well-defined, the specific material compositions will vary depending on the battery type. [26] Processes to recycle the components of diverse batteries must be able to separate critical minerals from one another for efficient downstream re-manufacture. The optimization of membrane cascades is one of the significant benefits that PSE offers within the critical mineral space.

### Multiscale and Interdisciplinary Challenges

The ability to rapidly develop, optimize, and scale up processes can accelerate circular supply chain development and the ability to meet clean energy (and, by extension, critical mineral production) goals. Multiscale optimization can be achieved with bidirectional (bottom-up and top-down) feedback to link technology breakthroughs at materials and device scales with process scale-up, supply chain optimization, and policy development at process and infrastructure scales. We now elaborate on these themes in five specific research opportunities.

### Machine Learning to Accelerate Material Design

Due to the number of different membrane materials (e.g., organic, inorganic) [27], materials should be chosen optimally for critical mineral separations and process



development. Membrane optimization is usually inefficient when pursued solely through Edisonian experimentation due to the complexity of the design space (i.e., the parameters needed to describe material properties and reaction conditions) and the physical phenomena (i.e., the parameters within the mathematical model). [27] As such, machine learning (ML's) capabilities with high-dimensional data make it well-suited to tackle membrane problems with many degrees of freedom. [27] Membrane development efforts present an opportunity to further improve and enhance the design, environmental impact, and performance of critical mineral and material production systems, with the performance of membrane materials being the limiting factor in new, more environmentally responsible processes. [28] Recently, there have been significant developments in ML techniques to push the functionality of membrane materials. For example, Bayesian optimization (BO) [29] can search the vast design space (i.e., monomer and fabrication conditions) for NF membranes. [30] The advanced capabilities of BO increased the efficiency of membrane design and allowed for the fabrication of membranes that exceeded the upper limit of the trade-off between water selectivity and permeability. [30]

Mathematical modeling and ML can connect material design with larger length scales to enhance technological success. [27,31] For example, Eugene *et al.* [31] explored adsorptive membranes for removing lithium from water, comparing the properties of several sorbents to identify feasibility ranges. Further, the framework enables analysis of membrane performance after scale-up, emphasizing the importance of characterizing materials within the desired operating range. As another example, Rall *et al.* [32] used artificial neural network surrogate models to optimize membranes at the material and process levels. Testing new materials within the desired system will become vital within critical mineral separations as the feed streams evolve. For instance, process development and optimization must occur as new critical mineral sources emerge in recycled products. Employing similar frameworks for newly identified feed streams can accelerate material testing, process development, and scale-up.

## Optimizing Membrane Modules

Computational fluid dynamics (CFD) helps optimize transport and flows in membrane module and device designs. [33] For example, Shirazi *et al.* [34] applied CFD to membrane distillation (MD), a thermally-driven membrane separation process, offering guidance on the physical design of the MD module and improving membrane performance. Recently, Choi *et al.* [35] used CFD with design of experiments (DoE) to identify variables affecting performance and determine the optimal MD module design. Similar optimization strategies using CFD

simulations have been applied to membrane bioreactors. [36,37] The success of CFD techniques in optimizing the physical membrane module design can extend to the critical mineral space to enhance overall process development further.

## Increasing Experimental Efficiency

PSE paradigms motivate the transition from Edisonian research to optimized and automated experiments.

We posit that membrane science is ripe to benefit from experiment automation and optimization advances. Automating experiments can accelerate material discovery by reducing the required time. For example, Muetzel *et al.* [38] automated a diafiltration experiment, reducing experimentation time by 40% and discovering concentration dependence of transport parameters. Similarly, Ouimet *et al.* [39] dosed a diafiltrate (of higher or lower concentration than the feed) into their automated diafiltration system to evaluate an expansive concentration space for water purification, leading to parameter identification five times faster than traditional filtration methods.

More robust efforts in enhanced experimental automation focus on self-driving laboratories (SDLs), which combine robotics and automated lab work with ML techniques to iterate experiments until the desired objective is reached. [40] SDLs improve computational efficiency (compared to the scientist performing the analysis) and increase the speed and precision of repetitive experimental techniques. [40] Further, SDLs can improve material discovery. For example, MacLeod *et al.* [41] recently developed a SDL for palladium film synthesis to quantify the trade-offs between temperature and conductivity. Their SDL required human action every 40-60 experiments, providing a robust dataset of information and overcoming human limitations. [41] Alternatively to SDLs using data-driven ML models, experimentation can be optimized using model-based design of experiments (MBoE). MBoE exploits the mathematical structure of science-based models to identify the experiments that will generate the most informative data efficiently. [42] We anticipate that SDLs driven by MBoE will be critical to navigating the complex (often competing) phenomena that govern efficient membrane separations.

## Process Design and Optimization

Membrane cascades typically have several feasible designs [43], and as the system complexity or size increases, it becomes unfavorable for the researcher to optimize each process individually. Superstructure optimization (SSO) is a popular strategy to assess all configurations using mathematical programming [44], subject to the desired objective (e.g., minimizing energy or cost, maximizing recovery or revenue) and constraints (e.g.,

transport phenomena, number of stages).

The success of SSO for membrane-based processes in other fields (e.g., carbon capture [43,45,46]) can inform future work in the critical mineral space. SSO has optimized membrane cascades within post-combustion carbon capture to identify more efficient designs. [43,45] SSO routinely accommodates multiple objectives (e.g., cost, emissions, water usage). [46] The generalizability of SSO also allows for increased flexibility (e.g., configuration size, separation type, individual module performance). For example, Lee *et al.* [45] found that varying the membrane properties in different stages (higher permeance upstream, higher selectivity downstream) improved process performance and efficiency. The generalized framework of SSO enables the application to diverse applications, allowing adaptation within critical mineral separations. Specifically, diafiltration cascades have used SSO for lithium-cobalt separation systems [47], efficiently optimizing staged separations and evaluating trade-offs in material property targets.

### Supply Chain Optimization

Finally, supply chain optimization can identify the best opportunities for PI to maximize the impact at a global scale by promoting resilient and circular supply chains for critical minerals. For example, Canales-Bustos *et al.* [48] considered economic and environmental factors to optimize a mining supply chain subject to decarbonization goals. Fattahi *et al.* [49] investigated uncertainty in the critical mineral supply chain through multi-stage stochastic programming. Similar studies performing optimization under uncertainty will be essential to developing resilient supply chains, as variability in feed concentration, desired product specifications, market prices, and consumer demand is inevitable. After integrating membranes within process and infrastructure models, PSE methods enable the communication of performance targets back to the materials and device scales. This integrated feedback loop is the key to rapidly optimizing enhanced critical mineral processes within a circular economy.

### CONCLUDING REMARKS

We propose a molecular-to-systems perspective of critical mineral processing and process systems engineering to enable process intensification with membranes. Machine learning techniques enable accelerated material discovery by evaluating material properties and process targets across materials, devices, and systems scales. Material discoveries lead to more robust device design, which can be further enhanced using automation and optimization techniques. Superstructure optimization of membrane cascades increases the generalizability and flexibility of process design and can evaluate

membrane material property targets and design specifications. Finally, supply chain optimization enables multi-stakeholder optimization by managing material and energy flows and determining facility siting. [50] The advancements within the device and materials scales inform process design and infrastructure development, and vice versa, emphasizing the need for a holistic approach involving the materials, membrane, data, and computational science communities.

### ACKNOWLEDGEMENTS

M.D. recognizes support from the Arthur J. Schmitt Presidential Leadership Fellowship. L.L. received support from a Graduate Assistance in Areas of National Need fellowship from the Department of Education via Grant P200A210048 and the Remick Graduate Fellowship in Engineering administered by the University of Notre Dame. This work was kindly supported by the National Science Foundation (NSF) through award 2147605. J.A.O. gratefully acknowledges support from the CEST/Bayer Predoctoral Fellowship. This effort was funded by the U.S. Department of Energy's Process Optimization and Modeling for Minerals Sustainability (PrOMMiS) Initiative, supported by the Office of Fossil Energy and Carbon Management's Office of Resource Sustainability.

Disclaimer: This project was funded by the Department of Energy, National Energy Technology Laboratory an agency of the United States Government, through a support contract. Neither the United States Government nor any agency thereof, nor any of its employees, nor the support contractor, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

### REFERENCES

1. USGS Communications and Publishing. <https://www.usgs.gov/news/national-news-release/us-geological-survey-releases-2022-list-critical-minerals>
2. Balaram V. Rare earth elements: A review of applications, occurrence, exploration, analysis, recycling, and environmental impact. *Geoscience*

- Frontiers* 10: 1285–1303 (2019)
3. Opere EO, Struhs E, Mirkouei A. A comparative state-of-technology review and future directions for rare earth element separation. *Renewable and Sustainable Energy Reviews* 143:110917 (2021)
  4. Mudd GM. The Sustainability of Mining in Australia: Key Production Trends and Their Environmental Implications for the Future. Monash University (2007)
  5. Garrett DE. Handbook of Lithium and Natural Calcium Chloride. Elsevier (2004)
  6. Babbitt CW, Althaf S, Cruz Rios F, Bilec MM, Graedel TE. The role of design in circular economy solutions for critical materials, *One Earth* 4:353–362 (2021)
  7. Liu C, Lin J, Cao H, Zhang Y, Sun Z. Recycling of spent lithium-ion batteries in view of lithium recovery: A critical review. *Journal of Cleaner Production* 228:801–813 (2019)
  8. Lair L, Ouimet JA, Dougher M, Boudouris BW, Dowling AW, Phillip WA. Critical Mineral Separations: Opportunities for Membrane Materials and Processes to Advance Sustainable Economies and Secure Supplies. *Annu Rev Chem Biomol Eng Under Review* (2024)
  9. Prodius D, Gandha K, Mudring AV, Nlebedim IC. Sustainable Urban Mining of Critical Elements from Magnet and Electronic Wastes. *ACS Sustain Chem Eng* 8:1455–1463 (2020)
  10. Drioli E, Stankiewicz AI, Macedonio F. Membrane engineering in process intensification-An overview. *Journal of Membrane Science* 380:1–8 (2011)
  11. USGS. Mineral Commodity Summaries 2023 (2023)
  12. Fritz AG, Tarka TJ, Mauter MS. Assessing the economic viability of unconventional rare earth element feedstocks. *Nat Sustain* 6:1103–1112 (2023)
  13. Haque N, Hughes A, Lim S, Vernon C. Rare earth elements: Overview of mining, mineralogy, uses, sustainability and environmental impact. *Resources* 3:614–635 (2014)
  14. Kaunda RB. Potential environmental impacts of lithium mining. *Journal of Energy and Natural Resources Law* 38:237–244 (2020)
  15. Swain B. Recovery and recycling of lithium: A review. *Separation and Purification Technology* 172:388–403 (2017)
  16. Zhao Y, Wang H, Li Y, Wang M, Xiang X. An integrated membrane process for preparation of lithium hydroxide from high Mg/Li ratio salt lake brine. *Desalination* 493:114620 (2020)
  17. Xu W, Liu D, He L, Zhao Z. A comprehensive membrane process for preparing lithium carbonate from high Mg/Li brine. *Membranes* 10:1–14 (2020)
  18. Grageda M, Gonzalez A, Quispe A, Ushak S. Analysis of a process for producing battery grade lithium hydroxide by membrane electrodialysis. *Membranes* 10:1–21 (2020)
  19. Kumar R, Liu C, Ha G, Park Y, Ali Khan M, Jang M, Kim S, Amin M, Gacem A, Jeon B. Downstream recovery of Li and value-added metals (Ni, Co, and Mn) from leach liquor of spent lithium-ion batteries using a membrane-integrated hybrid system. *Chemical Engineering Journal* 447:137507 (2022)
  20. Foo ZH, Rehman D, Bouma AT, Monsalvo S, Lienhard JH. Lithium Concentration from Salt-Lake Brine by Donnan-Enhanced Nanofiltration. *Environ Sci Technol* 57:6320–6330 (2023)
  21. Smith RC, Taggart RK, Hower JC, Wiesner MR, Hsu-Kim H. Selective Recovery of Rare Earth Elements from Coal Fly Ash Leachates Using Liquid Membrane Processes. *Environ Sci Technol* 53:4490–4499 (2019)
  22. Sun Y, Wang Y, Liu Y, Xiang X. Highly Efficient Lithium Extraction from Brine with a High Sodium Content by Adsorption-Coupled Electrochemical Technology. *ACS Sustain Chem Eng* 9:11022–11031 (2021)
  23. Zhang F, Ma K, Li Y, Ran Q, Yao C, Yang C, Yu H, Hu S, Peng S. Selective separation of thorium from rare earths and uranium in acidic solutions by phosphorodiamidate-functionalized silica. *Chemical Engineering Journal* 392:123717 (2020)
  24. Park S, Kim J, Moon S, Jung J, Wang H, Ali A, Quist-Jensen C, Macedonio F, Drioli E, Lee Y. Lithium recovery from artificial brine using energy-efficient membrane distillation and nanofiltration. *J Memb Sci* 598:117683 (2020)
  25. Tracy BS. <https://www.energy.senate.gov/hearings/2022/3/>
  26. Ciez RE, Whitacre JF. Examining different recycling processes for lithium-ion batteries. *Nat Sustain* 2:148–156 (2019)
  27. Yin H, Xu M, Luo Z, Bi X, Li J, Zhang S, Wang X. Machine learning for membrane design and discovery. *Green Energy and Environment* (2022)
  28. Talukder MJ, Alshami AS, Tayyebi A, Ismail N, Yu X. Membrane Science Meets Machine Learning: Future and Potential Use in Assisting Membrane Material Design and Fabrication. *Separation and Purification Reviews* 1–14 (2023)
  29. Wang K, Dowling AW. Bayesian optimization for chemical products and functional materials. *Curr Opin Chem Eng* 36:100728 (2022)
  30. Gao H, Zhong S, Zhang W, Igou T, Berger E, Reid E, Zhao Y, Lambeth D, Gan L, Afolabi MA, Tong Z. Revolutionizing Membrane Design Using Machine Learning-Bayesian Optimization. *Environ Sci Technol*, 56:2572–2581 (2022)
  31. Eugene EA, Phillip WA, Dowling AW. Material

- Property Targets to Enable Adsorptive Water Treatment and Resource Recovery Systems. *ACS ES&T Engineering* 1:1171–1182 (2021)
32. Rall D, Schweidtmann AM, Kruse M, Evdochenko E, Mitsos A, Wessling M. Multi-scale membrane process optimization with high-fidelity ion transport models through machine learning. *J Memb Sci* 608:118208 (2020)
  33. Ghidossi R, Veyret D, Moulin P. Computational fluid dynamics applied to membranes: State of the art and opportunities. *Chemical Engineering and Processing: Process Intensification* 45:437–454 (2006)
  34. Shirazi MMA, Kargari A, Ismail AF, Matsuura T. Computational Fluid Dynamic (CFD) opportunities applied to the membrane distillation process: State-of-the-art and perspectives. *Desalination* 377:73–90 (2016)
  35. Choi J, Cho H, Choi Y, Lee S. Combination of computational fluid dynamics and design of experiments to optimize modules for direct contact membrane distillation. *Desalination* 524:115460 (2022)
  36. Jin Y, Liu CL, Song XF, Yu JG. Computational fluid dynamics simulation as a tool for optimizing the hydrodynamic performance of membrane bioreactors. *RSC Adv* 9:32034–32046 (2019)
  37. Liu M, Yang M, Chen M, Yu D, Zheng J, Chang J, Wang X, Ji C, Wei Y. Numerical optimization of membrane module design and operation for a full-scale submerged MBR by computational fluid dynamics. *Bioresour Technol* 269:300–308 (2018)
  38. Muetzel ZW, Ouimet JA, Phillip WA. Device for the Acquisition of Dynamic Data Enables the Rapid Characterization of Polymer Membranes. *ACS Appl Polym Mater* 4:3438–3447 (2022)
  39. Ouimet JA, Liu X, Brown DJ, Eugene EA, Popps T, Muetzel ZW, Dowling AW, Phillip WA. DATA: Diafiltration apparatus for high-throughput analysis. *J Memb Sci* 641:119743 (2022)
  40. Abolhasani M, Kumacheva E. The rise of self-driving labs in chemical and materials sciences. *Nature Synthesis* 2:483–492 (2023)
  41. MacLeod BP, Parlane FG, Rupnow CC, Dettelbach KE, Elliott MS, Morrissey TD, Haley TH, Proskurin O, Rooney MB, Taherimakhsoosi N, Dvorak DJ. A self-driving laboratory advances the Pareto front for material properties. *Nat Commun* 13:995 (2022)
  42. Wang J, Dowling AW. Pyomo.DOE: An open-source package for model-based design of experiments in Python. *AIChE Journal* 68:e17813 (2022)
  43. Chavez Velasco JA, Tumbalam Gooty R, Tawarmalani M, Agrawal R. Optimal design of membrane cascades for gaseous and liquid mixtures via MINLP. *J Memb Sci* 636:119514 (2021)
  44. Mencarelli L, Chen Q, Pagot A, Grossmann IE. A review on superstructure optimization approaches in process system engineering. *Comput Chem Eng* 136:106808 (2020)
  45. Lee S, Binns M, Kim JK. Automated process design and optimization of membrane-based CO<sub>2</sub> capture for a coal-based power plant. *J Memb Sci* 563:820–834 (2018)
  46. Ramezani R, Randon A, Di Felice L, Gallucci F. Using a superstructure approach for techno-economic analysis of membrane processes. *Chemical Engineering Research and Design* 199:296–311 (2023)
  47. Wamble NP, Eugene EA, Phillip WA, Dowling AW. Optimal Diafiltration Membrane Cascades Enable Green Recycling of Spent Lithium-Ion Batteries. *ACS Sustain Chem Eng* 10:12207–12225 (2022)
  48. Canales-Bustos L, Santibañez-González E, Candia-Véjar A. A multi-objective optimization model for the design of an effective decarbonized supply chain in mining. *Int J Prod Econ* 193:449–464 (2017)
  49. Fattahi M. Resilient procurement planning for supply chains: A case study for sourcing a critical mineral material. *Resources Policy* 74:101093 (2021)
  50. Dowling AW, Ruiz-Mercado G, Zavala VM. A framework for multi-stakeholder decision-making and conflict resolution. *Comput Chem Eng* 90:136–150 (2016)
  51. Meng X, Long Y, Tian Y, Li W, Liu T, Huo S. Electro-membrane extraction of lithium with D2EHPA/TBP compound extractant. *Hydrometallurgy* 202:105615 (2021)
  52. Ying J, Luo M, Jin Y, Yu J. Selective separation of lithium from high Mg/Li ratio brine using single-stage and multi-stage selective electro-dialysis processes. *Desalination* 492:114621 (2020)
  53. Zhao Z, Liu G, Jia H, He L. Sandwiched liquid-membrane electro-dialysis: Lithium selective recovery from salt lake brines with high Mg/Li ratio. *J Memb Sci* 596:117685 (2020)
  54. Liu G, Zhao Z, He L. Highly selective lithium recovery from high Mg/Li ratio brines. *Desalination* 474:114185 (2020)
  55. Qiu Y, Yao L, Tang C, Zhao Y, Zhu J, Shen J. Integration of selectrodialysis and selectrodialysis with bipolar membrane to salt lake treatment for the production of lithium hydroxide. *Desalination* 465:1–12 (2019)
  56. Torres WR, Díaz Nieto CH, PrévotEAU A, Rabaey K, Flexer V. Lithium carbonate recovery from brines using membrane electrolysis. *J Memb Sci* 615:118416 (2020)
  57. Shi W, Liu X, Ye C, Cao X, Gao C, Shen J. Efficient

- lithium extraction by membrane capacitive deionization incorporated with monovalent selective cation exchange membrane. *Sep Purif Technol* 210:885–890 (2019)
58. Lee DH, Ryu T, Shin J, Ryu JC, Chung KS, Kim YH. Selective lithium recovery from aqueous solution using a modified membrane capacitive deionization system. *Hydrometallurgy* 173:283–288 (2017)
  59. Xing L, Song J, Li Z, Liu J, Huang T, Dou P, Chen Y, Li XM, He T. Solvent stable nanoporous poly (ethylene-co-vinyl alcohol) barrier membranes for liquid-liquid extraction of lithium from a salt lake brine. *J Memb Sci* 520:596–606 (2016)
  60. Cerda A, Quilaqueo M, Barros L, Seriche G, Gim-Krumm M, Santoro S, Avci AH, Romero J, Curcio E, Estay H. Recovering water from lithium-rich brines by a fractionation process based on membrane distillation-crystallization. *Journal of Water Process Engineering* 41:102063 (2021)
  61. Wang B, Liu F, Zhang F, Tan M, Jiang H, Liu Y, Zhang Y. Efficient separation and recovery of cobalt(II) and lithium(I) from spent lithium ion batteries (LIBs) by polymer inclusion membrane electrodialysis (PIMED). *Chemical Engineering Journal* 430:132924 (2022)
  62. Bhave RR, Kim D, Peterson ES. Membrane Assisted Solvent Extraction for Rare Earth Element Recovery. US Patent 9,968,887. (2018)
  63. Alemrajabi M, Ricknell J, Samak S, Rodriguez Varela R, Martinez J, Hedman F, Forsberg K, Rasmuson ÅC. Separation of Rare-Earth Elements Using Supported Liquid Membrane Extraction in Pilot Scale. *Ind Eng Chem Res* 61:18475-18491 (2022)
  64. Hammache Z, Bensaadi S, Berbar Y, Audebrand N, Szymczyk A, Amara M. Recovery of rare earth elements from electronic waste by diffusion dialysis. *Sep Purif Technol* 254: 117641 (2021)
  65. Ni'am AC, Wang YF, Chen SW, Chang GM, You SJ. Simultaneous recovery of rare earth elements from waste permanent magnets (WPMs) leach liquor by solvent extraction and hollow fiber supported liquid membrane. *Chemical Engineering and Processing - Process Intensification* 148:107831 (2020)
  66. Kim D, Powell LE, Delmau LH, Peterson ES, Herchenroeder J, Bhave RR. Selective Extraction of Rare Earth Elements from Permanent Magnet Scraps with Membrane Solvent Extraction. *Environ Sci Technol* 49:9452–9459 (2015)
  67. Deshmane VG, Islam SZ, Bhave RR. Selective Recovery of Rare Earth Elements from a Wide Range of E-Waste and Process Scalability of Membrane Solvent Extraction. *Environ Sci Technol* 54:550-558 (2019)
  68. López J, Reig M, Gibert O, Cortina JL. Recovery of sulphuric acid and added value metals (Zn, Cu and rare earths) from acidic mine waters using nanofiltration membranes. *Sep Purif Technol* 212:180–190 (2019)
  69. López J, Reig M, Gibert O, Cortina JL. Integration of nanofiltration membranes in recovery options of rare earth elements from acidic mine waters. *J Clean Prod* 210:1249–1260 (2019)
  70. Middleton A, Hsu-Kim H. Separation of Rare-Earth Elements by Supported Liquid Membranes: Impacts of Soluble Iron, Aluminum, and pH in Low-Grade Feedstocks. *ACS ES and T Engineering* 3:1197–1204 (2023)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Constraint Formulations for Bayesian Optimization of Process Simulations: General Approach and Application to Post-Combustion Carbon Capture

Clinton M. Duewall<sup>ab\*</sup>, Mahmoud M. El-Halwagi<sup>a</sup>

<sup>a</sup> Texas A&M University, Artie McFerrin Department of Chemical Engineering, College Station, TX, USA

<sup>b</sup> Bryan Research & Engineering, LLC, Bryan, TX, USA

\* Corresponding Author: [clint.duewall@bre.com](mailto:clint.duewall@bre.com).

---

## ABSTRACT

Some of the most highly trusted and ubiquitous process simulators have solution methods that are incompatible with algorithms designed for equation-oriented optimization. The natively unconstrained Efficient Global Optimization (EGO) algorithm approximates a black-box simulation with kriging surrogate models to convert the simulation results into a reduced-order model more suitable for optimization. This work evaluates several established constraint-handling approaches for EGO to compare their accuracy, computational efficiency, and reliability using an example simulation of an amine post-combustion carbon capture process. While each approach returned a feasible operating point in the number of iterations provided, none of them effectively converged to a solution, exploring the search space without effectively exploiting promising regions. Using the product of expected improvement and probability of feasibility as next point selection criteria resulted in the best solution value and reliability. Constraining probability of feasibility while solving for the next sample point was the least likely to solve, but the solutions found were most likely to be feasible operating points.

---

**Keywords:** Derivative Free Optimization, Surrogate Modeling, Process Simulation, Global optimization, Carbon Capture.

## INTRODUCTION

Process simulators are trusted sources of property prediction and process design calculations [1]. Some process simulations have characteristics that challenge most optimization algorithms. Black-box simulators have no exploitable functional form, computationally expensive executions, and approximate derivatives. Coupled with highly constrained optimization formulations, these characteristics make most algorithms inefficient when applied to process simulations.

Many different optimization techniques have been applied to process simulators. Gradient-based optimizers have been applied using finite difference approximation for derivatives. Derivative-free optimization techniques can be applied, but many rely prohibitively large numbers of objective function evaluations without extensive parallelization [2]. One approach to mitigate the challenges of expensive black-box optimization is to approximate

the objective and constraint functions with surrogate models. In theory, these types of optimization algorithms should improve in accuracy each time the objective and constraint functions are sampled, learning characteristics of the problem each iteration [3].

Several types of surrogate models have been used to approximate process simulators. Automatic surrogate model generation has been achieved through several routes. ALAMO uses subset selection for constrained regression to generate algebraic models suitable for optimization [4–6]. Others have used parametric [7] and non-parametric regression methods like neural networks [8], radial basis functions [9], and kriging models [10–12] to approximate process simulations.

The Efficient Global Optimization (EGO) algorithm [13] is a prevalent example of surrogate-based optimization. EGO uses a kriging surrogate model in an inner loop optimization to pick the next point to sample. While it has been extensively used in aerospace engineering design

optimization, EGO has rarely been applied to process engineering systems.

Contemporary research studies probability of feasibility, mean kriging predictions, and lower confidence bounds to constrain EGO's inner loop [14, 15]. This work applies EGO with selected constraint approaches to a post-combustion carbon capture process modeled in ProMax®. The objective is to minimize process energy intensity subject to process and reliability constraints.

The results compare the merits of each constraint approach according to constraint faithfulness, final objective value, and solution reliability. Conclusions are drawn to recommend which methods are most effective for the post-combustion carbon capture process.

## CONSTRAINED EGO OVERVIEW

EGO was first published by Jones, Schonlau, and Welch in 1998 [13]. They developed the algorithm to globally optimize expensive objective functions. The overall structure of the algorithm is as follows. Let  $F(x)$  be function of independent variables  $x$  to be minimized that returns a value  $y$ .  $F(x)$  can be sampled at different points  $X = \{x_1, x_2, \dots, x_n\}$  yielding  $Y = \{y_1, y_2, \dots, y_n\}$ . This initial dataset is selected using methods called design of experiments to fill the search space.  $X$  and  $Y$  are used to generate a kriging model  $f$  with mean and variance functions. Equation 1 shows the structure of the kriging model used in this work where  $\beta_i$  and  $\theta_i$  are sets of hyperparameters tuned during the model building process.

$$\hat{\mu}(x) = \sum_{i=1}^n \beta_i y_i + \prod_{i=1}^n \exp(-\theta_i |x_t^{(i)} - x_t^{(j)}|) \quad (1)$$

EGO uses this predicted mean and a predicted variance within selection criteria to determine the next best point to evaluate using the actual objective function. EGO has traditionally maximized a property called expected improvement  $EI$  to select the next point to evaluate. The expected improvement function is written in equation 2, where  $\sigma(x)$  is the prediction standard deviation, and  $\Phi$  and  $\phi$  are the cumulative and probability density functions, respectively.

$$EI(x) = (f_{\min} - \hat{\mu}(x)) \Phi(a(x)) + \sigma(x) \phi(a(x)) \quad (2)$$

$$a(x) = \frac{f_{\min} - \hat{\mu}(x)}{\sigma(x)}$$

Once EGO determines the point that maximizes the expected improvement function, it samples the objective function with the point. The new sampled point is then added to the end of the dataset. As shown in Figure 1, the next iteration then begins by refitting the kriging model with the new dataset.

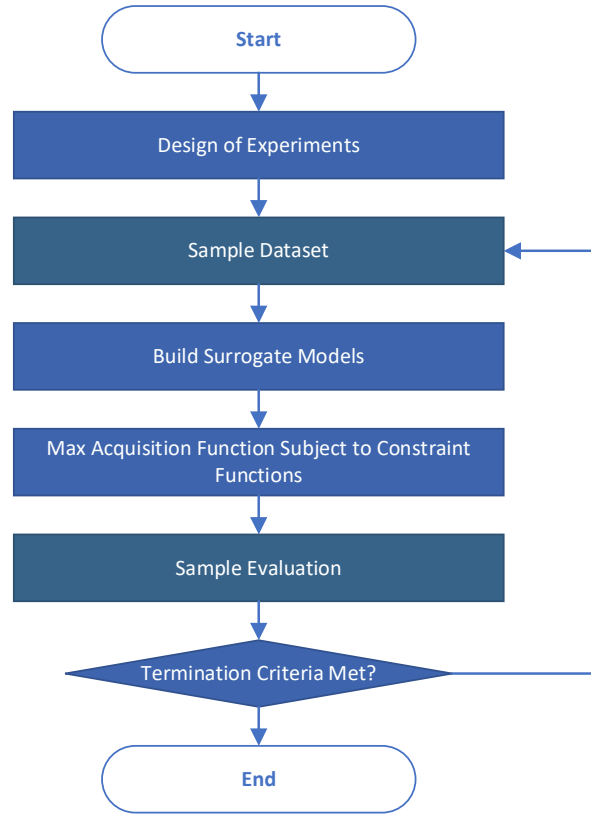


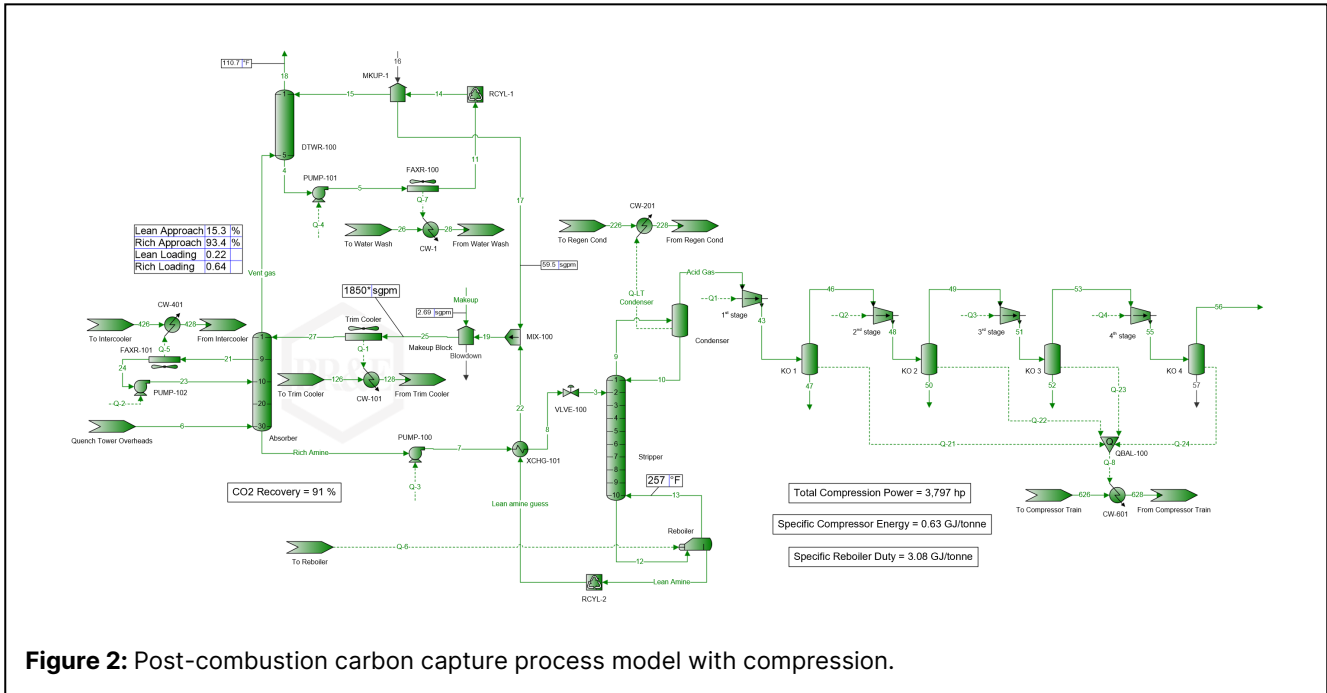
Figure 1. General EGO Algorithm Flowchart

Several constraint-handling methods have been developed for EGO. Modifying the next point selection criteria, the three approaches tested in this work rely on different properties of kriging models to enforce constraint boundaries. Two of these methods rely on a calculated probability of feasibility and were proposed by Sohst et al [14], who found that each method had its merits solving both mathematical test problems and aerodynamic shape optimization. Probability of feasibility can be conceptually understood as the probability that each constraint will be satisfied for a point based on the uncertainty in the kriging model for each constraint function. The formulation is shown in equation 3 where the constraint function is written in the form  $g(x) \leq 0$ . Overall probability of feasibility is just the product of each individual constraint's probability of feasibility.

$$PF(x) = \Phi\left(-\frac{\hat{g}(x)}{\sigma(x)}\right) \quad (3)$$

Probability of feasibility can be used in a few different ways to augment expected improvement within EGO's next point selection criteria. In one method, the product of expected improvement and probability of feasibility is maximized to select the next sample point. This method will be referred to as OFPF (Objective Function x Probability of Feasibility) in this work and is shown in equation 4.

$$\max OFPF(x) \quad (4)$$



**Figure 2:** Post-combustion carbon capture process model with compression.

$$OFPF(x) = EI(x) \cdot PF(x)$$

Another method tested in this work applies a probability of feasibility constraint to the maximization of expected in the next point selection suboptimization problem. This method will be referred to as PFCON (Probability of Feasibility Constraint).

$$\begin{aligned} \max EI(x) \\ \text{s. t. } PF(x) \geq 0.5 \end{aligned} \tag{5}$$

Constraints can also be enforced using a concept called the upper trust bound (UTB) of the kriging value. Given a desired confidence interval, the kriging model can calculate an upper bound on the constraint value by adding a factor of uncertainty. This constraint on the next point selection criteria is shown in equation 7 below where the constraint must be satisfied within three standard deviations of the kriging mean.

$$\begin{aligned} \max EI(x) \\ \text{s. t. } \hat{g}(x) \leq 3\sigma(x) \end{aligned} \tag{6}$$

If EGO selects a point where the objective function or constraint functions are undefined, a sufficiently suboptimal point approximating an infinite objective value can be used [16]. This approach should sufficiently penalize points in regions of the search space that cannot be evaluated.

## CARBON CAPTURE PROCESS MODEL

Gas processors have used amines to recover CO<sub>2</sub> and H<sub>2</sub>S from natural gas for nearly a century. After cooling the flue gas with a quench tower, the process of post-

combustion carbon capture with amines contacts an aqueous amine solvent with a CO<sub>2</sub>-rich stream to chemically absorb CO<sub>2</sub> with the amine. The CO<sub>2</sub>-rich amine is then heated in a still to regenerate the amine. Figure 2 provides a process flow diagram of the example unit used in the case study.

The main difference between post-combustion carbon capture and natural gas treating lies in the partial pressure of CO<sub>2</sub> in the feed gas. The concentration of CO<sub>2</sub> in the flue gas from a natural gas power plant is about 4 mol%. While this is easily within the range of concentrations found in natural gas, the pressure of the flue gas is near atmospheric pressure. The partial pressure of CO<sub>2</sub> in the flue gas will be 10-100 times less than in typical natural gas amine sweetening units, leading to the differences in the design of the two processes.

Flue gas carbon capture units are marked by larger absorber columns, due to the large volume of gas, with extensive water wash sections to reduce amine losses to the atmosphere. The amine solvent lost is not only expensive but also an air pollutant. Some designs also employ a side cooler, a feature not usually afforded for natural gas treating. Cooling the amine in the absorber removes heat of reaction and shifts equilibrium towards absorption in the liquid.

The example facility in this case study has a 30 ft diameter absorber packed with 60 ft of Sulzer IMTP #50 Metal random packing solved with ProMax’s Mass + Heat Transfer calculation method. The regenerator is a 9 ft diameter column modeled with 10 ideal stages of separation. The amine circulated is the CESAR-1 blend of 27 wt% amino methyl propanol (AMP) and 13 wt% piperazine.

**Table 1:** Inlet flue gas properties and composition

<b>Feed Property</b>	
Temperature	117 °F
Pressure	2.5 psig
Standard Vapor Volumetric Flow	406 MMSCFD
<b>Feed Composition</b>	
	mol %
Carbon Dioxide	3.99
Nitrogen	75.03
Oxygen	11.97
Water	9.00

Operators control the plant by manipulating the amine circulation rate, water wash rate, and reboiler duty. Operators are economically incentivized to minimize energy use while achieving a CO<sub>2</sub> recovery greater than 90% and amine losses below a permitted threshold. Utilities are integrated with the power generation facility. The amine reboiler can be powered by a slipstream of low-pressure steam from the heat recovery network of the powerplant. When integrated into the low-pressure steam circuit, the steam dedicated to the reboiler single-handedly reduces net power generation from 100 MW to 87 MW for the base-case simulation. Pumps and compressors can be powered by electricity generated on-site.

**Table 2:** Base carbon capture model power use summary.

<b>Unit Operation</b>	<b>Power</b>
Gas turbine	68 MW
Steam turbines (no CC)	32.7 MW
<b>Total power production</b>	<b>100.7 MW</b>
Reboiler steam power reduction	-12.8 MW
Flue gas blower	-5.1 MW
Amine and process water pumps	-0.1 MW
CO <sub>2</sub> Compression Power	-2.8 MW
Cooling water pumps	-0.6 MW
<b>Total power usage</b>	<b>-21.4 MW</b>
<b>Net power production</b>	<b>79.3 MW</b>

The desired CO<sub>2</sub> recovery depends on several factors including but not limited to local laws, plant scale, CO<sub>2</sub> sequestration or utilization strategy, and investor pressures. For plants located in the United States, the IRS's Section 45Q tax credit for carbon sequestration credits a dollar per ton amount based on the ultimate CO<sub>2</sub> use for eligible facilities [17]. In 2023 for new non-direct air capture processes, 45Q credits up to is \$85/tonne for geologically sequestered CO<sub>2</sub> and \$60/tonne for other qualified uses of CO<sub>2</sub> for the first 12 years of operation. Eligible power plants must capture at least 18,750 tonnes of CO<sub>2</sub> and be designed to capture at least 75% of the plant's emissions. In practice, many of these facilities are designed to capture at least 90% of baseline CO<sub>2</sub> emissions.

Once CO<sub>2</sub> is captured, it will require compression for transportation and its ultimate use. CO<sub>2</sub> transmission lines are designed to operate over 1000 psig in the so called supercritical "dense phase" where the fluid can be effectively pumped like a liquid. In this case, the captured CO<sub>2</sub> leaves the facility at 1500 psig.

## OPTIMIZATION FORMULATION

The implementation of constrained EGO uses the Surrogate Modeling Toolbox [18] package in Python. SMT includes an unconstrained implementation of EGO that was augmented to generate kriging models of each constraint function and use PFCON, UTB, and OFPF as next point selection criteria. The formulations used for next point selection criteria are described in equations (4) – (6) previously. These formulations are solved using a python SLSQP implementation [19] from 20 different starting points.

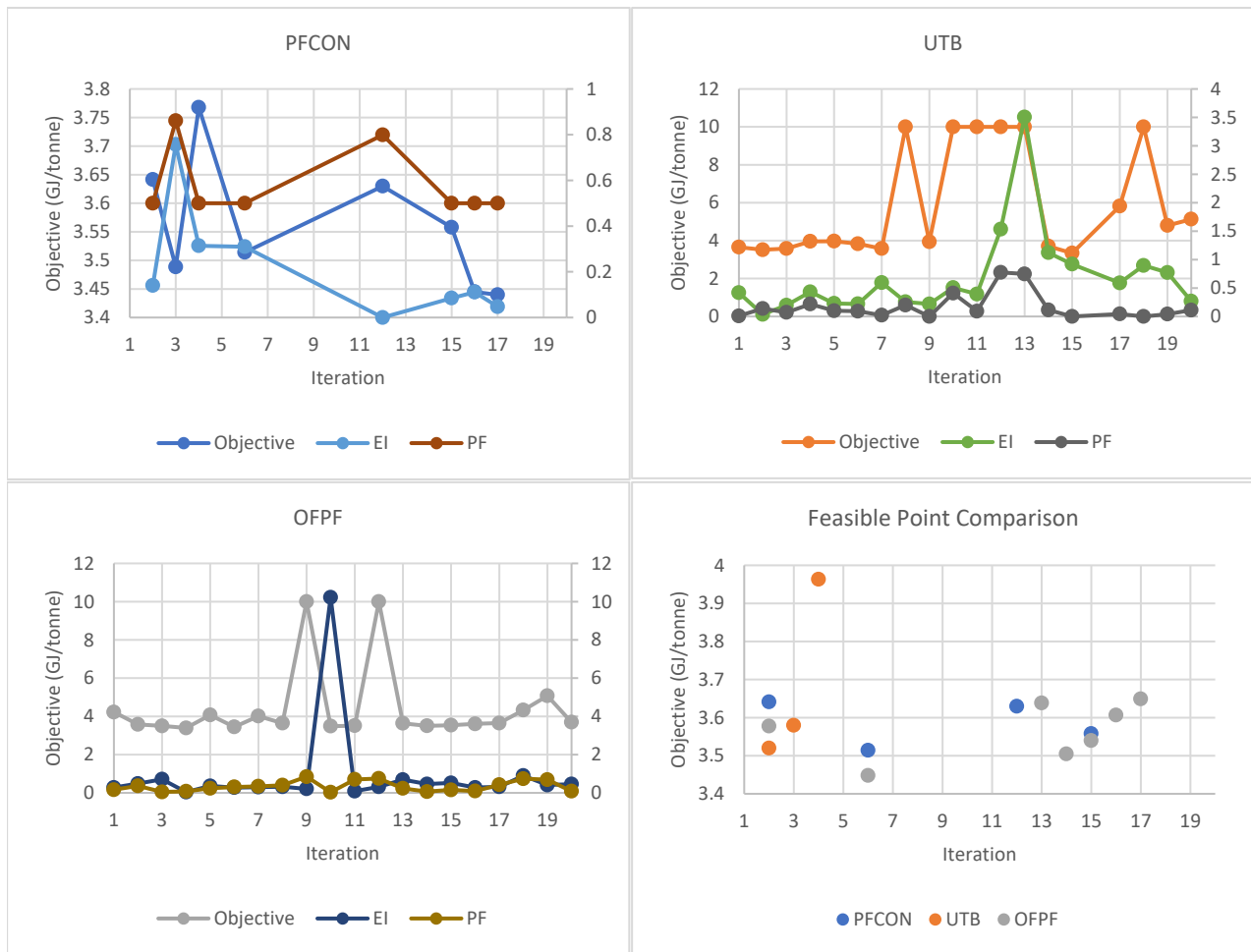
Any simulation run that did not converge in ProMax was given an objective value of 10 GJ/tonne, which was far greater than any observed objective value. Constraints for these unconverged samples were given values of 10 in standard form ( $g(x) \leq 0$ ) to denote a violation of each constraint.

The first task for developing an optimization formulation for a process is determining the most appropriate objective function and the most necessary constraints. For post-combustion carbon capture on power plant flue gas, one useful objective function is to minimize the energy used for carbon capture. Likely all this energy will be supplied by the power plant, so minimizing energy consumption will minimize the impact of carbon capture on plant profitability.

The overwhelming power requirement in the carbon capture process is the heat used to drive the reboiler. Of the power requirements shown in Table 2, the only significant costs that could be affected by the carbon capture unit itself are the reboiler duty and compression power. To avoid penalizing higher CO<sub>2</sub> recovery inherently, the objective function  $\hat{E}(x)$  will be the sum of the energy consumed by both the CO<sub>2</sub> compressors and the reboiler duty per unit of recovered CO<sub>2</sub> in units of GJ/tonne CO<sub>2</sub>.

Constraints are selected to ensure that equipment and operating constraints are met. The facility will be required to exceed 90% CO<sub>2</sub> recovery  $C(x)$ , emit less than 300 tonne/yr of amine solvent  $S(x)$ , and not exceed a calculated 95% fraction flooding in both the absorber  $F_a(x)$  and stripper  $F_s(x)$ .

Decision variables are limited to properties under operational control, leaving amine circulation  $a$ , water wash circulation  $w$ , and reboiler duty as candidates. In this study, the lean loading  $l$ , defined as the molar ratio of CO<sub>2</sub> and amine in the amine entering the absorber, will



**Figure 3:** Comparison of objective function, expected improvement, probability of feasibility, and feasible solution progress by iteration. Missing iterations failed to yield a sample point during sample point selection.

be used as a decision variable since it lends itself better to linear bounds than reboiler duty and should be more independent of other decision variables.

$$\begin{aligned}
 & \min \hat{E}(x) \text{ where } x = (a, w, l) \\
 & \text{subject to } C(x) \geq 90\% \\
 & \quad S(x) \leq 300 \text{ tonne/yr} \\
 & \quad F_a(x), F_s(x) \leq 95\% \quad (8) \\
 & \quad 1500 \leq a \leq 3000 \\
 & \quad 3000 \leq w \leq 5000 \\
 & \quad 0.05 \leq l \leq 0.2
 \end{aligned}$$

Each constraint function is passed to the algorithm in the form  $g(x) \leq 0$ . A Latin hypercube design of experiments is used to sample 30 points. This initial dataset becomes training points for the first kriging model. Next, each constraint-handling method is allowed 20 objective function evaluations to evaluate the search space and find an optimal solution.

One familiar with amine units would expect lower amine circulation rates and higher lean loadings to result in lower objective values. While the heat of absorption per ton of  $\text{CO}_2$  is effectively constant in this system, heating more solvent to the temperature required to desorb  $\text{CO}_2$  requires extra heat. One would also expect the water wash circulation should converge to the minimum required to meet the amine loss constraint.

## RESULTS AND DISCUSSION

The performance of each constraint method is evaluated using three criteria: constraint faithfulness, solution reliability, and best optimal solution. For this case study, each constraint method provides a feasible solution, but none of the methods reliably converged in the allotted number of iterations.

Objective function values for each iteration are shown in Figure 3. Optimal solutions for each run, plus the best solution observed for the problem in the study are stated in Table 3. This best solution was found by



chance in the initial design of experiments. The most optimal solutions had an intermediate value of lean loading between 0.12 – 0.15 mol CO<sub>2</sub>/mol amine which allowed lower circulation rates to meet the 90% CO<sub>2</sub> recovery constraint.

**Table 3:** EGO optimal solutions by constraint method.

	PFCON	UTB	OFFP	Best
Optimal Solution (GJ/tonne)	3.56	3.52	3.50	3.418
Amine Flow (sgpm)	1722	1626	1586	1575
Water Wash Flow (sgpm)	3976	4596	4421	4367
Lean Loading (mol/mol)	0.189	0.139	0.120	0.146

Despite not converging on a consistent operating point, the solutions provided by each method favor the expected minimal amine circulation rate. What was not expected was the extent to which lower lean loadings were preferred. Contacting flue gas with leaner than the maximum required lean loading to achieve 90% CO<sub>2</sub> recovery increased the capacity of each gallon of amine in circulation enough to offset the increased regeneration energy per gallon of amine. This resulted in lower energy requirements per metric ton of CO<sub>2</sub> captured.

Further analysis on the objective function values showed that specific compression power was effectively constant at about 0.63 GJ/tonne CO<sub>2</sub> captured. This indicates that compression power per unit of CO<sub>2</sub> captured was unaffected by regenerator operation while the overhead pressure was held constant.

**Table 4:** EGO run characteristics by constraint method.

	PFCON	UTB	OFFP
Failed Iterations	12	1	0
Feasible Points	4	4	8
Infeasible Points	4	15	12
Simulator	0	6	2
Execution Error			

Table 4 shows the incident rate of sample point selection failures and the tendency of each method to generate feasible solutions. OFFP produced the best optimal solution and the most feasible points with the least inner-loop failures. PFCON rarely converged to a sample point with its inner-loop optimization; however, when a sample point was selected, it had the highest rate of feasible sample points. UTB constraints rarely selected feasible points, and frequently sampled points where the simulator could not converge to a solution. These trends observed between the different constraint methods are consistent with the conclusions of Sohst et al [14].

## CONCLUSIONS

Each of the three tested constraint-handling methods yielded a feasible optimal solution; however, none of these methods were able to converge on an optimum or surpass a solution found in the design of experiments in the allowed number of iterations.

OFFP resulted in the best optimal solution found and had the lowest incidence of failures to select a new sample point. Finding a similar solution to OFFP, UTB sampled far more infeasible points than any of the other algorithms. PFCON suffered from several incidences where the next point selection algorithm failed to yield a sample point. However, when a sample point was determined, PFCON was the most likely to yield a feasible solution. The numerical issues of PFCON could be addressed by modifying the implementation of EGO used.

One potential improvement is to replace the multi-start SLSQP optimization algorithm to find the next sample point. The expected improvement objective function is highly multimodal and evaluates very quickly, lending itself well to heuristic optimization methods that would fight through local minima more effectively.

The effects of different constraint parameter values and larger numbers of EGO iterations on algorithm performance were not studied. More iterations could result in convergence to a better solution. There is a chance that different lower bound for probability of feasibility in PFCON would improve its reliability. Future work could compare parameters for each constraint method on different types of functions over more iterations.

Ultimately, applying these optimization methods to a chemical process simulator produced similar results to previous work on aerodynamic shape optimization. Further improvements on the algorithm would be necessary for constrained EGO to be of practical use for optimizing process simulations.

## ACKNOWLEDGEMENTS

The authors acknowledge Bryan Research & Engineering, LLC for providing the resources to conduct this work.

## REFERENCES

1. Elbashir, N.O., El-Halwagi, M.M., Economou, I.G., Hall, K.R.: Natural Gas Processing from Midstream to Downstream. Wiley (2018).
2. Vu, K.K., D'Ambrosio, C., Hamadi, Y., Liberti, L.: Surrogate-based methods for black-box optimization. *Int. Trans. Oper. Res.* 24, 393–424 (2017).
3. Pardalos, P.M., Rasskazova, V., Vrahatis, M.N. eds: Black Box Optimization, Machine Learning, and No-

- Free Lunch Theorems. Springer International Publishing, Cham (2021).
- Cozad, A., Sahinidis, N.V., Miller, D.C.: Learning surrogate models for simulation-based optimization. *AIChE J.* 60, 2211–2227 (2014).
  - Wilson, Z.T., Sahinidis, N.V.: The ALAMO approach to machine learning. *Comput. Chem. Eng.* 106, 785–795 (2017).
  - Ma, K., Sahinidis, N.V., Amaran, S., Bindlish, R., Bury, S.J., Griffith, D., Rajagopalan, S.: Data-driven strategies for optimization of integrated chemical plants. *Comput. Chem. Eng.* 166, 107961 (2022).
  - Brambilla, A., Vaccari, M., Pannocchia, G.: Analytical RTO for a critical distillation process based on offline rigorous simulation. *IFAC-Pap.* 55, 143–148 (2022).
  - Thon, C., Finke, B., Kwade, A., Schilde, C.: Artificial Intelligence in Process Engineering. *Adv. Intell. Syst.* 3, 2000261 (2021).
  - Regis, R.G.: A Survey of Surrogate Approaches for Expensive Constrained Black-Box Optimization. In: Le Thi, H.A., Le, H.M., and Pham Dinh, T. (eds.) *Optimization of Complex Systems: Theory, Models, Algorithms and Applications*. pp. 37–47. Springer International Publishing, Cham (2020).
  - Palmer, K., Realff, M.: Metamodeling Approach to Optimization of Steady-State Flowsheet Simulations. *Chem. Eng. Res. Des.* 80, 760–772 (2002).
  - Palmer, K., Realff, M.: Optimization and Validation of Steady-State Flowsheet Simulation Metamodels. *Chem. Eng. Res. Des.* 80, 773–782 (2002).
  - Caballero, J.A., Grossmann, I.E.: An algorithm for the use of surrogate models in modular flowsheet optimization. *AIChE J.* 54, 2633–2650 (2008).
  - Jones, D.R., Schonlau, M., Welch, W.J.: Efficient Global Optimization of Expensive Black-Box Functions. *J. Glob. Optim.* 13, 455–492 (1998).
  - Sohst, M., Afonso, F., Suleman, A.: Surrogate-based optimization based on the probability of feasibility. *Struct. Multidiscip. Optim.* 65, 10 (2022).
  - Durantín, C., Marzat, J., Balesdent, M.: Analysis of multi-objective Kriging-based methods for constrained global optimization. *Comput. Optim. Appl.* 63, 903–926 (2016).
  - Audet, C., Caporossi, G., Jacquet, S.: Binary, unrelaxable and hidden constraints in blackbox optimization. *Oper. Res. Lett.* 48, 467–471 (2020).
  - Jones, A.C., Marples, D.J.: The Section 45Q Tax Credit for Carbon Sequestration. Congressional Research Service (2023).
  - Saves, P., Lafage, R., Bartoli, N., Diouane, Y., Bussemaker, J.H., Lefebvre, T., Hwang, J.T., Morlier, J., Martins, J.R.R.A.: SMT 2.0: A Surrogate Modeling Toolbox with a focus on Hierarchical and Mixed Variables Gaussian Processes. *ArXiv Prepr.* (2023).
  - Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., SciPy 1.0 Contributors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods.* 17, 261–272 (2020).

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Industrial Biosolids from Waste to Energy: Development of Robust Model for Optimal Conversion Route – Case Study

Hesan Elfaki<sup>a\*</sup>, and Dhabia M. Al-Mohannadi<sup>b</sup>

<sup>a</sup> Texas A&M University, Texas Sea Grant College Program, College Station, Texas, United States

<sup>b</sup> Texas A&M University at Qatar, Chemical Engineering Program, Education City, Doha, Qatar

\* Corresponding Author: [h.elfaki@tamu.edu](mailto:h.elfaki@tamu.edu)

## ABSTRACT

Utilizing sustainable energy sources is crucial for expanding the range of solutions available to meet the growing energy demand and reducing reliance on environmentally damaging and depleting conventional fuels. Biosolids, a type of biomass, are generated as secondary effluent during wastewater treatment process in municipal and industrial sites. These solids possess the potential to serve as a sustainable energy source due to their richness of carbon. For an extended period, biosolids have been landfilled, even though it can be considered a wasteful use of a precious resource and a possible mean for contamination to the food supply chain. This has served as an extra impetus to investigate the potential for harnessing the capabilities of these substances. While many research studies have looked at different ways to put biomass waste to use, very little has been written on biosolids, especially those derived from industrial sources. This research assesses the feasibility of transforming GTL derived biosolids into value-added commodities that can serve as raw materials in chemical manufacturing or be employed energy generation. The study primarily examines widely recognized thermal conversion processes, pyrolysis and gasification. An evaluation is carried out to analyze the economic, technological, and environmental aspects of the treatment methods utilizing these technologies. The aim is to demonstrate the potential of GTL biosolids conversion and to determine associated costs and environmental impacts. The ASPEN simulation tool is utilized to model thermal treatment pathways, allowing for the generation of economic and environmental estimations for each route.

**Keywords:** Biosolids, Energy, Utilization, Simulation

## INTRODUCTION

Biosolids are a significant waste stream that is discharged from wastewater treatment plants and water management facilities around the world. Although biosolids have traditionally been regarded as a waste by-product and disposed of in landfills for many years, current research has demonstrated that they possess significant value as a result of their nutrient and energy content [1, 2]. They are now formally classified as renewable sources of energy because of the energy value derived from their organic component [3, 4]. The utilization of biosolids for energy generation has a significant potential to promote a circular economy and reduce environmental waste [4]. Possible uses for the recovery of mass and energy from biosolids include using them as fertilizer for

land, composting, anaerobic digestion, combustion, gasification, pyrolysis, and hydrothermal treatment. These applications transform biosolids into valuable resources, such as composts for nutrients or biofuels and char for energy. Each of these uses necessitates a designated preliminary treatment to prepare the biosolids stream. The treatment of biosolids is a significant obstacle in water management and treatment plants due to the substantial electrical and operational expenses involved. These costs make up 20% and 53% respectively of the total wastewater treatment process [1]. Pyrolysis and gasification are two thermal conversion methods that have great potential for transforming biosolids into products that have increased value.

Pyrolysis is a process that involves the combination of thermal cracking and catalytic conversion. It is

conducted at relatively moderate temperatures (400 – 700 °C) and atmospheric pressure [5]. The process of producing bio syngas, oil, and fuel char from biosolids involves heating them in an oxygen-free environment [6, 7]. Gasification is a combustion process that involves burning biosolids with a limited amount of oxygen or air at very high temperatures that can reach up to 1200 °C and 0.6-2.6 MPa. Conceptually, this technology has resemblance to pyrolysis, albeit it differs by utilizing a sub-stoichiometric oxygen input and operating at considerably higher temperature settings [8].

Considerable research is readily accessible that examines the costs and economic evaluations of treating and reusing biosolids. AlNouss et al. performed a techno-economic assessment of biomass gasification for hydrogen production [9]. Ghiat et al. provided a comprehensive economic analysis of a biomass gasification system that is combined with power generation and carbon dioxide recovery. The analysis was conducted using aspen simulation software [10]. [11] conducted a study on the techno-economic factors of power generation using biogas produced by gasifying biosolids, utilizing actual data from a facility [11]. However, there is a scarcity of research that has examined the process of converting biosolids produced by industrial facilities in oil and gas sector.

This study assesses the feasibility of transforming GTL industrial biosolids into high-value products through gasification and pyrolysis processes. Analytical investigation is conducted to compare the technological, economic, and environmental aspects of two treatment methods in order to identify the demonstrate efficient paths for treating GTL derived biosolids. The conversion processes are simulated using Aspen process simulation tool.

## PROBLEM STATEMENT

Gas-to-liquids (GTL) technology transforms natural gas, which is the most environmentally friendly fossil fuel, into premium liquid products that are typically derived from crude oil. The process produces biosolid waste streams that need to be treated to uncover their energy potential and reduce the need for solid waste management.

To assess the feasibility of treating a continuous flow of GTL biosolids, which includes information on flowrate, solid content, and moisture content, it is necessary to examine the viability of employing thermal pyrolysis and gasification methods. Additionally, it is crucial to quantify the costs and environmental impacts associated with these treatment approaches.

## ASPEN MODEL

## Process Flowsheet

The diagram in Figure 1 depicts the process of converting GTL biosolids through pyrolysis and gasification, as simulated using Aspen Plus. The primary feed stream consists of GTL generated biosolids. The thermo-chemical parameters of the GTL biosolids have been determined using proximate and ultimate analysis, as described in the study by Zuhara et al. [12]. The characteristics are succinctly outlined in Table 1. Figure 1 describes the pyrolysis flowsheet. The initial step involves the treatment of biosolids in a decomposition unit, which is simulated using the Ryield reactor model in Aspen Plus. The device disassembles the non-conventional flow of biosolids into its primary elemental components, namely hydrogen, oxygen, carbon, nitrogen, and sulfur. The output of this unit is sent to the pyrolysis reactor, which is simulated using the RGibbs reactor model. This model uses Gibbs energy minimization to estimate the products produced by the reactions. Subsequently, the pyrolysis products are sent to a separator, where the resultant gas is isolated from the solid phase products, namely ash and char. After the separation process, the gas undergoes a cooling process, which leads to the retrieval of biogas as the final outcome.

**Table 1:** GTL biosolids proximate and ultimate analysis [12]

<b>Proximate Analysis</b>	
Moisture	12.29
Volatile matter	47.47
Ash	26.07
Fixed carbon	14.16
<b>Ultimate Analysis</b>	
C	33.69
H	6.18
O	29.03
N	5.03
S	-

The gasification flow sheet illustrated in Figure 2, diverts the elemental components produced in the decomposition unit to an alternative treatment pathway. Post decomposition, a solid separation unit is used to initially separate ash and char products from the gas stream. Subsequently, the gas stream leaving the separation unit is combined with steam and air before being introduced into the gasification reactor. Steam and air act as gasification agents in this process. The gasifier is simulated using the RGibbs reactor model. The gasifier products undergo a cooling process and are subsequently separated in a solid separation unit to obtain biogas and char products.

The simulation utilized thermodynamic calculations employing the Peng Robinson property method with the Boston modification to improve the accuracy of estimating the properties of non-conventional inputs, specifically GTL biosolids and Ash. The operational parameters are concisely outlined in Table 2.

## RESULTS AND DISCUSSION

The simulation model was constructed under the assumption of steady-state operation under atmospheric pressure, disregarding pressure variations throughout unit operations and heat dissipation in gasification and pyrolysis units.

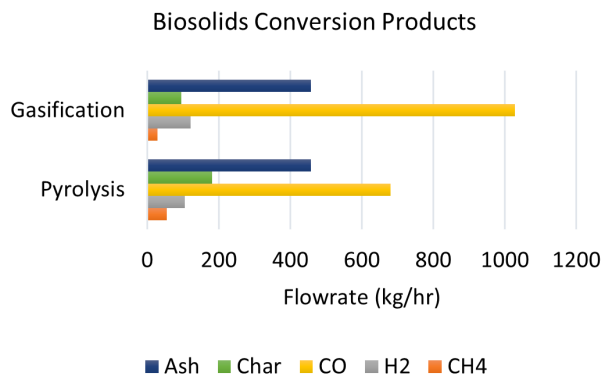
**Table 2** Pyrolysis and Gasification flowsheet operation conditions

Parameter	Unit	Value
<b>Biosolids Feed</b>		
Feed flowrate	Kg/hr	2000
Temperature	°C	25
<b>Pyrolysis</b>		
Inert flowrate	Kg/hr	40
Inert temperature	°C	65
Pyrolysis unit temperature	°C	700
Decomposer temperature	°C	700
<b>Gasification</b>		
Air flowrate	Kg/hr	40
Air temperature	°C	65
Gasifier temperature	°C	800
Decomposer temperature	°C	700

The model presupposes that the char consists entirely of solid carbon.

The simulation model was executed using a feed flowrate of 2000kg/hr of GTL biosolids, as described by the provided approximate and ultimate analysis. Prior to being sent to the conversion units, pyrolysis and gasifier, the decomposition unit transformed the nonconventional biosolid material into its primary components. After the conversion has taken place, The solid splitters were used to separate the ashes and char, while the resulting biogas was pre-cooled to recover heat before being collected as the end product. The composition of the biogas stream was analyzed to determine the amounts of syngas and methane generated during the process. These gases are commonly used for energy production or as raw materials in downstream petrochemical processes. The accumulated ash can be beneficial in several applications for material utilization, such as in the manufacturing of cement.

Char, in its solid state, is a useful energy source that can be used for carbon sequestration and land reclamation, depending on its specific properties [13]. The production rate of each of these components is illustrated in Figure 3.

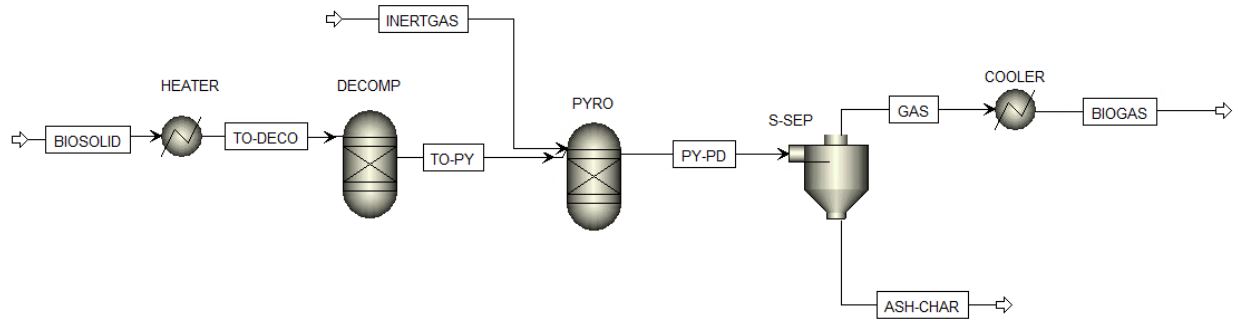


**Figure 3.** Production rate of various thermal conversion products for GTL biosolids feed

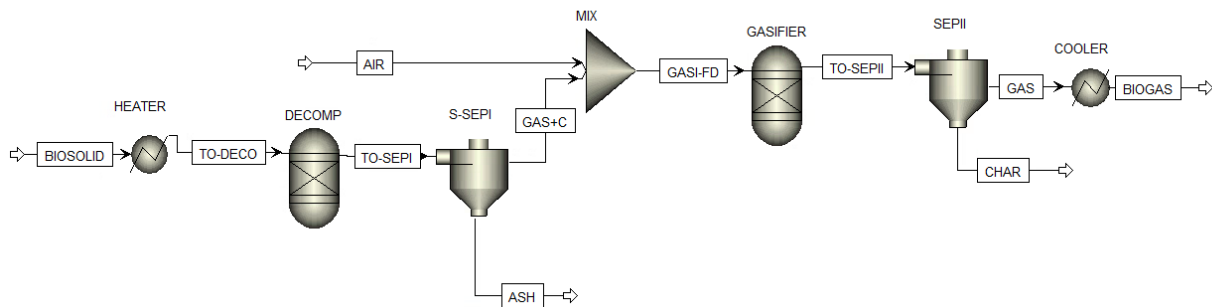
The findings exhibited total transformation of the biosolids into biogas, char, and ash. Upon analyzing the syngas composition in the effluent of the two processes, it is observed that gasification yields 1029 Kg/hr of carbon monoxide (CO) and 121 Kg/hr of hydrogen (H<sub>2</sub>). The generation of CO has increased by approximately 50%, while the production of H<sub>2</sub> has increased by approximately 15% compared to pyrolysis. The increase in syngas creation was achieved at the cost of a 48% decrease in CH<sub>4</sub> output during gasification. The overall quantity of ash exiting the process remains constant in both process outlets, as they possess identical characteristics and receive the same amount of biosolids feed with a predetermined nonconventional ash content, according to the approximate and ultimate analysis of GTL biosolids. Gasification is preferred over pyrolysis when considering the use of biosolids as an extra feedstock stream for a petrochemical plant.

Post process simulation, Aspen program was used to conduct a techno-economic and environmental assessment of the two processes. The model employs the integrated economic analyzer and environmental analysis tool. The tool provided estimates for the financial expenses, capital and operational costs of both scenarios and presented the greenhouse gas (GHG) emissions as carbon equivalents, measured in terms of their ability to contribute to global warming (GWP) [14]. The capital cost encompassed equipment cost estimation via an integrated unit operation mapping and equipment sizing methodology. Figure 4 below provides a concise summary of the cost estimation analysis and the CO<sub>2</sub> equivalent emissions for the two conversion scenarios.

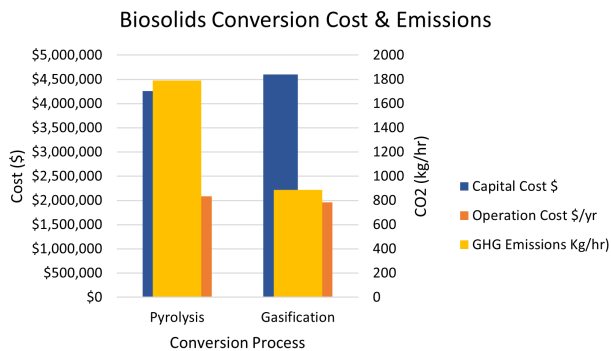




**Figure 1:** Biosolids pyrolysis flowsheet developed in Aspen Plus



**Figure 2:** Biosolids gasification flowsheet developed in Aspen Plus



**Figure 4.** Cost estimation and GHG emissions from biosolids pyrolysis and gasification processes

The results indicate that the capital cost for biosolids gasification is around \$4.6 million, which slightly exceeds the capital required for pyrolysis by 8%. The discrepancy can be rationalized by the disparity in unit operations employed in each scenario's flowsheet and the

chosen materials capable of withstanding the operational temperatures of the conversion units. The model allocated suitable utilities to the various unit operation as per the required duties and pressure and temperature ranges. In order to meet the conversion process requirements in both scenarios, three primary utility categories were chosen: fuel, cooling water, and power. The overall operational cost, with utility expenses being the primary factor, showed an insignificant variance between the two scenarios. The anticipated cost for pyrolysis was \$2.09 million per year, while the cost for the other method was \$1.97 million per year. The CO<sub>2</sub> emissions depicted in Figure 4 exhibit a significant disparity in emissions between pyrolysis at a rate of approximately 1800 kg/hr and gasification 900 kg/hr. According to this model, pyrolysis results in double the amount of gasification GHG emissions. In gasification, the conversion process results in the production of additional CH<sub>4</sub> and CO<sub>2</sub>. However, these gases are then further converted into H<sub>2</sub> and CO products, which helps to reduce greenhouse gas emissions from the process side of operation.

## CONCLUSION

Biosolids are a valuable byproduct that must be effectively treated in order to be utilized. There is a dearth of research exploring the possibilities of utilizing industrial biosolids in literature. This study focuses on examining the transformation of biosolids derived from GTL (Gas-to-Liquid) into biogas products, specifically hydrogen (H<sub>2</sub>), carbon monoxide (CO), and methane (CH<sub>4</sub>). The investigation is conducted through the utilization of an Aspen simulation flowsheet that has been specifically developed for this purpose. Two primary thermal conversion pathways have been developed, which involve the integration of gasification and pyrolysis processes. Gasification emerges as the ideal method for treating biosolids and using them as an additional source of feedstock in a petrochemical plant. The results suggest that gasification produces greater amounts of carbon monoxide (CO) and hydrogen (H<sub>2</sub>) in comparison to pyrolysis, which is beneficial for specific industrial uses. Gasification reduces methane (CH<sub>4</sub>) emissions, while maintaining effective conversion of biosolids into biogas, char, and ash.

The Aspen software allows for technoeconomic and environmental studies, which offer additional insights. Although gasification has a somewhat larger initial investment compared to pyrolysis, the ongoing expenses for both techniques are practically the same. This implies that although the initial cost for gasification may be slightly greater, the operational effectiveness and long-term advantages compensate for this disparity. The environmental impact connected with each activity is of paramount significance. Gasification exhibits a notable superiority over pyrolysis in relation to greenhouse gas emissions since it leads to reduced emissions of carbon dioxide (CO<sub>2</sub>) equivalents. The decrease in emissions is credited to the transformation of excess methane and carbon dioxide into hydrogen and carbon monoxide, effectively reducing greenhouse gas emissions from the operational process.

Ultimately, when it comes to utilizing biosolids as an extra input for petrochemical operations, gasification proves to be a beneficial choice. Gasification, despite having somewhat higher initial costs, has better conversion efficiency, similar running expenses, and reduced environmental effect when compared to pyrolysis.

In conclusion, this study demonstrates the capacity to transform industrial biosolids wastes into a viable energy source or valuable raw material, while simultaneously reducing the amount of solid waste generated. These findings provide valuable insights for decision-making processes as it enables well-informed judgments regarding biosolids treatment strategy that achieves a balance between the efficiency of conversion efficiency, economic feasibility, and environmental impacts.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge funding provided by Qatar Shell Research and Technology Centre (QSTRC).

## REFERENCES

1. Zhao, G., Garrido-Baserba, M., Reifsnnyder, S., Xu, J.C., Rosso, D. Comparative energy and carbon footprint analysis of biosolids management strategies in water resource recovery facilities. *Science of the Total Environment*. 665, 762–773 (2019).
2. Wang, H., Brown, S.L., Magesan, G.N., Slade, A.H., Quintern, M., Clinton, P.W., Payn, T.W. Technological options for the management of biosolids, *Environ Sci Pollut Res* 15:308–317 (2008)
3. Edwards, J., Othman, M., Crossin, E., Burn, S. Anaerobic co-digestion of municipal food waste and sewage sludge: A comparative life cycle assessment in the context of a waste service provision. *Bioresour Technol*. 223, 237–249 (2017).
4. Egan, M. Biosolids management strategies: An evaluation of energy production as an alternative to land application. *Environ Sci Pollut Res* 20:4299–4310 (2013)
5. Spinosa, L. From sludge to resources through biosolids. *Water Sci Technol*. 50 (9): 1–9 (2004)
6. He, W., Yang, H., Han, D. Thermodynamic investigation and optimization of a heat pump coupled open-air, open-water humidification dehumidification desalination system with a direct contact dehumidifier. *Desalination*. 469: 114101 (2019).
7. Liu, Z., McNamara, P., Zitomer, D. Autocatalytic Pyrolysis of Wastewater Biosolids for Product Upgrading. *Environ Sci Technol*. 51, 9808–9816 (2017).
8. Winchell, L.J., Ross, J.J., Brose, D.A., Pluth, T.B., Fonoll, X., Norton, J.W., Bell, K.Y. Pyrolysis and gasification at water resource recovery facilities: Status of the industry, *Water Environ Res*. 94:e10701(2022)
9. AlNouss, A., McKay, G., Al-Ansari, T. Enhancing waste to hydrogen production through biomass feedstock blending: A techno-economic-environmental evaluation. *Appl Energy*. 266 (2020).
10. Ghat, I., AlNouss, A., McKay, G., Al-Ansari, T. Biomass-based integrated gasification combined cycle with post-combustion CO<sub>2</sub> recovery by potassium carbonate: Techno-economic and environmental analysis. *Comput Chem Eng*. 135 (2020).
11. Carlos, L., Villalba, O., Acevedo, E.R. Technical and

economical assessment of power generation technologies firing syngas obtained from biosolid gasification. *Prod. Limpia* 9: 31-43 (2014)

12. Zuhara, S., Pradhan, S., Pasha, M., McKay, G. Potential of GTL-Derived Biosolids for Water Treatment: Fractionization, Leachate, and Environmental Risk Analysis. *Water (Switzerland)*. 14 (2022).
13. Collivignarelli, M.C., Canato, M., Abbà, A., Carnevale Miino, M. Biosolids: What are the different types of reuse? *J. Clean. Prod.* 238: 117844, (2019)
14. Shahbaz, M., Alnouss, A., Parthasarathy, P., Abdelaal, A.H., Mackey, H., Mckay, G., Al-Ansari, T. Investigation of biomass components on the slow pyrolysis products yield using Aspen Plus for techno-economic analysis. *Biomass Convers. Biorefin.* 12:669–681 (2022)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Analysis of Infrastructures for Processing Plastic Waste using Pyrolysis-Based Chemical Upcycling Pathways

Evan D. Erickson<sup>a</sup>, Jiaze Ma<sup>a</sup>, Philip Tominac<sup>a</sup>, Horacio Aguirre-Villegas<sup>b</sup>, and Victor M. Zavala<sup>a\*</sup>

<sup>a</sup> University of Wisconsin-Madison, Department of Chemical & Biological Engineering, 1415 Engineering Dr, Madison, WI 53706, USA

<sup>b</sup> University of Wisconsin-Madison, Department of Biological Systems Engineering, 460 Henry Mall, Madison, WI 53706, USA

\* Corresponding Author: [victor.zavala@wisc.edu](mailto:victor.zavala@wisc.edu).

## ABSTRACT

Modern mechanical recycling infrastructure for plastic is capable of processing only a small subset of waste plastics, reinforcing the need for parallel disposal methods such as landfilling and incineration. Emerging pyrolysis-based chemical technologies can “upcycle” plastic waste into high-value polymer and chemical products and process a broader range of waste plastics. In this work, we study the economic and environmental benefits of deploying an upcycling infrastructure in the continental United States for producing low-density polyethylene (LDPE) and polypropylene (PP) from post-consumer mixed plastic waste. Our analysis aims to determine the market size that the infrastructure can create, the degree of circularity that it can achieve, the prices for waste and derived products it can propagate, and the environmental benefits of diverting plastic waste from landfill and incineration facilities it can produce. We apply a computational framework that integrates techno-economic analysis, life cycle assessment, and value chain optimization. Our results demonstrate that the infrastructure generates an economy of nearly 20 billion USD and positive prices for plastic waste, opening opportunities for compensation to residents who provide plastic waste. Our analysis also indicates that the infrastructure can achieve a plastic-to-plastic degree of circularity of 34% and remains viable under various external factors (including technology efficiencies, capital investment budgets, and polymer market values). Finally, we present significant environmental benefits of upcycling over alternative landfill and incineration waste disposal methods, and comment on ongoing work expanding our modeling methodology to other chemical upcycling pathway case studies, including hydroformylation of specific plastics to chemicals.

**Keywords:** Polymers, Optimization, Supply Chain, Modelling, Interdisciplinary

## INTRODUCTION

With global rates of plastic production reaching upwards of 400 million tonnes per year and growing [1], nations require integrated, reliable plastic waste management systems [2]. Despite reductions in the environmental impacts of plastic production [3], global plastic recycling rates remain well below 10% [4], necessitating the continued use of parallel disposal methods (such as landfilling and incineration) and allowing significant plastics to accumulate in the environment [5].

Most current recycling occurs by well-established mechanical methods, which are designed to accept bottles of polyethylene terephthalate (PET) and high-density polyethylene (HDPE), together making up less than 12%

of post-consumer plastic waste [6]. However, these methods also produce degradation and allow these plastics to only be recycled a finite number of times [7]. Mechanical recycling strategies cannot process most other resins or flexible, multilayer plastics (which account for 35% of plastics currently produced [8]). The parallel disposal methods required for these plastics cause significant environmental and social issues.

Chemical upcycling has been proposed as a complement to mechanical systems. These pathways apply a complex sequence of chemical processing steps to decompose plastics into constitutive chemical species that can be further refined into value-added products, such as fuels, chemicals, and plastic precursors. Due to their improved feedstock versatility, chemical upcycling

pathways can be more scalable than mechanical recycling pathways. We analyze a pathway which requires pre-sorting and purification of mixed plastic waste from post-consumer recyclable streams, occurring at material recovery facilities (MRFs) and plastic reprocessing facilities (PRFs). Sorted plastic waste is baled and mechanically reduced to flakes before undergoing pyrolysis and steam cracking to produce plastic monomers, which can finally be repolymerized into resins (here, low-density polyethylene and polypropylene) with identical quality as the starting materials and the same original properties as virgin plastics [4]. Thus, chemical upcycling pathways can help to achieve a circular economy for plastic waste.

Many recent analyses have explored the potential contributions of upcycling to broader plastic waste management systems [4, 7, 9, 10, 11] and its economic and environmental performance [12, 13, 14]. However, a significant gap exists in our understanding of the deployment of upcycling technologies at scale. Previous works have developed models to analyze plastic waste supply chains [15, 16] and supply chain optimization frameworks for studying these problems at regional scales [6]. Large supply chain models can help to analyze complex economic behavior of networks of processing facilities and track the transportation of products in the value chain.

In this paper, we summarize the methods and results of our recent full paper on the national-scale modeling of an upcycling pathway for the continental US [17], and present updates in ongoing comparative analysis of other upcycling pathways with different process details and outcomes on the same national scale, including the recently explored hydroformylation process [18].

We overview a computational framework to examine the potential for deploying a national-scale upcycling infrastructure for post-consumer plastic waste in the continental US, and present an optimal upcycling value chain solution (see Figure 1) for techno-economic analysis and lifecycle assessment. The infrastructure is optimized to maximize total profit while minimizing investment, operating, and transportation costs. We show that this infrastructure can be interpreted as a value chain (an economy) under which stakeholders exchange plastic waste and derived products to maximize profit. In our model, upcycling creates high value from the production of virgin polymers from post-consumer waste, producing a total economy (market) size of 19.7 billion USD per year and remaining viable under various changes to market inputs.

Upcycling as a national plastic waste management strategy achieves a 34% plastic-to-plastic degree of circularity for converting residential plastic waste into high quality polymers. The system compares favorably to existing disposal methods for mixed plastic by environmental metrics, including 69–75% reductions in greenhouse gas (GHG) emissions relative to incineration and 86 avoided new landfills over the next 50 years).

In contrast to this form of pyrolysis-based upcycling (a plastic-to-plastic method for mixed plastics), hydroformylation has been envisioned as a plastic-to-chemicals pathway for more specific plastic feedstocks. To date, hydroformylation has only been explored for one post-consumer resin (HDPE) and three virgin resins (HDPE, LDPE, and PP) [18]. Thus, hydroformylation would compete with mechanical recycling for HDPE. We comment that current market prices for the products of hydroformylation may not support the pathway's use relative to prices for mechanically recycled HDPE.

## METHODOLOGY

We model the US as a collection of county-level waste-producing centers of population. Our chemical upcycling pathway comprises a series of six processing technologies, including MRFs, PRFs, pyrolysis, steam cracking, LDPE polymerization, and PP polymerization. The value chain contains 16 different products, including inputs (post-consumer recyclables), intermediates (e.g., plastic waste bales, plastic flake, pyrolysis oil, ethylene), and outputs (e.g., LDPE, PP, fuels). Only feedstock recyclables, plastic bales, plastic flake, and pyrolysis oil are modeled as transportable to represent the infeasibility of transporting many intermediates (e.g., ethylene and propylene gas) and simplify analysis of market demand.

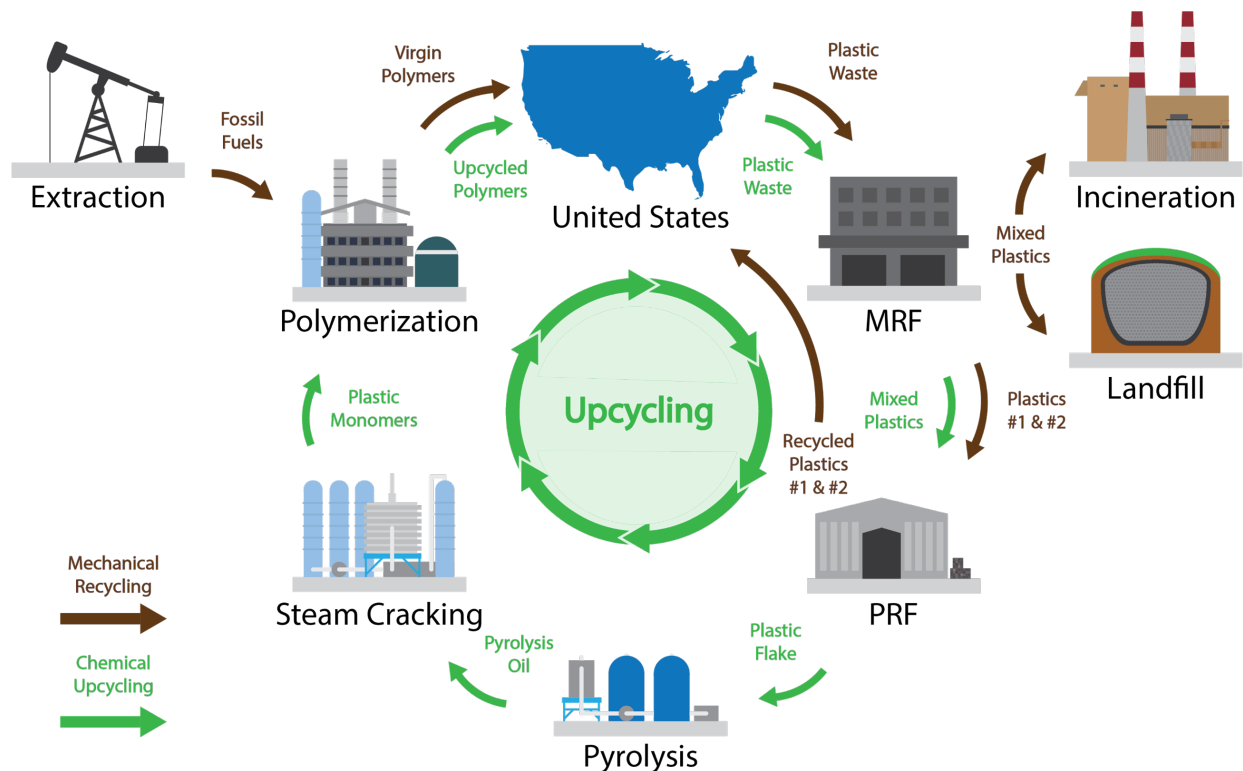
Each component of the value chain has techno-economic definitions from Ma et al. [6]. We applied a mixed-integer program (MIP) formulation to our supply chain optimization model, also proposed by Ma et al. [6] to study upcycling in the US Midwest region. The MIP designs a value chain by placing processing facilities in locations that best capture residential waste, maximize profit, and minimize investment, processing, and transportation costs. Furthermore, we apply LCA methodologies to compare GHG emissions of upcycling, landfilling, and incineration. Emission factors for each technology are adapted from Ma et al. [6].

To limit model size and computational complexity for a national case study, we allow transportation routes between adjacent county centers. Still, the national scale presents challenges for current solvers, so we report close-to-optimal solutions with optimality gaps demonstrating their measured distance from the estimated upper bound on total profit.

The model comprises 3,108 county centers and 4,010 sites available for technology placement, with two sizes of technology possible at each potential site. In total, the MIP problem contains 796,756 continuous variables, 8,020 binary variables, and 587,405 constraints.

To implement and run the MIP formulation, we utilize the algebraic modeling package JuMP [17] for the Julia programming language and solve using Gurobi 9.1.0 [20]. We use the MIP to design the physical arrangement





**Figure 1:** Overview of a pyrolysis-based chemical upcycling value chain. In chemical upcycling, waste plastics are recovered as fuels, chemicals, and virgin-quality polymers in a sequence of processing steps. Our case study examines a combination of mechanical recycling for PET (plastic type #1) and HDPE (plastic type #2), and chemical upcycling of mixed plastics (plastic types #3 - #7). This combined network creates an economy worth nearly 20 billion USD annually and processes 100% of US residential plastic waste, achieving a degree of circularity of 33.9% (more than one-third of input waste plastics are recovered as mechanically or chemically recycled plastics) while most other remaining inputs are converted into valuable byproduct fuels and chemicals.

of the upcycling value chain and then perform economic and environmental analysis using the **CoordinatedSupplyChains.jl** package (**CSC.jl**) developed by Tominac and Zavala [21] and modified for this work with visualization capabilities. After 180 hours of computational time, the MIP was solved to an optimality gap of 4.66%, meaning that the identified system design provides an economic result within 5% of the optimal solution. The optimal design reaches all 3,108 counties and captures 100% of residential plastic waste; in other words, the model indicates that there are sufficient economic incentives to harness all the post-consumer waste to obtain virgin plastics.

## RESULTS

### Circularity & Economic Evaluation

The final upcycling network captures 100% of residential waste, recovering 11.6% of input waste plastics (the fraction of PET and HDPE bottles) as mechanically-recycled HDPE and PET through MRFs and PRFs, and 22.3% as LDPE and PP through pyrolysis, steam cracking,

and repolymerization. This leads to a total plastic-to-plastic degree of circularity of 33.9%, but overall, 94.6% of residential plastics are recovered as valuable products (including significant fuel and chemical byproducts of upcycling which enhance upcycling's economic viability).

Prices for plastic waste are positive for all county centers, demonstrating that the upcycling network is economically viable without added incentives through policy, largely due to the production of high-value LDPE and PP through chemical upcycling. Upcycling creates an economy worth 19.7 billion USD annually (a measure of total economic surplus, excluding capital costs) and nearly 10 billion in residential profit, which could be applied to compensate residents for recycling properly (or donating their waste plastics to upcycling systems) at a rate of 31 USD / year.

### Sensitivity Analysis

External market factors and technical specifications impact the economic performance of the upcycling network. Total waste capture remains above 99% for resident offering costs up to 40 USD / tonne (rather than

donating plastic waste for free). Improving the efficiencies of key technologies (such as pyrolysis and steam cracking) increases economic surplus but can exhaust the processing capacity of the network faster, excluding some residential waste providers from the market. Increasing transport costs for feedstocks and intermediates has a minimal impact on process economics (transport costs can be doubled and the system processes 99.47% of residential waste plastic). The network can withstand 20% reductions in LDPE and PP market value (major economic drivers of viability) without significant decreases in waste capture. Finally, limiting the overall capital cost investment available for constructing the network limits its extent of waste capture, prioritizing regions of higher waste production density excluding regions including the Great Plains and Northwest.

### Comparison to Existing Processing Methods

The upcycling network emits 1.387 tonne CO<sub>2</sub> eq / tonne of plastic, including processing and transportation contributions. Landfills produce limited direct GHG emissions (0.0236 tonne CO<sub>2</sub> eq / tonne of plastic), but impose other categories of environmental impacts including land use, leakage and microplastic pollution. Incineration or waste-to-energy (WtE) systems produce 2.3-3.1 tonnes CO<sub>2</sub> eq of direct emissions per tonne of plastic combusted depending on resin type.

However, upcycling and WtE fulfill additional functions not achieved by landfills: upcycling produces virgin-quality polymers and WtE can produce electricity. Applying LCA methodology to credit upcycling for its avoided fossil fuel use and WtE for its avoided grid electricity consumption, we determine revised emissions factors of 0.508 tonne CO<sub>2</sub> eq / tonne plastic (including transport) for upcycling and 1.6-2.0 tonnes CO<sub>2</sub> eq per tonne of plastic combusted for WtE. With credits, upcycling emissions compare more favorably to WtE, whose credits based on grid electricity will gradually disappear as the grid becomes decarbonized.

### Comparison to Hydroformylation

While we model a plastic-to-plastic upcycling pathway with fuel and chemical byproducts, other pathways with other aims also attract the interest of researchers. In particular, Li et al. [18] studied hydroformylation (another pyrolysis-based pathway) to produce aromatics, paraffin, alcohols, and dialcohols from post-consumer HDPE or virgin HDPE, LDPE, and PP. The inputs of this pathway also determine its applications: as developed in this publication for a HDPE waste feedstock, hydroformylation would not be an alternative to pyrolysis-based upcycling in this work and Ma et al. [6] but to mechanical recycling. However, since the pathway is still based upon pyrolysis, it could be applied to mixed plastic wastes in the future.

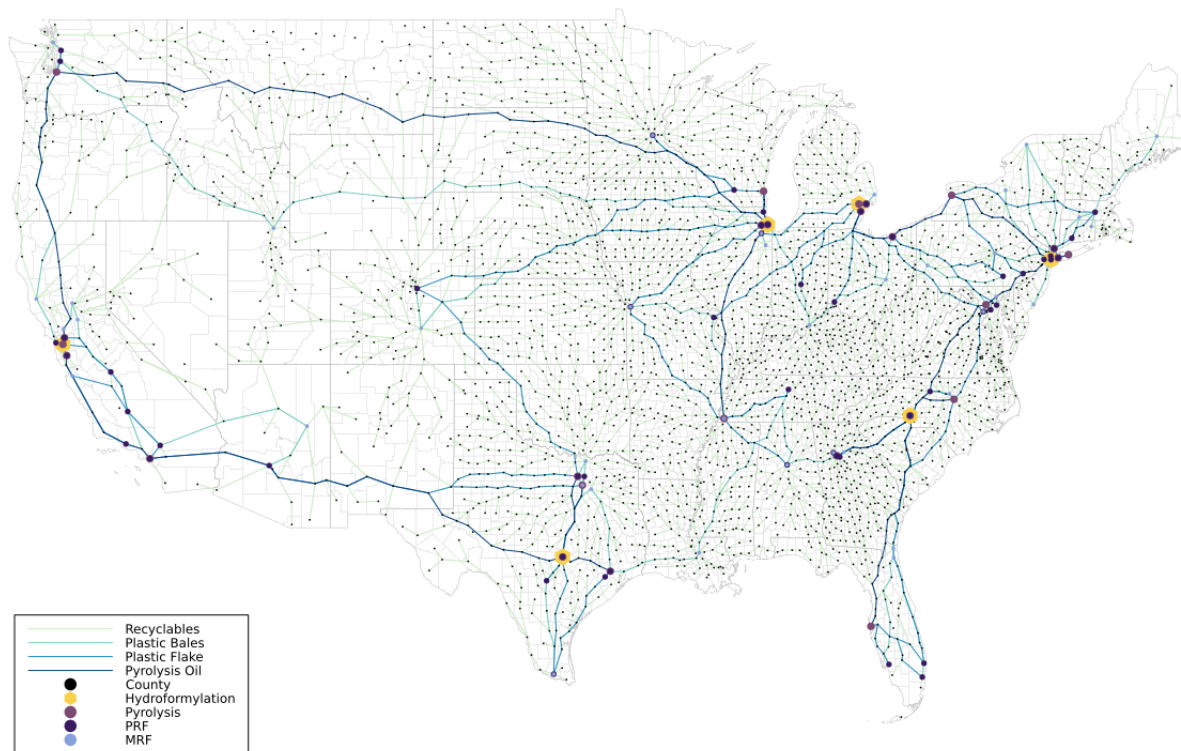
Li et al. [18] report yields and market values of the

four chemical categories produced by hydroformylation of waste HDPE, which we apply as preliminary target yields for the hydroformylation of mixed plastic wastes. We determine that hydroformylation provides a high combined revenue for pyrolysis oil (2154 USD / tonne) from which additional operating costs (644 USD / tonne) detract significantly, leading to an adjusted pyrolysis oil value of 1509 USD / tonne and a corresponding post-MRF-and-PRF plastic waste value of 1035 USD / tonne. This is slightly above the market value prices for recycled HDPE produced at MRFs and PRF used in Ma et al. [6] and this work (991 USD / tonne), but is significantly above the market value produced for post-MRF-and-PRF mixed plastic wastes from plastic-to-plastic upcycling (711 USD / tonne).

We extend our model to consider a hydroformylation pathway and map its optimal infrastructure arrangement and product transportation flows in Figure 2. The comparatively higher value for mixed plastic waste supported by hydroformylation allows plastic waste to sustain greater transportation distances and costs, and thus, the resulting arrangement possesses fewer decentralized components. For example, the plastic-to-plastic upcycling solution placed 336 mostly small-scale MRFs (the first step in each pathway) whereas the hydroformylation solution placed only 38 mostly large-scale MRFs. The total MRF capacity of the hydroformylation solution was nearly 18% lower than that of plastic-to-plastic upcycling, but the utilization of available MRF capacity was much greater (98.8% compared to 81.2%). In contrast, the number of hydroformylation hubs in this solution is roughly comparable to the number of steam cracking and polymerization combined hubs in the plastic-to-plastic upcycling solution.

A more centralized network entails trade-offs with supply, as in the hydroformylation solution, only 99.36% of suppliers (3088 counties out of 3108) participate in the network, compare to 100% in the plastic-to-plastic upcycling solution. Future work will aim to refine this solution to achieve 100% waste capture. For the residents that participate in the hydroformylation network, a potential compensation rate of 30 USD / year is generated through the value chain, which is just below that of plastic-to-plastic upcycling.

Hydroformylation achieves different production objectives than plastic-to-plastic upcycling, and with different environmental performance. The uncredited emissions of the hydroformylation network reach 1.473 tonne CO<sub>2</sub> eq / tonne of plastic processed, which is greater than that of plastic-to-plastic upcycling. We note that this analysis does not consider that hydroformylation products would otherwise be produced by expensive fossil-based pathways. Furthermore, hydroformylation produces only chemicals, so its plastic-to-plastic degree of circularity is 0%.



**Figure 2:** Optimal infrastructure placement and product transportation patterns for a hydroformylation alternative to plastic-to-plastic upcycling.

## DISCUSSION & IMPLICATIONS

Previous research has defined the individual steps of upcycling and modeled their implementation at smaller scales. Our analysis reveals that upcycling remains viable when expanded to the continental US, contributing to a circular economy for plastics driven by the high value of upcycling's end products. Upcycling also presents environmental impact reductions in land use relative to landfills and GHG emissions relative to WtE plants.

As demonstrated in this work, system-level modeling projects have the potential to illuminate not only economic and environmental behaviors of potential value chains, but also comparative results between processing options. Our future work will continue expanding our case studies in hydroformylation and other proposed chemical upcycling pathways, while also incorporating ways to relax some of the major simplifying assumptions in our model structure. We are also interested in more directly modeling policy interventions (including plastic waste source reduction or incentives for upcycling technologies) to better represent the role of policy in national

scale systems and the considerations of current US waste management, recycling, and plastic pollution prevention strategies.

## DIGITAL SUPPLEMENTARY MATERIAL

The CSC.jl analysis package is publicly available at <https://juliahub.com/ui/Packages/CoordinatedSupply-Chains>. All scripts and data needed to reproduce the results are available at [https://github.com/zavalab/Julia-Box/tree/master/National\\_Plastic](https://github.com/zavalab/Julia-Box/tree/master/National_Plastic).

## ACKNOWLEDGEMENTS

This material is based upon work supported by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Bioenergy Technologies Office under Award Number DEEE0009285.

## REFERENCES

1. Statista. Plastic waste in the U.S. Technical report, Statista, 2022.
2. UNEP. From Pollution to Solution: A global

- assessment of marine litter and plastic pollution. Technical report, United Nations Environmental Programme, Oct. 2021. Section: publications.
- ACC. Report: Shrinking Environmental Footprint in Plastics Manufacturing. Technical report, American Chemistry Council, Sept. 2022.
  - Li, Houqian et al. Expanding Plastics Recycling Technologies: Chemical Aspects, Technology Status and Challenges, June 2022.
  - Q. Hou, M. Zhen, H. Qian, Y. Nie, X. Bai, T. Xia, M. Laiq Ur Rehman, Q. Li, and M. Ju. Upcycling and catalytic degradation of plastic wastes. *Cell Reports Physical Science*, 2(8):100514, Aug. 2021.
  - J. Ma, P. A. Tominac, H. A. Aguirre-Villegas, O. O. Olafasakin, M. M. Wright, C. H. Benson, G. W. Huber, and V. M. Zavala. Economic evaluation of infrastructures for thermochemical upcycling of post-consumer plastic waste. *Green Chemistry*, 25(3):1032–1044, Feb. 2023. Publisher: The Royal Society of Chemistry.
  - C. Jehanno, J.W. Alty, M. Roosen, S. De Meester, A. P. Dove, E. Y.-X. Chen, F. A. Leibfarth, and H. Sardon. Critical advances and future opportunities in upcycling commodity polymers. *Nature*, 603(7903):803–814, Mar. 2022. Number: 7903 Publisher: Nature Publishing Group.
  - Z. Xu, K. L. Sanchez-Rivera, A. Munguia-Lopez, M. Ochs, K. Nelson, R. V. Lehn, E. Bar-Ziv, H. A. Aguirre-Villegas, and G.W. Huber. Recycling of Plastic Films through Solvent Targeted Recovery and Precipitation Fact Sheet. Technical report, Chemical Upcycling of Waste Plastics, 2023.
  - X. Chen, Y. Wang, and L. Zhang. Recent Progress in the Chemical Upcycling of Plastic Wastes. *ChemSusChem*, 14(19):4137–4151, 2021. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/cssc.202100868>.
  - L. T. J. Korley, T. H. Epps, B. A. Helms, and A. J. Ryan. Toward polymer upcycling—adding value and tackling circularity. *Science*, 373(6550):66–69, July 2021. Publisher: American Association for the Advancement of Science.
  - P. S. Roy, G. Garnier, F. Allais, and K. Saito. Strategic Approach Towards Plastic Waste Valorization: Challenges and Promising Chemical Upcycling Possibilities. *ChemSusChem*, 14(19):4007–4027, 2021. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/cssc.202100904>.
  - B. Hernández, P. Kots, E. Selvam, D. G. Vlachos, and M. G. Ierapetritou. Techno-Economic and Life Cycle Analyses of Thermochemical Upcycling Technologies of Low-Density Polyethylene Waste. *ACS Sustainable Chemistry & Engineering*, 11(18):7170–7181, May 2023. Publisher: American Chemical Society.
  - Y. Luo, E. Selvam, D. G. Vlachos, and M. Ierapetritou. Economic and Environmental Benefits of Modular Microwave-Assisted Polyethylene Terephthalate Depolymerization. *ACS Sustainable Chemistry & Engineering*, 11(10):4209–4218, Mar. 2023. Publisher: American Chemical Society.
  - A. d. C. Munguia López, D. Göreke, K. L. S´anchez-Rivera, H. A. Aguirre-Villegas, S. Avraamidou, G. Huber, and V. M. Zavala. Quantifying the Environmental Benefits of a Solvent-Based Separation Process for Multilayer Plastic Films, Nov. 2022.
  - O. A. Badejo, B. Hernandez, D. G. Vlachos, and M. Ierapetritou. Design of sustainable supply chains for managing plastic waste: A case study for low density polyethylene. Available at SSRN 4516671.
  - Z. Chen, Y. Kimura, and D. T. Allen. Recycled polymers as a feedstock for chemical manufacturing supply chains in the united states: A network analysis for polyethylene pyrolysis. *ACS Sustainable Chemistry & Engineering*, 2023.
  - E. D. Erickson, P. A. Tominac, J. Ma, H. Aguirre-Villegas, and V. M. Zavala. Evaluating the Economic and Environmental Benefits of Deploying a National-Scale, Thermo-Chemical Plastic Waste Upcycling Infrastructure in the United States, Oct. 2023.
  - H. Li, J. Wu, Z. Jiang, J. Ma, V. M. Zavala, C. R. Landis, M. Mavrikakis, and G. W. Huber. Hydroformylation of pyrolysis oils to aldehydes and alcohols from polyolefin waste. *Science*, 381:660–666, 2023.
  - J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah. Julia: A fresh approach to numerical computing. *SIAM review*, 59(1):65–98, 2017.
  - Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual, 2022.
  - P. A. Tominac and V. M. Zavala. Economic properties of multi-product supply chains. *Computers & Chemical Engineering*, 145:107157, Feb. 2021.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Biofuels with Carbon Capture and Storage in the United States Transportation Sector

Caleb H. Geissler<sup>a,b</sup> and Christos T. Maravelias<sup>a,b,c,\*</sup>

<sup>a</sup> Princeton University, Department of Chemical and Biological Engineering, Princeton, New Jersey, USA

<sup>b</sup> DOE Great Lakes Bioenergy Research Center, USA

<sup>c</sup> Princeton University, Andlinger Center for Energy and the Environment, Princeton, New Jersey, USA

\* Corresponding Author: [maravelias@princeton.edu](mailto:maravelias@princeton.edu).

---

## ABSTRACT

There is a need to drastically reduce greenhouse gas emissions. While significant progress has been made in electrifying transport, heavy duty transportation and aviation are not likely to be capable of electrification in the near term, spurring significant research into biofuels. When coupled with carbon capture and storage, biofuels can achieve net-negative greenhouse gas emissions via many different conversion technologies such as fermentation, pyrolysis, or gasification to produce ethanol, gasoline, diesel, or jet fuel. However, each pathway has a different efficiency, capital and operating costs, and potential for carbon capture, making the optimal pathway dependent on policy and spatial factors. We use the Integrated Markal-EFOM System model applied to the USA, adding a rich suite of biofuel and carbon capture technologies, region-specific CO<sub>2</sub> transportation and injection costs, and government incentives from the Inflation Reduction Act. We find that under current government incentives, biofuels and carbon capture from biorefineries are primarily focused in the Midwest and South of the USA, but play a relatively small role in the overall USA transportation sector even in 2055. However, increased government incentives, biomass availability, or oil price could lead to increased biofuel production and reduced transportation emissions.

---

**Keywords:** Carbon Dioxide Sequestration, Biofuels, Biomass, Optimization, Energy Systems

## INTRODUCTION

Reducing greenhouse gas (GHG) emissions in the transportation sector is critical for reducing global warming. While improvements in electrification and electric vehicles are expected to play a critical role in reducing emissions from light duty vehicles, heavy duty transportation and aviation are not likely to be electrified in the near term [1]. In these sectors, the high energy density of fuel is important, so liquid fuels are expected to be essential for years to come [2]. Therefore, there has been significant research into renewable liquid fuels.

Renewable liquid fuels can be derived from a variety of sources. One of the most affordable sources has been waste lipids, which can be upgraded through the hydro-processed esters and fatty acids (HEFA) conversion pathway. However, the availability of waste lipids, and purpose-grown oil crops, is currently limited, and expected to remain low [3]. Meanwhile, cellulosic biomass

is widely available and can be converted into ethanol, gasoline, diesel, or jet fuel through a variety of biological and thermochemical pathways [4].

Biofuels have dramatically lower GHG emissions than their fossil fuel counterparts, but there are still emissions from the growth, harvesting, transportation, and conversion of the biomass. However, by coupling biofuel production with carbon capture and storage (CCS), net-negative GHG emissions can be achieved [5]. CCS incurs significant capital and operating expenses, so it is generally only economically viable with government incentives. With the passing of the Inflation Reduction Act (IRA), the USA has a tax credit of \$85/Mg CO<sub>2</sub> captured and stored securely. There is also a higher credit of \$180/Mg CO<sub>2</sub> for CO<sub>2</sub> captured directly from the atmosphere. The USA also provides incentives for the production of renewable liquid fuels, including ethanol, gasoline, diesel, and jet fuel. Notably, the credit for sustainable aviation fuel (SAF) depends on its emission reduction compared for fossil-



based jet fuel. SAF with 50% emission reduction receives a \$1.25/gallon credit, which increases by 1 cent for each additional percent of emission reduction, up to a maximum of \$1.75/gallon at net-zero emissions. This credit exceed that for ethanol, gasoline, and diesel at \$1/gallon.

Many techno-economic analysis have been performed on biofuels with carbon capture and storage (BECCS) systems and the impact of sequestration credits on their economic viability, but many of these studies pick a single conversion pathway and option for CCS. They also typically study a single biorefinery with a single feedstock price for a representative year [6]. However, biomass availability, locations for carbon sequestration, and competition with other fuels are key factors that can only be included with larger systems-level models.

Biofuels, with and without CCS, have featured heavily in integrated assessment models (IAMs). However, IAMs model economic, energy, and environmental sectors and interactions for the globe over long periods of time, so are often limited to a coarse representation of the energy sector [7]. Instead, detailed energy systems model can include a rich suite of conversion pathways.

Other studies of biofuels in energy systems models, and especially those focusing on the USA, have several crucial limitations. First, they typically simplify the wide range of combinations of biomass type, conversion pathway, and CCS to one or two conversion technologies. Second, they consider CCS from only some streams at the biorefinery even for pathways that generate multiple distinct CO<sub>2</sub> streams, such as those based on microbial conversion and gasification. Third, they often simplify the cost of CO<sub>2</sub> transportation and injection to a single value that is independent of region.

In this work, we present a study including a rich suite of biofuel conversion pathways to ethanol, gasoline, diesel, and jet fuel with the option for CCS from each CO<sub>2</sub> stream. We add these pathways to the Integrated MARKAL-EFOM System (TIMES) model applied to the USA. We further improve the model by adding transportation costs and region-specific injection cost curves for CO<sub>2</sub>. We also implement the most recent government incentives to study the role of biofuels and CCS in the USA transportation sector.

## METHODS

### Optimization Model

TIMES is a bottom-up long-term energy system linear optimization model that includes transportation, electricity generation, residential, commercial, and industrial sectors. It is solved considering 5-year timesteps from 2020 through 2055 with perfect foresight to minimize the discounted cost of meeting energy demands. It divides years into seasons, which are further divided into daily time slices. These times slices are used to consider both

seasonal and daily balances for energy demands [8].

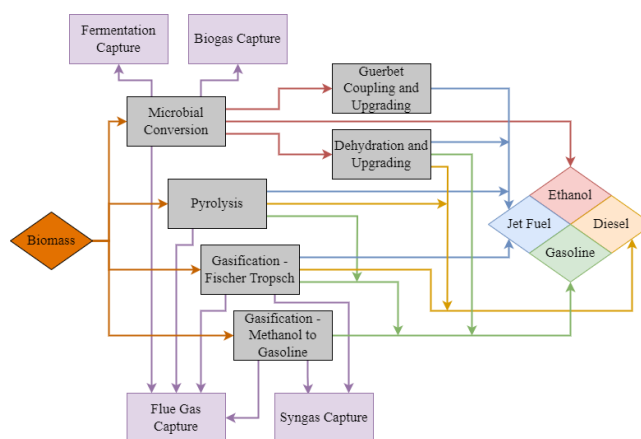
TIMES can be applied to different regions depending on the technologies available, and we use the EPAUS9rT database, which includes the existing energy technologies in the USA and their associated GHG emissions, divided into nine regions [9]. The inclusion of existing capacity and costs for conventional fuel technologies are essential for investigating how biofuel would compete in the current and future energy sector. However, because advanced biofuels are not currently widespread in the USA, the database has little representation for these fuels. Thus, this work adds these advanced biofuel technologies, as outlined in the following sections.

### Biomass Availability and Oil Price

The amount of biomass available and its cost is critical for studying biofuels. The Billion Ton report, updated in 2016, determined supply curves of biomass availability as a function of price for a variety of biomass types under different scenarios [10]. It generates these results assuming that sufficient land for food production is ensured before land is used for biomass production so that there is no competition with food. The study includes agricultural residues, grassy energy crops, corn stover, urban wood waste, forest resources, and woody energy crops. We note that utilizing waste biomass requires no additional land or water. For energy crops, additional water and land is required, but these are not tracked in TIMES and therefore beyond the scope of this study. We generate four scenarios of reference biomass availability and oil price, high biomass availability and reference oil price, reference biomass availability and high oil price, and high biomass availability and oil price.

### Biofuel Conversion Pathways

We include a wide variety of conversion technologies and options for CCS, summarized in Figure 1.



**Figure 1.** Biofuel production and carbon capture technologies considered in this study.

One mature pathway is microbial conversion, in

which fermentation is used to produce ethanol from sugars. In a biorefinery using this pathway, CO<sub>2</sub> is released from the fermentation process, in biogas from anaerobic digestion of wastewater, and in the flue gas from solid residue combustion [11]. Ethanol can be used as a fuel, or further upgraded. The upgrading options begin with Guerbet coupling to produce jet fuel, or with dehydration to produce a range of gasoline, diesel, and jet fuel [12].

Biomass can also be converted thermochemically. Pyrolysis produces a bio-oil that can be further upgraded into gasoline and diesel with hydrogen that can be produced onsite from the fuel gas. In this case, the only stream from which CO<sub>2</sub> can be captured is the dilute flue gas stream [13]. Alternatively, biomass can first undergo gasification to produce syngas. The syngas can then be converted into a range of gasoline, diesel, and jet fuel through Fischer-Tropsch synthesis [14], or to gasoline through a methanol intermediate [15]. For both pathways based on gasification, a very pure stream of CO<sub>2</sub> is produced when cleaning the syngas, and another dilute stream of CO<sub>2</sub> is produced from fuel gas combustion.

While thermochemical conversion technologies can process any type of biomass, the lower carbon content and higher ash content of herbaceous biomass increases costs. The nearly pure CO<sub>2</sub> streams from fermentation and syngas cleaning require only dehydration before compression, transportation, and sequestration. For all other CO<sub>2</sub> streams, we assume that amine absorption is used. We also assume that all the CO<sub>2</sub> in biogas is captured, while 85% of CO<sub>2</sub> in flue gas streams is captured.

### CO<sub>2</sub> Transportation and Sequestration

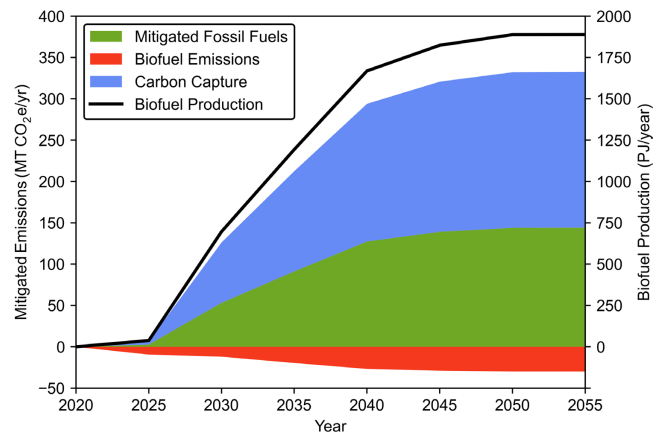
An important element in CO<sub>2</sub> sequestration not included in the EPAUS9rT database is the availability and injection cost of saline aquifers in each region. To add this to the model, we identified the potential injection sites with injection costs below \$100/Mg CO<sub>2</sub> in the literature. Assuming that each aquifer has a maximum annual injection limit of 1% of its total capacity, we constructed a cost curve of marginal cost of injection as a function of total CO<sub>2</sub> injection within each region. We calculated the CO<sub>2</sub> transportation costs within and between regions by the average distance between the potential locations of biorefineries and the potential injection sites.

## RESULTS

### Existing Credits

With the existing credits, we run the improved model through 2055. The results of biofuel production and the transportation emission reduction are shown in Figure 2. The total biofuel production increases rapidly beginning in 2030, almost entirely comprised of gasification and Fischer-Tropsch synthesis. Biofuel production continues to increase through 2050, but the growth slows as the all

the low-cost biomass sources are used and marginal cost of using additional biomass increases.

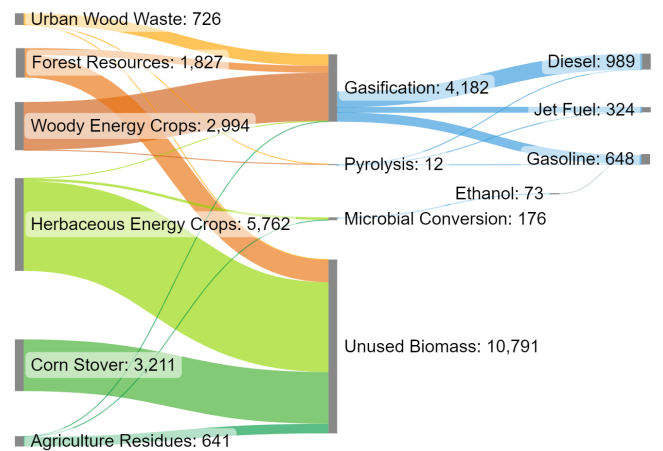


**Figure 2.** Biofuel production and corresponding GHG emission reduction over time.

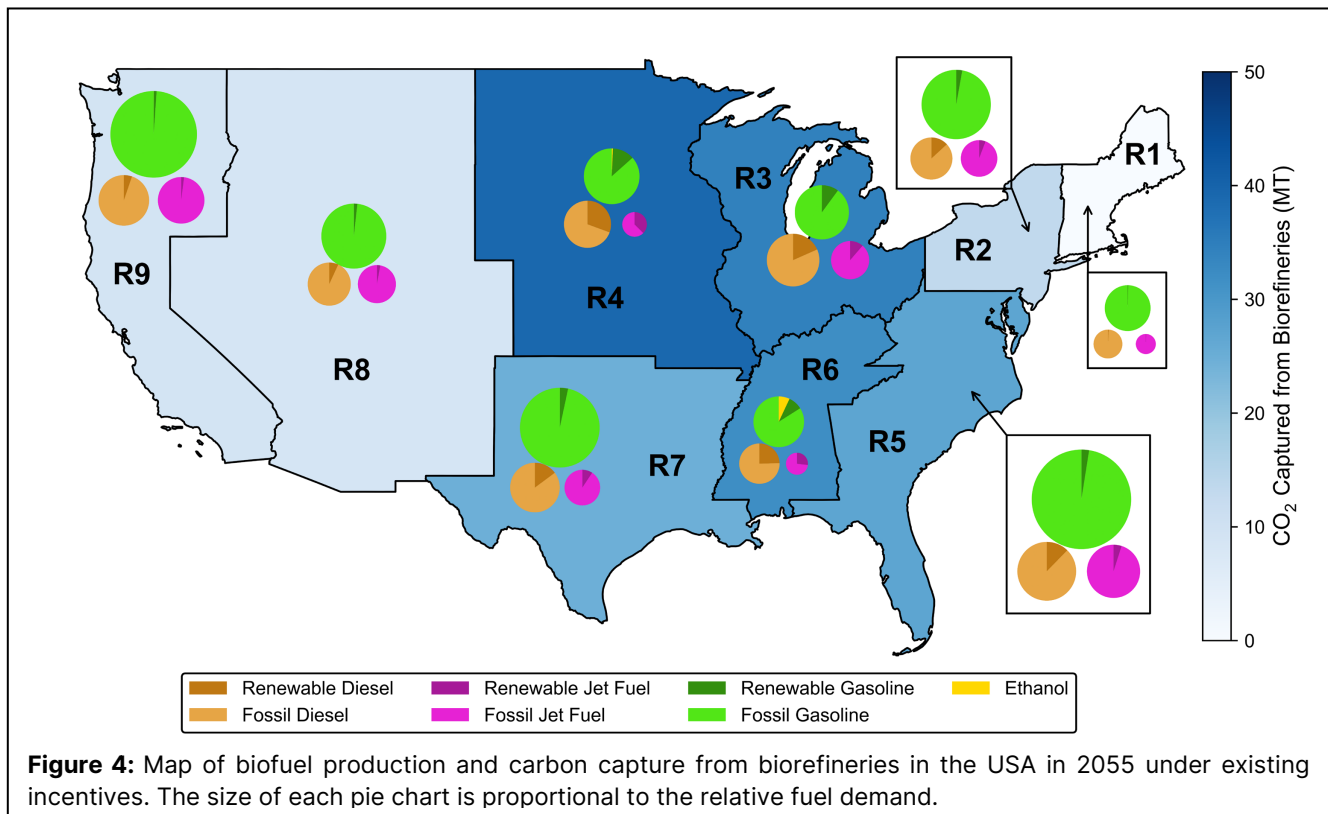
The amount of CCS varies by region depending on the cost of electricity and CO<sub>2</sub> transportation and injection, but on average 60% of biofuel production from Fischer-Tropsch includes CCS in both the stream from syngas cleaning and flue gas, while the remaining 40% includes capture from the syngas cleaning stream only.

Biofuels produce some emissions, but much fewer than fossil fuels. Accordingly, more than one third of emission reduction from BECCS come from the replacement of fossil fuels, while the remaining comes from the direct negative emissions of CCS.

Figure 3 shows the distribution of biomass sources and biofuel conversion pathways selected in 2055. Dedicated woody energy crops are the most used biomass sources, followed by urban wood waste and forest resources. These sources are selected due to their higher yields and lower costs than herbaceous sources.



**Figure 3.** Sankey diagram for energy flows of biomass and biofuels. Results are in petajoules (PJ) in 2055.



Very little herbaceous biomass is used and is limited primarily to producing small amounts of ethanol via microbial conversion. Overall, only a quarter of biomass available is used in 2055.

Figure 4 shows the regional distribution of biofuel production and CCS from biorefineries. Both biofuel production and CCS are focused in the Midwest and South (regions 3-7), which have lower costs and higher availability of both biomass and saline aquifer injection sites compared to the coastal regions. R4 has the highest biofuel production of any region at 445 PJ, and the highest amount of carbon captured at 50 MT/year. In New England (R1), there are no potential injection sites, so any carbon captured would have to be transported a significant distance, which would be expensive. Therefore, in that region a small amount of biomass is converted to fuel using pyrolysis without any CCS, which is the lowest-cost method of producing biofuels when CCS is not available or economically viable. The only region that continues to produce ethanol in 2055 is R6, which produces ethanol from grassy energy crops, instead of from corn as is currently the case in the USA. While there are options to upgrade ethanol to jet fuel via dehydration or Guerbet coupling, no ethanol upgrading pathways are selected due to their high costs compared to the thermochemical conversion pathways based on pyrolysis and gasification.

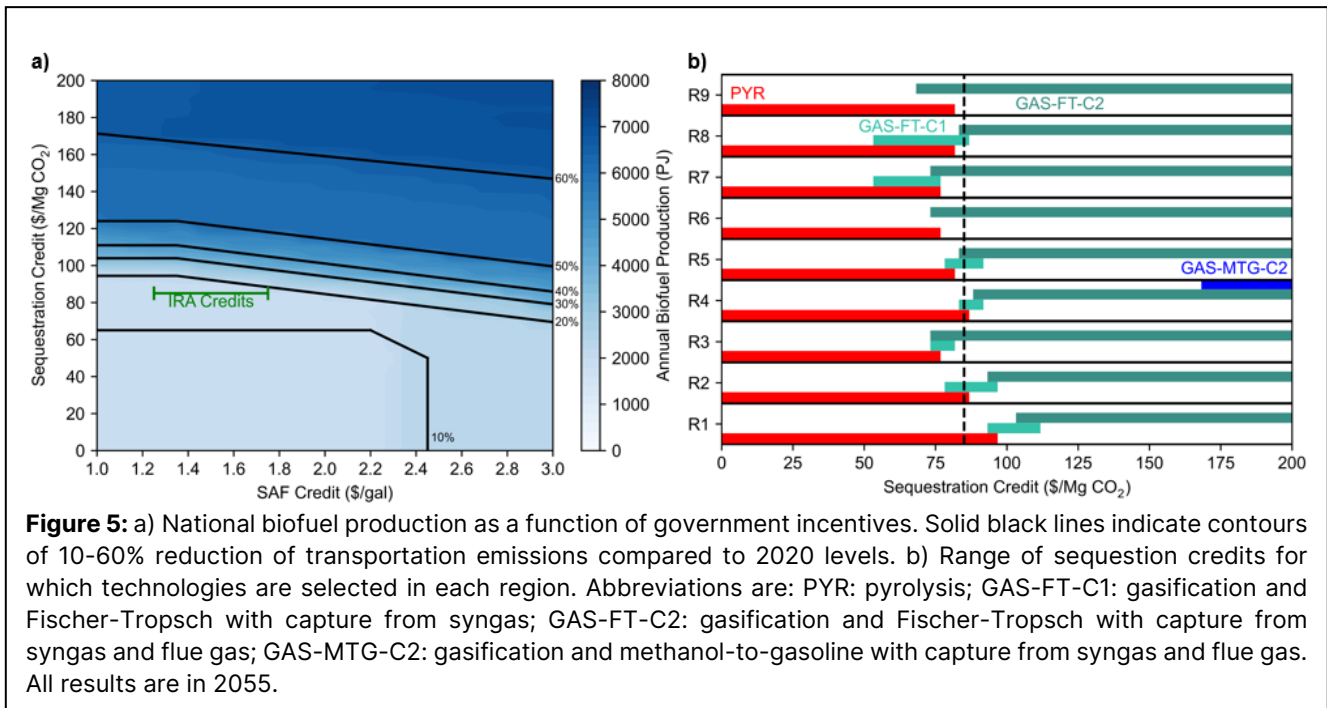
### Impact of changing credits

Relatively little biofuel production is incentivized

with existing credits, so we next explore how changing the value of the credits impacts the choice of technological pathway for biomass to biofuels, total biofuel production, and national transportation emission reduction.

Figure 5a shows that the sequestration credit is much more impactful on biofuel production and emission reduction than the SAF credit. In fact, at low sequestration credits, a SAF credit of over \$2.20 would be required to increase biofuel production by a significant amount. While at current incentives and assumptions in the EPAUS9rT database, annual transportation emissions were reduced by only 15% compared to 2020 levels, minor increases in the sequestration credit can dramatically reduce emissions, reaching 50% emission reduction at \$120/Mg CO<sub>2</sub> with current SAF credit levels.

Figure 5b shows the range of sequestration credits for which biofuel conversion pathways are selected. In all regions, pyrolysis is chosen at low credits, and this transitions to gasification and Fischer-Tropsch at higher credits, with the credit at which the transition occurs varying by region. The transition begins to occur at credits as low as \$55/Mg CO<sub>2</sub> in R7 and R8, where CCS from the syngas stream becomes incentivized. At credits above \$110/Mg CO<sub>2</sub>, gasification and Fischer-Tropsch synthesis with CCS from all sources preferred in all regions. Gasification and methanol-to-gasoline is incentivized in R4 at high sequestration credits because it has a higher ratio of CO<sub>2</sub> captured to fuel produced, allowing it to take greater advantage of the sequestration credit.



### Impact of biomass availability and oil price

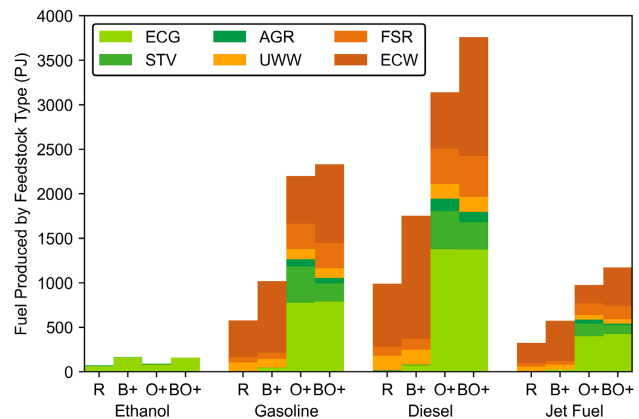
Even if sequestration and SAF credits do not change, there are other ways in which biofuel production might be increased. With continuing research in purpose-grown woody and herbaceous energy crops, and the potential of rising oil prices with depletion of reserves and environmental restrictions, biofuels may become economically attractive without increased government incentives. Accordingly, we show in Figure 6 biofuel production by different biomass sources under different scenarios of biomass availability and oil price.

With increasing biomass availability and oil price, the choice of biofuel conversion pathways does not change, though the magnitude of production changes significantly. The high biomass availability (B+) scenario has approximately double the biomass availability in 2055 compared to the reference (R) scenario, and biofuel production correspondingly doubles. The most significant increase between these scenarios is the increased use of woody energy crops, though there is also an increase in grassy energy crops undergoing gasification and Fischer-Tropsch synthesis and microbial conversion. However, R7 is the only region that uses grassy energy crops in this scenario.

When the oil price increases with reference biomass availability (scenario R to scenario O+), the most significant change is the dramatic increase in the use of herbaceous biomass, primarily in R3, R4, and R7. Herbaceous biomass has lower conversion efficiency and higher costs compared to woody biomass that makes it not competitive with fossil fuels in the reference scenario, but the increase in oil price is sufficient to change that. In

this scenario, R2 also begins to produce small amounts of ethanol from grassy energy crops and corn stover. Overall, an increase in the oil price alone increases total biofuel production by a factor of 4.

When an increase in biomass availability is also added on (scenario O+ to scenario BO+), there is less corn stover and significantly more woody energy crops available, which further increases biofuel production to nearly 7500 PJ. In the BO+ scenario, biofuel meets roughly 14% of gasoline, 55% of diesel, and 28% of jet fuel demand, resulting in 60% emission reduction compared to 2020 levels.



**Figure 6:** Biofuel production by biomass type under each biomass availability and oil price scenario in 2055. Abbreviations are: ECG: grassy energy crops; STV: corn stover; AGR: agriculture residues; UWW: urban wood waste; FSR: forest resources; ECW: woody energy crops.

## CONCLUSION

We improved the TIMES model applied to the USA energy sector by adding a rich suite of biofuel conversion pathways with the option for carbon capture from all CO<sub>2</sub> streams generated and region-specific CO<sub>2</sub> transportation and injection costs. With the improved model, we studied how the existing government credits in the Inflation Reduction Act would incentivize biofuel production and carbon capture, finding biofuels to meet only 15% of diesel, 4% of gasoline, and less than 1% of jet fuel demand in 2055. We explored changes in sequestration and SAF production credits and found the sequestration credit to be much more impactful, with a relatively minor increase to \$120/Mg CO<sub>2</sub> reducing emissions by 50%. Lastly, we examined how increases in biomass availability or oil price could increase biofuel production without and change in government incentives. Here, we found similar levels of emission reduction could potentially be achieved in the most biofuel-friendly scenario.

## ACKNOWLEDGEMENTS

This material is based upon work supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018409.

## REFERENCES

1. Oke D, Dunn JB, Hawkins TR. The contribution of biomass and waste resources to decarbonizing transportation and related energy and environmental effects. *Sustain Energy Fuels* 6:721-735 (2022)
2. Uludere Aragon NZ, Parker NC, VanLoocke A, Bagley J, Wang M, Georgescu M, Sustainable land use and viability of biojet fuels. *Nat Sustain* 6:158-68 (2023)
3. Quante G, Bullerdiel N, Bube S, Neuling U, Kaltschmitt M, Renewable fuel options for aviation – a system-wide comparison of drop-in and non drop-in fuel options. *Fuel* 333:126269 (2023)
4. Gelfand I, Sahajpal R, Zhang X, Izaurrealde RC, Gross KL, Robertson GP. Sustainable bioenergy production from marginal lands in the US Midwest. *Nature* 493:514-7 (2013)
5. Geissler CH, Maravelias CT, Economic, energetic, and environmental analysis of lignocellulosic biorefineries with carbon capture. *Appl Energy* 302:117539 (2021)
6. Geissler CH, Maravelias CT, Analysis of alternative bioenergy with carbon capture strategies: present and future. *Energy Environ Sci* 15:2679-89 (2022)
7. Daioglou V, Rose SK, Vauer N, Kitous A, Muratori M, Sano F, Fujimori S, Gidden MJ, Kato E, Keramidas K, Klein D, Leblanc F, Tsutsui J, Wise M, van Vuuren DP, Bioenergy technologies in long-run climate change mitigation: results from the EMF-33 study. *Clim Change* 160:3-20 (2020)
8. Loulou R, Goldstein G, Kanudia A, Lettila A, Remme U, Documentation for the TIMES model. <https://iea-etsap.org/index.php/documentation> (2016)
9. Lenox C, EPAUS9rT database for use with the TIMES modeling platform. [https://cfpub.epa.gov/si/si\\_public\\_record\\_report.cfm?Lab=NRMRL&dirEntryId=346478](https://cfpub.epa.gov/si/si_public_record_report.cfm?Lab=NRMRL&dirEntryId=346478) (2019)
10. Langholtz MH, Stokes BJ, Eaton LM, 2016 billion-ton report: advancing domestic resources for a thriving bioeconomy, volume 1: economic availability of feedstocks. *ORNL/TM-2016/160* (2016)
11. Humbird D, Davis R, Tao L, Kinchin C, Hsu D, Aden A, Schoen P, Lukas J, Olthof B, Worley M, Sexton D, Dudgeon D, Process design and economics for biochemical conversion of lignocellulosic biomass to ethanol: dilute-acid pretreatment and enzymatic hydrolysis of corn stover. *NREL/TP-5100-47764* (2011)
12. Restrepo-Flórez JM, Ryu J, Witkowski D, Rothamer DA, Maravelias CT, A systems level analysis of ethanol upgrading strategies to middle distillates. *Energy Environ Sci* 15:4376-4388 (2022)
13. Dutta A, Sahir A, Tan E, Humbird D, Snowden-Swan LJ, Meyer P, Ross J, Sexton D, Yap R, Lukas J, Process design and economics for the conversion of lignocellulosic biomass to hydrocarbon fuels. *NREL/TP-5100-62455* (2015)
14. Swanson RM, Platon A, Satrio JA, Brown RC, Hsu DD, Techno-economic analysis of biofuels production based on gasification. *NREL/TP-6A20-46587* (2010)
15. Tan ECD, Talmadge M, Dutta A, Hensley J, Schaidle J, Bidy M, Humbird D, Snowden-Swan LJ, Ross J, Sexton D, Yap R, Lukas J, Process design and economics for the conversion of lignocellulosic biomass to hydrocarbons via indirect liquefaction. *NREL/TP-5100-62402* (2015)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Sustainable Production of Fertilizers via Photosynthetic Recovery of Nutrients in Livestock Waste

Leonardo D. González<sup>a</sup>, Celeste Mills<sup>b</sup>, Aurora del C. Munguía-López<sup>a</sup>, and Victor M. Zavala<sup>a\*</sup>

<sup>a</sup> University of Wisconsin-Madison, Department of Chemical and Biological Engineering, Madison, WI, USA

<sup>b</sup> Texas A&M University, Artie McFerrin Department of Chemical Engineering, College Station, TX, USA

\* Corresponding Author: [zavatejeda@wisc.edu](mailto:zavatejeda@wisc.edu).

---

## ABSTRACT

Increases in population and improvements in living standards have significantly increased the demand for animal products worldwide. However, modern livestock agriculture exerts significant pressure on the environment due to high material and energy requirements. These systems also generate significant amounts of waste that can cause severe environmental damage when not handled properly. Thus, if we wish to enable farmers to meet this increased demand in a sustainable way, technology pathways must be developed to convert livestock agriculture into a more circular economy. With this end in mind, we propose a novel framework (which we call ReNuAI) for the recovery of nutrients from livestock waste. ReNuAI integrates existing technologies with a novel biotechnology approach that uses cyanobacteria (CB) as a multi-functional component for nutrient capture and balancing, purifying biogas, and capturing carbon. The CB can be applied to crops, reducing the need for synthetic fertilizers like diammonium phosphate. Using manure profiles obtained from dairy farms in the Upper Yahara region of Wisconsin, we construct a case study to analyze the environmental and economic impacts of ReNuAI. Our results illustrate that the minimum selling price (MSP) of CB fertilizer produced from deploying ReNuAI at a 1000 animal unit (AU) farm is significantly higher than the cost of synthetic fertilizers. We also observe that ReNuAI can return environmental benefits in areas such as climate change and nutrient runoff when compared to current practices. As a result, we see that consideration of environmental incentives can significantly increase the economic viability of the process.

---

**Keywords:** Process Design, Technoeconomic Analysis, Life Cycle Analysis

## INTRODUCTION

The UN Food and Agriculture Organization projects that demand for animal products will increase by up to 70% by 2050 [1]. This represents an opportunity for significant economic growth for areas that specialize in livestock agriculture like the Upper Yahara region of Wisconsin. However, meeting this demand also presents a significant sustainability challenge. Livestock agriculture is a resource-intensive venture: production of 1 kg of fat and protein corrected milk requires 2.7 MJ of energy, 2.61 kg of nitrogen (N), and 0.4 kg of phosphorus (P) [2, 3]. While nitrogen is a renewable resource, phosphorus and the required energy are usually sourced from non-renewable feedstocks [4]. Concurrently, these systems generate large amounts of waste, mostly in the form of animal

manure [5]. The most common method for managing this waste is to apply it to cropland where it can serve as a fertilizer. However, this approach leads to several undesirable environmental outcomes. Manure breaks down when exposed to sunlight and releases potent greenhouse gases (GHGs) like N<sub>2</sub>O and CH<sub>4</sub>; this practice has been measured to have a global warming potential (GWP) of 72.5 kg CO<sub>2</sub>-eq/tonne of manure. Additionally, because livestock waste has a low N:P ratio, it is often over-applied to meet crop N requirements; this results in the overloading of soils with P [6]. This, in turn, leads to excess nutrients being washed out into surrounding waterways where they cause harmful algal blooms (HABs). In addition to severe environmental damage to aquatic ecosystems, these eutrophication events also result in significant health and economic damage to the residents of

neighboring communities [7].

Technologies that provide alternative methods for handling livestock manure have existed for decades. The use of anaerobic digesters (ADs), which were developed in the early 20<sup>th</sup> century, in a farm setting was initially considered during the 1970s in response to an energy crisis [8]. ADs use microorganisms that feed on raw manure to generate a gas mixture, commonly referred to as biogas, largely composed of CO<sub>2</sub> and CH<sub>4</sub> that can serve as a renewable energy source. Additionally, by breaking down manure in a more controlled setting, ADs are able to prevent the release of GHG emissions seen in land application. However, the leftover material, known as digestate, has a similar N:P ratio as manure, and also leads to nutrient overloading when applied to cropland. Water recovery systems (WRS) offer a solution to this issue by providing further separation capabilities for the digestate. The central aim of a WRS is to recover clean water from manure. This is achieved using a solids-liquids separator (SLS) followed by ultrafiltration (UF) and reverse osmosis (RO) systems [9]. This allows for the separation of P and N nutrients. The solid product contains most of the P while the N is largely recovered in a concentrated solution following the UF and RO steps [10]. While this allows for variable nutrient loading, additional infrastructure and management protocols are required to ensure that these products are properly stored and applied [9].

Cyanobacteria (CB) are aquatic photosynthetic bacteria with high productivity compared to terrestrial plants [11, 12] that have been shown to be capable of processing nutrient-rich streams like manure or digestate [13, 14]. In fact, HABs are the result of the rapid growth of certain species of CB in nutrient-rich water. We propose harnessing this ability in a more controlled way to cultivate this organism for use as a biofertilizer. In addition to recycling nutrients and reducing nutrient leaching, CB fertilizer has been shown to improve soil health, reduce erosion, and has a smaller carbon footprint when compared to synthetic fertilizers [15-17]. To this end, we present the Renewable Nutrients from Algae (ReNuAl) process, a small and low-intensity operation for the production of CB fertilizer from manure. This technology pathway is centered around the use of CB strains that are engineered to have N and P content ratios that match crop nutrient requirements. We study the economic metrics of this process using a technoeconomic (TEA) model of ReNuAl at a hypothetical 1000 animal unit (AU) dairy farm in the Upper Yahara. The calculated minimum selling price (MSP) of the biofertilizer, while similar to the value reported in another study [18], is significantly higher than the cost of synthetic fertilizers. However, a preliminary life cycle assessment (LCA) of the process indicates that ReNuAl has the potential to deliver substantial environmental benefits when compared to direct land application. Thus, the consideration of environmental incentives

can reduce the fertilizer MSP significantly.

## COMPUTATIONAL FRAMEWORK

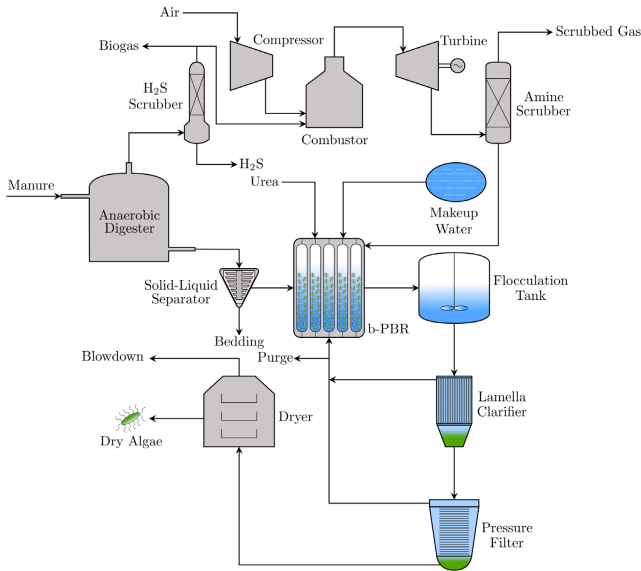
In this section we present the ReNuAl process, state the assumptions made in the process models, and provide a description of the proposed computational framework. As seen in Figure 1, the front end of ReNuAl pairs anaerobic digestion with CB cultivation and harvesting. This integration allows for the continued recovery of energy from manure via biogas in tandem with biofertilizer production, providing the process with multiples product streams. The CB are cultivated in bag photobioreactors (b-PBRs), and their growth is assumed to be light-limited. We assume that two thirds of the produced biogas is exported to the grid while the remaining third is burned for on-site power generation. The flue gas is scrubbed, and the CO<sub>2</sub> is fed to the cyanobacteria. Solids in the digestate are removed in a solids-liquids separation unit, and the liquid fraction is pumped into the b-PBRs along with any additional N required in the form of urea. Once the CB are ready to be harvested, they are sent to a dewatering train consisting of a flocculation tank, clarifier, and pressure filter that yields a concentrated CB solution. This stream is then fed to a thermal dryer to achieve the desired moisture level. CB growth in the b-PBRs is simulated using a detailed growth model based on the work of Straka [22] while the remaining units are simulated with linearized models using yield factors found in the literature. The ReNuAl process is fed 11.7 tonnes/yr of N and 10.6 tonnes/yr of P from 20800 tonnes/yr of manure. The b-PBRs receive 350 μmol/m<sup>2</sup>·s of sunlight and are harvested every 30 days. The P content of the CB is assumed to be 0.023 g P/g CB resulting in a production rate of 419 tonnes/yr of fertilizer. Using these unit operations models and system parameters, we calculate the mass and energy flows of the process and the required size of the process units. This information is then used to determine the capital investment as well the yearly operating costs and revenues of ReNuAl. Note that prices are set in terms of 2020 USD and the process is assumed to have a lifetime of 20 years.

The MSP is calculated by determining the price at which the fertilizer must be sold to achieve a discounted return on investment (DROI) of 15%. The DROI is defined as the discount rate that results in a net present value (NPV) of 0 at the end of the project life where:

$$NPV = C + \sum_{t=1}^T P(1+i)^{-t} \quad (1)$$

The DROI is denoted by  $i$ ,  $T$  is the project lifetime (in years),  $C$  is total capital investment (TCI) and  $P$  is the annual after-tax profit (AATP) and is formulated as:

$$P = (1-r)(p_f \dot{m}_f + p_g \dot{m}_g + p_e \dot{w} - p_o - d) + d \quad (2)$$



**Figure 1.** ReNuAl process configuration

where the production rates of fertilizer, biogas and electricity are represented by  $\dot{m}_f$ ,  $\dot{m}_g$ , and  $\dot{w}$  respectively and these are sold at  $p_f$ ,  $p_g$ , and  $p_w$ . The term  $r$  is the tax rate,  $p_o$  is the total annual operating cost (TOC), and  $d$  is the annual depreciation of the process equipment. Note that the values in (2) are held constant throughout the lifetime of the project. Thus, the MSP is defined as the value of  $p_f$  that results in a value 0 for (1) when  $i = 0.15$ .

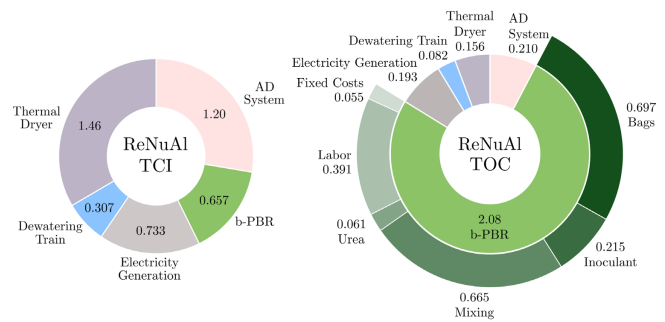
We use the material and energy flows calculated by the process model to develop a preliminary LCA to quantify the environmental benefits of ReNuAl over current practices. The LCA is performed using openLCA in conjunction with the Environmental Footprint and AGRIBALYSE databases and the Environmental Footprint impact assessment method [19, 20, 21]. The functional unit is set to be the size of the dairy farm (1000 AU), and we consider that the electricity demands are satisfied using the average electricity mix in the US. The impact categories considered in this study are climate change and water use. We valorize the observed benefits using existing frameworks, such as RIN and LCFS credits and a hypothetical P credit based on the economic impact of P runoff as quantified in [7], in order to determine the effect of these additional cash flows on the product MSP.

## RESULTS AND DISCUSSION

We determined that the construction and operation of the ReNuAl process have a cost of 4.36 MMUSD and 2.73 MMUSD/yr respectively. At these values, the fertilizer would have to be sold for 7.75 USD/kg to meet the DROI target. This is significantly higher than the 0.25 USD/kg that it would cost to obtain the equivalent N and P content using synthetic fertilizers.

From Figure 2, we can see that TCI and TOC are not

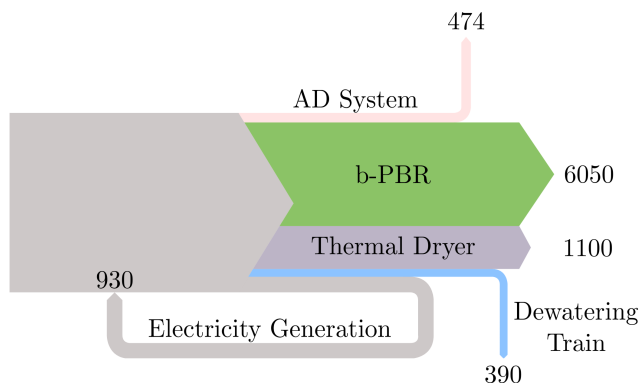
evenly distributed across the various sections of the process. The anaerobic digester and thermal dryer alone account for over 60% of the TCI while the operation of the b-PBRs represents about 75% of the TOC. This indicates that by exploring alternative process configurations or extending the capabilities of the CB, we might be able to significantly decrease these costs and, by extension, the MSP. For example, a CB strain that can clean the biogas by removing  $H_2S$  and  $CO_2$  can reduce the capital cost by 11% by eliminating the need for gas scrubbers. Similarly, if it is not necessary to completely dry the CB or if strains that are more easily recoverable via sedimentation and filtration are used, then it might be possible to reduce the size of the thermal dryer or eliminate it completely.



**Figure 2.** Distribution of total capital (left) and operating (right) costs in MMUSD across the sections of the ReNuAl process

While the economic metrics of ReNuAl seem unfavorable, it should be noted that there is a significant degree of uncertainty surrounding the values of the CB cultivation and harvesting sections; the design of these systems has traditionally focused on biofuels applications. As a result, the scale and complexity of the equipment used is significantly higher than what should realistically be required for ReNuAl; this is especially true of the bioreactors. From Figure 3, we can determine that operation of the b-PBRs accounts for approximately 75% of the energy consumed by the process. This is largely driven by the mixing requirements of the reactors, which alone represent 24% of the TOC. This level of energy demand is reasonable at a biofuel production facility given the large volume b-PBRs must fit to accommodate the high production rates required. However, ReNuAl will use significantly smaller reactors, and, as a result, these will likely require less energy to be well-mixed. Similarly, the labor, maintenance, and replacement costs of the b-PBRs, which together account for 40% of the TOC, would likely be lower than the values we calculated for the same reason.

The preliminary environmental analysis shows that the water use impact is higher for ReNuAl than for current practices. This is unsurprising as CB cultivation operations have high water requirements due to the achievable



**Figure 3.** Yearly energy flows (in MW-hr) for the different sections of the ReNuAI process

titers being fairly low (approximately 1-2g/L). Manure spreading, on the other hand, does not require any additional water as manure has a moisture content of approximately 90%, and, after the removal of suspended solids, it can be easily pumped through a liquids distribution system. However, we also observe that ReNuAI generates 39% fewer GHG emissions than land application of manure. A significant portion of this reduction is due to the production of biogas. Under our framework, ReNuAI is able to export 429 tonnes/yr of methane which make it eligible to receive compensation from government-funded incentives like RIN credits from the EPA's renewable fuel standards program and low carbon fuel standard (LCFS) credits offered by the state of California. Applying the value of the RIN and LCFS credits, we determine that the process is able to generate an additional 600,000 USD/yr of revenue; this lowers the MSP of the CB fertilizer 18% to 6.33 USD/kg.

When we consider the release of nutrients to the environment from the land application of livestock waste with those from ReNuAI, we observe that our process has two distinct advantages. First, the N:P ratio of the biofertilizer more closely matches crop needs than manure, reducing the need for overapplication. Second, CB is a more stable medium and releases nutrients more gradually than manure [15, 16]; this significantly reduces the amount of nutrients carried away during rain events. The HABs caused by nutrient runoff result in the affected waterbodies not being usable for economic activities like fishing and recreation and, as a result, represent a loss for the surrounding communities. Previous studies focusing on the Upper Yahara have estimated these losses to have a cost of 74.5 USD/kg P [7]. Thus, if we apply an equivalent credit, we determine that we can generate an additional 789,700 USD/yr if a CB strain that is able to capture all of the P in the manure can be engineered. If we combine this with the revenue from the RIN and LCFS credits, the MSP of the CB can be reduced further to 4.45 USD/kg. It is important to note that this represents a best-case scenario and additional work must be done to

determine nutrient uptake and leaching rates for the CB.

In addition to highlighting potential knowledge gaps as previously discussed, our framework allows us to understand the sensitivity of process economics to changes in various key parameters. For this analysis we considered the reactor batch time, the fraction of biogas sent to market, the reactor surface area to volume (SA:V) ratio, the light available to the bioreactors, and the P uptake of the CB. Note that this is not an exhaustive selection; rather, these choices are based on the variables we can meaningfully simulate using the current process model.

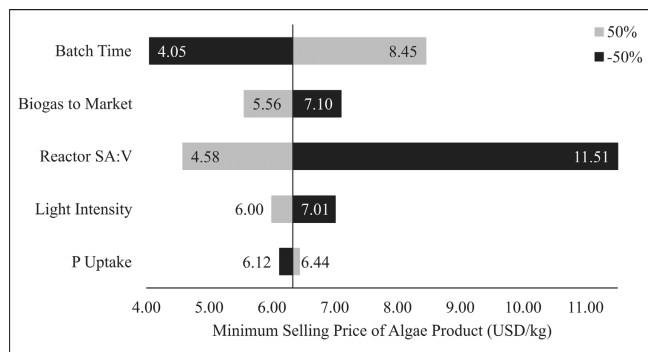
The results shown in Figure 4 (note we are including RIN and LCFS credits in the calculation of the MSP) indicate that the parameters associated with the design and operation of the b-PBRs, batch time and reactor SA:V ratio, have the largest impact on MSP. Given the cost of operating the reactors, these results are not surprising as these parameters directly affect the amount of reactor volume required.

After the batch time and SA:V ratio, the next most important variable to consider is the fraction of biogas sent to market. Due to the low value of electricity compared to the value of biogas and the cost of the electricity generation section, we observe a 13% decrease in the MSP when we export all of the biogas offsite. However, we should note that a significant portion of the revenue from the biogas comes from RIN and LCFS credits, and changes in policy can reduce the value of these credits.

From the variation in the light intensity, we can conclude that the process is not severely light-limited as shifting light availability up or down by 50% only moves the MSP -5% to +11% from the baseline. This indicates that in the current design it is the reactor surface area to volume ratio that limits the titer that can be achieved. This is due to shading effects which limit the distance that light can penetrate the vessel. As a result, the reactor design should be optimized before exploring methods for increasing light availability.

The results seem to indicate that lowering the P content of the CB might be favorable as this lowers the MSP. However, in this context, it is important to think about the value of CB which is dependent on its nutrient content. The reason the MSP decreases at lower P uptake values is because more of it must be produced to consume all of the available nutrients. As a result, on a nutrient content basis, the biomass is less valuable. A more appropriate question is then if this drop in value is matched by the drop in MSP. From Figure 4, we observe that the answer is no. Conversely, we also see that when the P content of the cells is increased, the MSP increases at a lower rate. Additionally, this decrease in throughput translates to a lower TCI and TOC. Thus, it might be possible to reverse this trend at higher P loading values.





**Figure 4.** Change in MSP of CB product with changes to selected operating parameter

## CONCLUSIONS AND FUTURE WORK

This study presented a framework for the capture and recycling of nutrients in livestock waste. The goal of this process, which we refer to as ReNuAl, is to provide a pathway towards a more circular economy that will allow livestock farmers to meet the projected increase in demand for animal products in a sustainable manner. An initial economic analysis of ReNuAl indicates that the produced biofertilizer is significantly more expensive than current synthetic fertilizers. However, we also determined that there is a significant degree of uncertainty around the economic values calculated for the b-PBR section due to the difference in scales that existing CB cultivation facilities operate at compared to ReNuAl. Thus, we are interested in developing prototypes of the b-PBRs to better understand their operations and obtain more accurate performance metrics.

Our environmental analysis indicates that ReNuAl has the capability to reduce carbon and nutrient-based pollution. However, we also observe that CB cultivation has a high water requirement. As result, identifying methods that allow for high degrees of water reuse and recovery as well as locations that are not water-strained for deployment are important factors to consider moving forward. It is also important to note that these results are preliminary. Moving forward we will focus on improving the accuracy of these values by working with microbial, agronomy, and environmental science experts who can perform field trials to measure variables of interest like maximum attainable titer, nutrient uptake and leaching, and GHG emissions. Additionally, we will also extend our analysis by evaluating other environmental metrics and performing a sensitivity analysis on the energy sources used to power the process as previous studies have shown that changing the energy source to non-fossil alternatives can decrease the climate change impact [23].

Using our TEA model, we were able to identify key operating parameters that can have a significant impact on the viability of the ReNuAl process. Our results indicate that optimization of b-PBR design and operation,

specifically vessel geometry and batch times, can significantly reduce the MSP of the biofertilizer. We also observe that by increasing P-uptake we can make a more nutrient-dense product that can be sold at a lower price (on a nutrient content basis) while also reducing capital and operating expenses. From these results we can conclude that developing fast-growing CB strains that have a high P content using a reactor with as high an areal mass density (high SA:V value) as possible will significantly improve the economics of ReNuAl. Additionally, this work demonstrates how a computational framework serves as a goal-oriented tool for determining promising directions that researchers can proceed in. Moving forward we are interested in exploring the impact of additional CB-specific traits that can potentially improve the economic and environmental performance of ReNuAl such as maximum harvest titer and settling velocity as well as the capacity of the organism to remove CO<sub>2</sub> and H<sub>2</sub>S from the biogas stream; we plan to achieve this by making use of new experimental data to refine our existing process models. In other words, our future work centers around the integration of process and economic models with experiments to allow for continuous model improvement which in turn allows for the accurate identification of key variables and regions of diminishing returns.

## ACKNOWLEDGEMENTS

We acknowledge support from NSF-EFRI #2132036, the Advanced Opportunity Fellowship from the UW-Madison Graduate Engineering Research Scholars program, and the PPG Fellowship.

## REFERENCES

1. Makkar HPS. Review: Feed demand landscape and implications of food-not feed strategy for food security and climate change. *Animal* 12:1744-1754 (2018)
2. Kim D, Stoddart N, Rotz CA, Veltman L, Chase J, Ingraham P, Izaurralde RC, Jones CD, Gaillard R, Aguirre-Villegas HA, Larson RA, Ruark M, Salas W, Joillet O, Thoma GJ. Analysis of beneficial management practices to mitigate environmental impacts in dairy production systems around the Great Lakes. *Agricultural Systems* 176:102660 (2019)
3. Leytem AB, Williams P, Zuidema S, Martinez A, Chong YL, Vincent A, Vincent A, Cronan D, Kliskey A, Wulfhorst JD, Alessa L, Bjorneberg D. Cycling phosphorus and nitrogen through cropping systems in an intensive dairy production region. *Agronomy* 11:1005 (2021)
4. Cordell D, Drangert JO, White S. The story of



- phosphorus: Global food security and food for thought. *Global Environmental Change* 19:292-305 (2009)
5. Larson RA, Sharara M, Good LW, Porter P, Runge R, Zavala V, Sampat A, Smith A. *Evaluation of Manure Storage Capital Projects in the Yahara River Watershed*. University of Wisconsin-Extension, University of Wisconsin-Madison College of Agricultural and Life Sciences, Biological Systems Engineering technical report (2016)
  6. Wang H, Aguirre-Villegas HA, Larason RA, Alkan-Ozkaynak. Physical properties of dairy manure pre- and post-anaerobic digestion. *Applied Sciences* 9:2703 (2019)
  7. Sampat AM, Hicks A, Ruiz-Mercado GJ, Zavala VM. Valuing economic impact reductions of nutrient pollution from livestock waste. *Resources, Conservation and Recycling* 164:105199 (2021)
  8. Lusk PD. *Methane recovery from animal manures: The current opportunities casebook*. National Renewable Energy Laboratory, Resource Development Associates technical report (1998)
  9. Larson RA, Aguirre-Villegas H. *Treating Manure to Produce Clean Water*. University of Wisconsin-Extension, University of Wisconsin-Madison Nelson Institute for Environmental Studies technical report (2022)
  10. Ledda C, Schievano A, Salati S, Adani F. Nitrogen and water recovery from animal slurries by a new integrated ultrafiltration, reverse osmosis and cold stripping process: A case study. *Water Research* 47:6157-6166 (2013)
  11. Biller P, Ross A. Potential yields and properties of oil from the hydrothermal liquefaction of microalgae with different biochemical content. *Bioresource Technology* 102:215-225 (2011)
  12. Knoot CJ, Ungerer J, Wangikar PP, Pakrasi HB. Cyanobacteria: Promising biocatalysts for sustainable chemical production. *Biological Chemistry* 293:5044-5052 (2018)
  13. Wang L, Li Y, Chen P, Min M, Chen Y, Zhu J, Ruan RR. Anaerobic digested dairy manure as a nutrient supplement for cultivation of oil-rich green microalgae *Chlorella* sp. *Bioresource Technology* 101:2623-2628 (2010)
  14. Álvarez X, Arévalo O, Salvador M, Mercado I, Velázquez-Martí B. Cyanobacterial biomass produced in the wastewater of the dairy industry and its evaluation in anaerobic co-digestion with cattle manure for enhanced methane production. *Processes* 8:1290 (2020)
  15. Mulbry W, Westhead EK, Pizarro C, Sikora L. Recycling of manure nutrients: Use of algal biomass from dairy manure treatment as slow-release fertilizer. *Bioresource Technology* 96:451-458 (2005)
  16. Coppens J, Grunert O, Van Den Hende S, Vanhoutte I, Boon N, Haesaert G, De Gelder L. The use of microalgae as a high-value organic slow-release fertilizer results in tomatoes with increase carotenoid and sugar levels. *Applied Phycology* 28:2367-2377 (2016)
  17. Goemann HM, Gay JD, Mueller RC, Brookshire ENJ, Miller P, Poulter B, Peyton BM. Aboveground and belowground responses to cyanobacterial biofertilizer supplement in a semi-arid, perennial bioenergy cropping system. *GCB-Bioenergy* 13:1908-1923 (2021)
  18. Lee JC, Lee B, Kim HW, Jeon BH, Lim H. Techno-economic analysis of livestock urine and manure as a microalgal growth medium. *Waste Management* 135:276-286 (2021)
  19. Ciroth A, Di Noi C, Lohse T, Srocka M. "openLCA 1.10 Comprehensive user manual." GreenDelta, Berlin, Germany (2020)
  20. Fazio S, Castellani V, Sala S, Schau E, Secchi M, Zampori L, Diaconu E. "Supporting information to the characterisation factors of recommended EF Life Cycle Impact Assessment methods: new methods and differences with ILCD." (2018)
  21. Colomb V, Ait-Amar S, Basset-Mens C, Gac A, Gaillard G, Koch P, Mousset J, Salou T, Tailleur A, Van Der Werf HM. "AGRIBALYSE®, the French LCI Database for agricultural products: high quality data for producers and environmental labelling." (2015)
  22. Straka L. Light-dependent growth kinetics and mathematical modeling of *Syncehocystis* sp. PCC 6803. *Ph.D. thesis* Arizona State University (2017)
  23. Munguia-Lopez AC, Göreke D, Sánchez-Rivera KL, Aguirre-Villegas HA, Avraamidou S, Huber GW, Zavala VM. "Quantifying the environmental benefits of a solvent-based separation process for multilayer plastic films." *Green Chem.* 25(4): 1611-1625. (2023)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Techno-Economic Analysis of Methane Production from Pulp and Paper Sludge

Erfan Hosseini<sup>a</sup>, Selen Cremaschi<sup>a\*</sup>, and Zhihua Jiang<sup>ab</sup>

<sup>a</sup> Department of Chemical Engineering, Auburn University, Auburn, AL, US

<sup>b</sup> Alabama Center for Paper and Bioresource Engineering, Auburn University, Auburn, AL, US

\* Corresponding Author: [szc0113@auburn.edu](mailto:szc0113@auburn.edu).

---

## ABSTRACT

This study investigates the feasibility of valorizing pulp and pulp sludge (PPS) into methane through anaerobic digestion (AD) with a focus on techno-economic analysis (TEA). Three scenarios are evaluated: (A) the base case, (B) sludge AD with alkaline pretreatment using green liquor dregs (GLD), and (C) co-digestion with nitrogen-rich feedstocks. The evaluation is applied to a common PPS, consisting of 70% primary sludge (PS) from the primary clarifier and 30% secondary sludge (SS) from biological treatments from a kraft mill. Theoretical methane potential (TMP) is determined using the Buswell equation. The study highlights the significance of co-digestion with nitrogen-rich feedstocks in enhancing the economic viability of the AD process for PPS, providing valuable insights for sustainable waste management and resource recovery in the pulp and paper industries.

---

**Keywords:** Pulp and paper sludge, biomethane, anaerobic digestion, techno-economic analysis, valorization.

## INTRODUCTION

The pulp and paper industry (PPI) generates a significant volume of wastewater that undergoes physical and biological treatment [1,2], resulting in a considerable quantity of pulp and paper sludge (PPS) [3], which is one of the major waste streams of PPI [4]. As global energy demand escalates and concerns over energy security and climate change intensify, anaerobic digestion (AD), which transforms organic materials into methane and carbon dioxide in the absence of oxygen [5], has emerged as a versatile technology for renewable energy production. This microbial-mediated process presents a promising solution for managing PPS, offering an alternative energy source to supplement industrial fossil fuels. PPS contains 45-55% organic matter and various nutrients, including nitrogen and phosphorus [6], making it a potential resource for microorganisms in the production of bio-products and biofuels. The conversion of pulp and paper mill by-products, such as sludge, into value-added goods and bioenergy can significantly contribute to the commercial growth of biorefineries. However, for effective utilization, these by-products must meet specific minimum standards. To enhance efficiency and

profitability, there is a growing interest in evaluating opportunities to generate new revenues from innovative, value-added products within the existing infrastructure of pulp and paper mills. Consequently, this study aims to identify a promising process pathway for a biorefinery integrated with an existing pulp and paper mill, considering various scenarios in pursuit of improved efficiency and profitability.

## METHODOLOGY OVERVIEW

### Process synthesis and design

The techno-economic analysis (TEA) of AD for PPS considers three distinct scenarios (Fig. 1), each representing a unique process aimed at evaluating the economic and technical feasibility of methane production. The sludge mixture constituted 70% PS from the primary clarifier and 30% SS obtained from biological treatments. The sludge generation varies widely among mills. In this context, the input mass flow for the sludge is assumed to be 500 tons per day with a dry matter content of 12%. The dry matter content and mass flowrate specified conform to data obtained from a United States Environmental Protection Agency (EPA) investigation encompassing

104 bleached Kraft mills and are consistent with established industry norms [7,8].

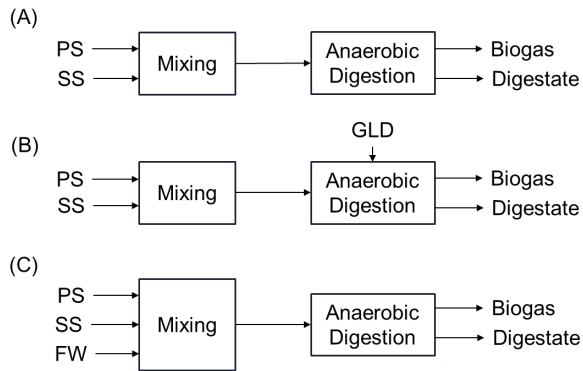


Figure 1. Three TEA scenarios for AD of mixed kraft mill primary and secondary sludge.

### Scenario A: Base case

The base case scenario mirrors the conventional operational conditions of AD. As a baseline condition, scenario A serves as a reference point for assessing the shortcomings and limitations of the existing process in terms of economic viability and methane production efficiency.

### Scenario B: Sludge AD with alkaline pretreatment

Among the various pretreatment methods, alkaline pretreatment is a promising approach that can successfully improve the enzymatic hydrolysis for many lignocellulosic materials [9,10].

As reported in [9], alkaline pretreatment effectively disrupted the floc structure of pulp and paper sludge, leading to a reduction in fiber size. According to reference [9], effective biodegradation in bioreactors was evidence in terms of soluble chemical oxygen demand (SCOD) removal efficiency, which attained the range of 83–93% in bioreactors with pretreated PPS compared with 70% removal for untreated PPS, indicating a significant improvement of 18.6–32.8% in organic removal. These findings show the efficacy of alkali pretreatment as a promising method for enhancing methane yield in the context of PPS.

An alkaline pretreatment typically involves the use of hydroxides like NaOH or KOH [11]. However, it has been suggested that the GLD, a by-product of the kraft pulping process, could be used as a substitute for powerful and expensive alkaline agents like NaOH [12].

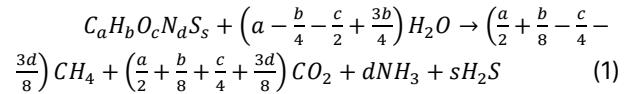
### Scenario C: Co-digestion with nitrogen-rich feedstocks

In this scenario, food waste (FW) is specified as a promising nitrogen-rich feedstock for co-digestion with PPS, supported by existing studies on the co-digestion of FW with PPS [13]. Previous work has mentioned nitrogen

deficiency as a major challenge in AD of PPS [14,15]. The ideal carbon-to-nitrogen ratio for AD feedstock typically falls within the 20–30% range [16]. A high C:N ratio, as observed in PPS, accelerates nitrogen consumption by methanogens, adversely affecting microbial population growth and prolonging the carbon digestion process. Conversely, a low C:N ratio leads to elevated ammonia release in the digester and inhibits the AD process.

## Theoretical methane potential

Theoretical methane potential was assessed using the Buswell equation (Eq. (1)), which relies on the stoichiometric balance between biodegradable organic matter and the resulting gaseous products from anaerobic biodegradation [17]. Volatile solids (VS) represent the organic matter content of the sludge that is readily converted to gas during the anaerobic digestion process, and total solids (TS) refer to the total mass of both organic and inorganic matter present in the sludge. These key parameters, VS and TS, play a significant role in determining the efficiency of anaerobic digestion and the potential methane yield. In the specific case of PPS, the Buswell equation can be applied, assuming a 30% volatile solids (VS) removal based on experimental findings [14]. In this context, biodegradation efficiency signifies the proportion of VS degraded during the process.



We can deduce the maximum theoretical methane potential (TMP) using Eq. (2):

$$TMP \left( \frac{m^3}{kgVS} \right) = \left( \frac{(4a+b-2c-3d) \times 22.415}{12a+b+16c+14d} \right) \quad (2)$$

## Economic analysis

### Gross economic potential (GEP)

The gross economic potential is used to assess the economic benefits and potential value generated by the process. The GEP is defined in Eq. (3).

$$GEP = VP + STF - VF \quad (3)$$

where VP is the value of products, STF is savings on tipping fee, and VF is the value of feeds. The value of products refers to the economic value generated from the products resulting from the AD process. In the case of converting sludge to methane, VP would include the revenue or value obtained from selling the methane produced, as well as any other by-products that can be monetized, such as organic fertilizers or other valuable substances extracted from the process. Tipping fees are charges imposed for disposing of waste in landfills. By diverting sludge from landfills and converting it into biogas through AD, the process effectively reduces or

eliminates the need to dispose of the waste in landfills. This saving is a significant factor in the economic assessment, especially given the specific cost of the tipping fee (\$58.47/ton [18]). Value of feeds represents the value of any feeds or materials used in the process. In the context of PPS converted to biogas, the VF is considered zero because PPS is a by-product with no value as a feed or input material.

### Manufacturing cost

The sizing and cost estimation of the pretreatment reactor followed the methodology proposed by Ulrich et al. [19]. However, in the case of the digester, its large size exceeds the range covered by the graphs outlined. Therefore, for the digester, we referred to a similar farm digester documented in the literature as a case study to estimate costs [20]. Capital cost of this case study is scaled to the capacity of AD systems assumed in our scenarios based on the usual 0.6-power rule (Eq. (4)):

$$CC_A = CC_B \times \left(\frac{Cap_A}{Cap_B}\right)^{0.6} \quad (4)$$

In Equation 4,  $CC_A$  and  $CC_B$  represent the capital costs of equipment A and B, respectively, while  $Cap_A$  and  $Cap_B$  denote the capacities of equipment A and B. All cost calculations are based on the chemical engineering plant cost index (CEPCI) of 816 (for 2022). Finally, assuming 18% of the bare module costs for contingency costs and fees based on reference [21], the total module cost ( $C_{TM}$ ) is calculated as Eq. (5) where  $n$  represents the total number of pieces of equipment and  $C_{BM,i}$  is the bare module cost for each piece of equipment  $i$ :

$$C_{TM} = 1.18 \times \sum_{i=1}^n C_{BM,i} \quad (5)$$

In this work, the CAPCOST method was applied to estimate the manufacturing costs [21]. Following the methodology proposed by Ulrich [19], the typical labor requirement for a continuous reactor is estimated at 0.3 workers per unit per shift. The operating labor cost ( $C_{OL}$ ) estimation assumptions considered are:

- On average, a worker at this plant operates five shifts per week for 52 weeks annually.
- The plant operates 365 days a year with three shifts per day.
- The 2022 Mean Annual Wage for Chemical plant and system operators is reported as \$79,290 per year, according to the U.S. Bureau of Labor Statistics [22].

The equation used to evaluate the cost of manufacture (COM) is:

$$COM = \text{Direct Manufacturing Costs} + \text{Fixed Manufact Costs} + \text{General Expenses} \quad (6)$$

Each individual cost item can be estimated

considering the costs of utilities ( $C_{UT}$ ), waste treatment ( $C_{WT}$ ), raw materials ( $C_{RM}$ ) and fixed capital investment (FCI). Turton et al. provide typical ranges for constants (multiplication factors) (Table 8.2. of [21]) to estimate these individual cost items. Since no other information is accessible regarding these costs in our study, we utilize the midpoint value within each range. Depreciation allowance is added separately to compute the cost of manufacturing (COM) using Eq. (7).

$$COM = 0.28 \times FCI + 2.76 \times C_{OL} + 1.23 \times (C_{UT} + C_{WT} + C_{RM}) \quad (7)$$

In Eq. (7), the FCI cost equals the total module cost, given that we are making alterations to an existing facility. In the specific context of our preliminary feasibility study or conceptual design, we adopt a simplified model or scenario wherein the costs of utilities ( $C_{UT}$ ), waste treatment ( $C_{WT}$ ), and raw materials ( $C_{RM}$ ) are ignored. However, it is essential to acknowledge that in real-world scenarios, these costs would typically be significant factors contributing to the overall manufacturing expenses.

## RESULTS AND DISCUSSION

### Buswell method

Based on the results of the elemental composition analysis for sludge derived from Lopes et al. [23] (refer to table 1), we can calculate the theoretical volume of methane using Eq. (2).

**Table 1.** Elemental compositions of the primary sludge (PS) and secondary sludge (SS) and mixed sludge (7:3) [23]

Parameters	PS	SS	Mix
VS/TS (g/g)	0.99	0.85	0.97
C (% TS)	44.10	45.20	44.41
H (% TS)	6.04	5.83	5.98
O (% TS)	48.80	29.80	34.86
N (% TS)	0.06	4.85	1.43
S (% TS)	0.40	1.82	0.81
Ash(%TS)	0.60	12.50	12.51
Total	100	100	100

VS: Volatile Solids; TS: Total Solids

According to the data presented in Table 2 calculated using Eqs. (1) and (2), the TMP for the mixed primary and secondary sludge, with a ratio of 7:3, was approximately 130.95 mlCH<sub>4</sub>/gVS<sub>fed</sub>, assuming a 30% removal of volatile solids (VS). This aligns with experimental findings documented in the literature [14,24]. This



theoretical value provides a preliminary insight into biogas and methane production.

**Table 2.** Biogas product yield and composition from the Buswell equation

Biogas	Yield (ml/gVS <sub>fed</sub> )	Composition (%)
CH <sub>4</sub>	130.95	49.12
CO <sub>2</sub>	126.41	47.42
NH <sub>3</sub>	7.43	2.79
H <sub>2</sub> S	1.79	0.67
Biogas	265.63	100

## Techno-economic analysis

### GEP of the different scenarios

For the base case scenario, the methane yield amounts to 130.95 mlCH<sub>4</sub>/gVS<sub>fed</sub>, equivalent to 269,141.78 ft<sup>3</sup>CH<sub>4</sub>/day and 17.46 tons PPS conversion per day. In a simplified context, assuming full upgrading and considering the lower heating value of methane at 910 Btu/ft<sup>3</sup> [25], the resulting energy production would be approximately 244.92 million Btu/day. Considering a natural gas price of \$6.45 per million Btu [26] and the average cost of landfilling municipal solid waste in the US in 2022 at \$58.47 per ton [18], the GEP of the base case is 0.95 million dollars per year.

In the second scenario, given the absence of specific Volatile Solid Removal (VSR) data in the literature, we approximated the improvement in VSR by referencing the enhancement in Soluble Chemical Oxygen Demand (SCOD) removal. Both SCOD and VSR serve as crucial indicators of organic material removal efficiency, and according to the literature [27], they tend to exhibit similar trends. According to [9] the highest organic conversion rate corresponds to a 33% increase in SCOD removal efficiency. This translates to approximately a 40% conversion of volatile solids for the pretreated PPS in our study, leading us to anticipate a methane yield of 174.60 mlCH<sub>4</sub>/gVS<sub>fed</sub>, equivalent to 358855.71 ft<sup>3</sup>CH<sub>4</sub>/day and 23.28 tons PPS conversion per day. Taking these factors into account, the GEP for the second scenario is estimated to be \$1.26 million per year.

According to the literature, AD of food waste gave a specific methane yield of 470 mlCH<sub>4</sub>/gVS<sub>fed</sub>, which is equal to approximately 70% of the theoretical value (660 mlCH<sub>4</sub>/gVS<sub>fed</sub> with a biogas methane content of 58%) based on the Buswell equation. Also, the food waste exhibited a TS content of 23.9% and a VS content of 21.6% [28]. Assuming the absence of any synergistic effects, the biodegradation efficiency, calculated by averaging the efficiencies of AD for both FW and PPS, stands at 50%. Using the Buswell equation and the elemental composition for the mixture of PPS and FW at a TS ratio of 1:1 (refer to Table 3), the cumulative methane yield reaches 291.76 mlCH<sub>4</sub>/gVS<sub>fed</sub>, aligning with the findings reported

in the literature [29]. With a fixed ratio of 0.94 gVS/gTS [23,28] and a total solids content of 12%, the calculated methane yield amounts to 581,107.51 ft<sup>3</sup>/day, resulting in a converted PPS of 28.2 tons/day. Consequently, for the third scenario, the GEP is \$1.85 million per year.

**Table 3:** Elemental compositions of the food waste and mixed Pulp and paper sludge

Parameters	FW [28]	FW: Mixed PPS (1:1)
VS/TS (g/g)	0.90	0.94
C (% TS)	51.10	47.75
H (% TS)	6.41	6.19
O (% TS)	32.50	33.68
N (% TS)	3.10	2.26
S (% TS)	0.00	0.40
Ash (% TS)	6.89	9.72
Total	100	100

## Sizing and Capital Cost

### Digester

Considering 12% TS and an organic loading rate (OLR) of 5 kgVS/m<sup>3</sup>day, the required digester volume is 11,220 m<sup>3</sup> or approximately 3 million gallons (MG) [30].

The hydraulic residence time (HRT) was calculated as 22.44 days for a 3 MG digester processing 500 tons of feed per day, aligning with established industry standards. The cost estimation is derived from the Synergy Biogas, LLC Case Study [20]. The purchased cost from the case study has been adjusted to suit the assumed capacity of the AD systems in this study. The capital cost for a digester vessel with a volume of 2.2 million gallons, as indicated in the report from the year 2011, is specified at \$1.25 million. The cost was adjusted to fit the 3 MG capacity using Eq. (4) yielding a value of \$1.50 million. Subsequently, converting this cost estimation to the 2022 price, the adjusted cost would be approximately \$2.01 million.

### Pretreatment reactor

The sizing of the pretreatment reactor, determined in accordance with established norms, for a given input flow rate (q) and hydraulic retention time (HRT) of 1 hour is computed using Eq. (8), resulting in a volumetric capacity of 20.83 m<sup>3</sup> (Equivalent to a process vessel with a height of 5 m and an inside diameter of 2.3 m).

$$Volume(m^3) = HRT \times q \quad (8)$$

Based on Ulrich's graphs (see Fig 5.44 of reference [19]), the estimated purchase cost for a carbon steel vessel of this size under atmospheric pressure is approximately \$150,000.

For Scenario B, the total bare module cost for this initial TEA is assessed by combining the purchased costs of digester and pretreatment reactor, totaling \$2.16



**Table 4:** Comparative economic analysis of a PPS-based biogas plant in different scenarios.

Scenarios	CTM (\$MM)	Labor cost (\$MM)	COM (\$MM)	GEP (\$MM)			Payback period (years)	CFRR (%)	NPV (\$MM)
				VP	STF	Total			
A	2.37	0.12	0.99	0.58	0.37	0.95	Undefined	NA	-2.04
B	2.55	0.20	1.26	0.77	0.49	1.26	Undefined	NA	-1.98
C	2.37	0.12	0.99	1.25	0.60	1.85	4.3	22.55	2.50

NA: Not Applicable; C<sub>TM</sub>: Total Module Cost; COM: Cost of Manufacturing; GEP: Gross Economic Potential; CFRR: Cash Flow Rate of Return; NPV: Net Present Value

million. Utilizing Eq. (5), the total module costs for scenarios A, B, and C amount to \$2.37 million, \$2.55 million, and \$2.37 million, respectively.

## Operating Cost

### Labor cost

Based on the specified assumptions, the base case scenario requires 1.5 full-time operators, whereas the scenario involving alkaline pretreatment necessitates 2.5 full-time operators due to an additional reactor.

According to the 2022 yearly mean wage estimates from the US bureau of labor statistics, the reported annual mean wage for chemical plant and system operators is \$79,290 [22]. Using this information, the estimated annual labor cost is \$118,935 for the base case and scenario C and \$198,225 for the scenario incorporating alkaline pretreatment.

### Fixed capital investment

The fixed capital investment (FCI) for the process is equal to the total module cost for each scenario. Finally, using Eq. (7), the total manufacturing cost is estimated to be \$0.99, 1.26, and \$0.99 million for scenarios A, B, and C respectively.

## Cash Flow Analysis

The feasibility of different scenarios is analyzed based on key parameters – i.e., net present value (NPV), cash flow rate of return (CFRR), and payback period. Table 4 summarizes revenue generation through each scenario, and other critical parameters for a 20-year project life (years after Startup). The assumed values, parameters, and equations for the analysis are summarized in Table 5.

Scenario C stands out as the most promising scenario as it was the only one to yield a positive NPV. This favorable outcome is principally due to its high conversion rates and methane production yield, resulting in increased revenue. Although Scenario B has higher conversion rates than the base case scenario, its overall performance is still overshadowed by higher total module cost and comparatively higher COM due to higher labor

costs associated with the additional reactor.

Further research is needed to explore the synergistic effects of co-digesting FW and PPS, ensuring a more comprehensive understanding of their potential economic benefits in AD processes. In addition, to enhance the robustness of our findings and pave the way for future advancements, several additional aspects demand further attention. Considering the potential benefits of co-digesting FW and PPS, along with alkaline pretreatment, presents an opportunity to reveal novel insights for optimizing AD processes. For more accurate cost estimation, it's crucial to integrate biogas upgrading, utility and chemical expenses (covering items such as inoculum for bacterial cultures, buffering agents, antifoaming agents, and, where applicable, enzymes) into the analysis. These efforts will collectively contribute to advancing the economic and environmental sustainability of AD processes in the PPI.

**Table 5:** Parameters and equations employed for the cash flow analysis

Parameter	Unit	Value
Project life	Years	20
Construction period	Years	3
Taxation rate	%	21
Annual interest rate	%	10
Depreciation method	-	MACRS (5-year)
Annual interest rate	%	10
Salvage value	\$	0.1*FCI [21]
Working capital	\$	0.1*(C <sub>RM</sub> + FCI + C <sub>OL</sub> ) [21]

## CONCLUSION

This study investigates the potential of AD to valorize PPS through TEA. Three distinct scenarios are compared: (A) the base case, (B) sludge AD with an alkaline pretreatment, and (C) co-digestion with nitrogen-rich

feedstocks.

The study highlighted the enhanced economic viability of PPS digestion through the integration of food waste (Scenario C). Alkaline pretreatment (Scenario B) also showed potential with a 33% increase in volatile solids conversion.

Further research is needed to investigate the synergistic effects of co-digesting food waste and PPS in AD and integrate biogas upgrading, utility, and chemical expenses in the TEA. Conducting experiments to optimize process parameters such as temperature, pH, TS content, OLR, and retention time could enhance methane production efficiency in the pulp and paper industry's AD processes. In addition, exploring potential applications for the AD digestate could lead to additional revenue streams or beneficial reuse options.

## ACKNOWLEDGMENTS

The authors have not reported funding for this work.

## REFERENCES

1. Oel PR, Hoekstra AY. Towards Quantification of the Water Footprint of Paper: A First Estimate of its Consumptive Component. *Water Resour Manag* 26: 733–749 (2012)
2. Pokhrel D, Viraraghavan T. Treatment of pulp and paper mill wastewater—a review. *Sci. Total Environ.* 333: 37–58 (2004)
3. Priadi C, Wulandari D, Rahmatika I, and Moersidik SS. Biogas Production in the Anaerobic Digestion of Paper Sludge. *APCBEE Procedia* 9: 65–69 (2014)
4. Gottumukkala LD, Haigh K, Collard FX, Rensburg EV, and Görgens J. Opportunities and prospects of biorefinery-based valorisation of pulp and paper sludge. *Bioresour Technol* 215: 37–49 (2016)
5. Kamali M, Gameiro T, Costa MEV, Capela I. Anaerobic digestion of pulp and paper mill wastes – An overview of the developments and improvement opportunities. *J Chem Eng* 298:162–182 (2016)
6. Jackson MJ, Line MA, S. Wilson, Hetherington SJ. Application of Composted Pulp and Paper Mill Sludge to a Young Pine Plantation. *J Environ Qual* 29:407–414 (2000)
7. Whittemore RC, LaFleur LE, Gillespie, WJ, Amendola, GA, Helms J. USEPA/paper industry cooperative dioxin study: The 104 mill study, *Chemosphere* 20:1625–1632 (1990)
8. Paul Greene. <https://www.biocycle.net/digester-sizing-rng-performance/>
9. Lin Y, Wang D, Wu S, Wang C. Alkali pretreatment enhances biogas production in the anaerobic digestion of pulp and paper sludge. *J Hazard Mater* 170:366–373 (2009)
10. Navia R, Soto M, Vidal G, Bornhardt C, Diez MC. Alkaline Pretreatment of Kraft Mill Sludge to Improve Its Anaerobic Digestion. *Bull Environ Contam Toxicol* 69:869–876 (2002)
11. Loow Y, Wu TY, Jahim JM, Mohammad AW, Teoh WH. Typical conversion of lignocellulosic biomass into reducing sugars using dilute acid hydrolysis and alkaline pretreatment. *Cellulose* 23:1491–1520 (2016)
12. Sewsynker-Sukai Y, David AN, Kana EBG. Recent developments in the application of kraft pulping alkaline chemicals for lignocellulosic pretreatment: Potential beneficiation of green liquor dregs waste. *Bioresour Technol* 306:123225 (2020)
13. Lin Y, Wang D, Liang J, Li G. Mesophilic anaerobic co-digestion of pulp and paper sludge and food waste for methane production in a fed-batch basis. *Environ Technol* 33:2627–2633 (2012)
14. Bayr S, Rintala J. Thermophilic anaerobic digestion of pulp and paper mill primary sludge and co-digestion of primary and secondary sludge. *Water Res* 46:4713–4720 (2012)
15. Meyer T, Edwards EA, Anaerobic digestion of pulp and paper mill wastewater and sludge. *Water Res* 65: 321–349 (2014)
16. Veluchamy C, Kalamdhad AS. Influence of pretreatment techniques on anaerobic digestion of pulp and paper mill sludge: A review. *Bioresour Technol* 245:1206–1219 (2017)
17. Achinas S, Euverink GJW. Theoretical analysis of biogas potential prediction from agricultural waste. *Resource-Efficient Tech* 2:143–147 (2016)
18. Bruna Alves. <https://www.statista.com/statistics/692063/cost-to-landfill-municipal-solid-waste-by-us-region/>
19. Ulrich G, Vasudevan P. Chemical Engineering: Process Design and Economics; a Practical Guide. Process Publ. (2004)
20. Labatut R, Gooch C. Anaerobic Digestion at Synergy Biogas, LLC: Case Study. (2012)
21. Turton R. CAPCOST software to accompany: Analysis, synthesis, and design of chemical processes. (2010).
22. U.S. BUREAU OF LABOR STATISTICS. <https://www.bls.gov/oes/current/oes518091.htm>
23. A. Lopes ACP, Silva CM, Rosa AP, Rodrigues FA. Biogas production from thermophilic anaerobic digestion of kraft pulp mill sludge. *Energy J* 124:40–49 (2018)
24. Bayr S, Kaparaju P, Rintala J. Screening pretreatment methods to enhance thermophilic anaerobic digestion of pulp and paper mill wastewater treatment secondary sludge. *J Chem Eng* 223:479–486 (2013)

25. The Engineering ToolBox. Fuel Gases - Heating Values.  
[https://www.engineeringtoolbox.com/heating-values-fuel-gases-d\\_823.html](https://www.engineeringtoolbox.com/heating-values-fuel-gases-d_823.html)
26. Henry Hub Natural Gas Spot Price.  
<https://www.eia.gov/dnav/ng/hist/rngwhhdm.htm>
27. Dasgupta A, Chandel M. Enhancement of biogas production from organic fraction of municipal solid waste using alkali pretreatment. *J Mater Cycles Waste Manag* 22: 757-767 (2020)
28. Yirong C, Banks CJ, Heaven S. Comparison of mesophilic and thermophilic anaerobic digestion of food waste. *AD13 Recovering (bio) Resources for the World* (2013)
29. Lin Y, Wang D, Liang J, Li G. Mesophilic anaerobic co-digestion of pulp and paper sludge and food waste for methane production in a fed-batch basis. *Environ Technol* 33: 2627-2633 (2012)
30. Gautam R, Nayak JK, Daverey A, Ghosh UK. Emerging sustainable opportunities for waste to bioenergy: an overview. In: Waste-to-Energy Approaches Towards Zero Waste. Ed: Hussain CM, Singh S, Goswami L. Elsevier (2022)
31. Corporate income tax (CIT) rates.  
<https://taxsummaries.pwc.com/quick-charts/corporate-income-tax-cit-rates>

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Integrated Ex-Ante Life Cycle Assessment and Techno-Economic Analysis of Biomass Conversion Technologies Featuring Evolving Environmental Policies

Dat T. Huynh<sup>a</sup>, Marianthi Ierapetritou<sup>a\*</sup>,

<sup>a</sup> Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, Delaware, USA

\* Corresponding Author: [mgi@udel.edu](mailto:mgi@udel.edu).

## ABSTRACT

Biorefineries can reduce carbon dioxide emissions while serving the global chemical demand market. Governments are also using carbon pricing policies, such as carbon taxes, cap-and-trade models, and carbon caps, as a strategy to reduce emissions. The use of biomass feedstocks in conjunction with carbon capture usage and storage technologies are mitigation strategies for global warming. Businesses can invest in these technologies to accommodate the adoption of these policies. Rapid action is necessary to halt global warming, which results in aggressive policies. In this work, a multi-period process design and planning problem is developed for the design and capacity expansion of biorefineries. The three carbon pricing policies are integrated into the model and parameters are selected according to the aggressive scenario denoted by the Paris Agreement. The results show that the cap-and-trade policy achieves a higher net present value evaluation over the carbon tax model across all pareto points due to the flexibility of the allowances in the cap-and-trade policy. The carbon cap model substantial investments are required in carbon capture technologies to adhere to the emissions constraints.

**Keywords:** Biomass, Life Cycle Analysis, Technoeconomic Analysis, Technoeconomic Analysis, Process Design

## INTRODUCTION

CO<sub>2</sub> emissions from energy combustion and industrial processes have risen from 24.9 Gt CO<sub>2</sub> to 36.8 Gt CO<sub>2</sub> from 2000 to 2022[1]. The Intergovernmental Panel on Climate Change (IPCC) reported that urgent action is necessary to curb global warming to 1.5°C[2]. Towards that end, scientists and policymakers are developing solutions to mitigate CO<sub>2</sub> contributions to the global warming crisis.

Traditional chemical manufacturing uses petroleum-based feedstocks, which are unsustainable resources and result in high CO<sub>2</sub> emissions. In an effort to reduce reliance on petroleum-based feedstocks, scientists have been researching lignocellulosic biomass as a feedstock alternative. Lignocellulosic biomass is an abundant resource and has the potential to be sustainable with low emissions. The biorefinery concept proposes that each major component from lignocellulosic biomass, i.e.,

cellulose, hemicellulose, and lignin, can be fractionated and valorized into chemicals, like petroleum refinery and chemical plant operations. Biorefineries supports decarbonization by transitioning towards a sustainable feedstock and lower emissions processes.

Significant research has been conducted in the Carbon Capture, Utilization, and Storage (CCUS) field that aims to reduce the amount of CO<sub>2</sub> currently emitted by industrial processes and capture CO<sub>2</sub> already existing in the atmosphere. For example, Yusuf et al. evaluated the economic feasibility of producing soda ash from CO<sub>2</sub> heavy flue gas generated from power plants, a Carbon Capture and Utilization (CCU) technology[3]. Wang et al. performed a technoeconomic analysis on the sequestration of CO<sub>2</sub> flue gas from power plants via compression and storage, a Carbon Capture and Storage (CCS) technology[4].

Governments are increasingly leveraging environmental policies to reduce CO<sub>2</sub> emissions. As of 2022, 23%

of all CO<sub>2</sub> emissions are under some form of carbon pricing policy[5]. Fifty-two countries enforce a carbon tax, Emissions Trading System (ETS), or both policies[5]. Twenty countries are currently considering the implementation of these policies as they can provide not only environmental but also social and economic benefits[5]. Under a carbon tax policy, carbon dioxide emitters are charged a financial penalty per ton of CO<sub>2</sub> emitted. An ETS is a system enforcing a cap-and-trade model where the government provides allowances, an amount of permitted CO<sub>2</sub> emissions, for manufacturers. They can purchase additional allowances or sell unused allowances on an open market. Benchmarks have been set via carbon pricing to limit global warming to 2°C. According to the Paris Agreement, emission levels should be reduced by 45% by 2030 and reach net-zero carbon emissions by 2050[5]. An additional benchmark provided by the World Bank states that carbon pricing should be between 61 and 122\$ by 2030[5].

Superstructure optimization is used as a framework for exploring multiple process design alternatives. Luo et al. utilized neural networks to model the biorefinery flexibility index facilitating operational flexibility constraints in superstructure optimization[6]. Multi-period optimization can be used for considering planning problems over a long-time horizon. Sabet et al. used a multi-period formulation to model a global manufacturing capacity management problem[7]. These two approaches can be integrated with environmental policy to optimize process designs.

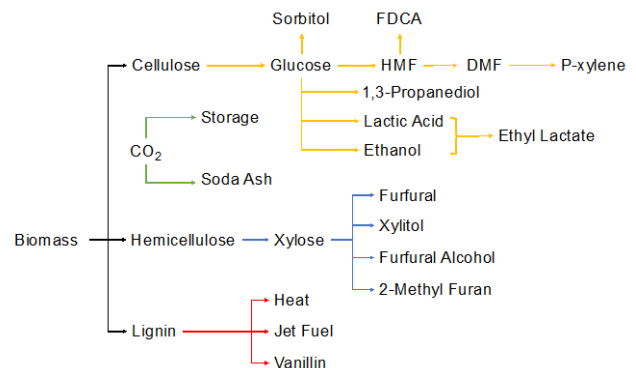
This work incorporates the benchmarks provided by the World Bank and IPCC into a multi-period biorefinery design and optimization problem[2, 5]. Three different carbon emissions policies, namely carbon cap, cap and trade, and carbon tax, are considered as constraints in the formulation. Pareto fronts are constructed for economic and environmental objective functions.

## MULTI-PERIOD PROGRAMMING FOR BIOREFINERY

The multi-period programming formulation is utilized for the long term biorefinery construction and expansion optimization problem. The planning horizon is set for thirty years corresponding to the Paris Agreement goals. Each time period represents one year. In the first year, an initial biorefinery is constructed. In each subsequent time period, the biorefinery can experience capacity expansion or the construction of new units. The problem is constrained by three environmental policies that are increasingly restrictive over each time period to match the IPCC and Paris Agreement benchmarks while maximizing net present value.

A superstructure approach is used. In this work, the superstructure represents all process alternatives

consisting of chemical transformations and separation sequences. The reactions were selected to represent a broad range of chemicals, which is displayed in Figure 1. Commodity chemicals, such as ethanol, and biomass platform chemicals, such as furfural, were included. Drop-in chemicals, such as para-xylene, and biomass derived alternatives, such as furan-dicarboxylic acid, were also included. Different separation steps were considered consisting of crystallization, distillation, extraction, membrane separation, and pervaporation. Shortcut methods and surrogate models are used to characterize the utility usages. Carbon capture storage and carbon capture and utilization technologies were also incorporated into the superstructure to accommodate the dynamic environmental policies.



**Figure 1.** Biorefinery superstructure with CCUS technologies.

## Objective Functions

The objective of the optimization problem is to maximize net present value (NPV) and minimize cumulative emissions (CE) while adhering to the carbon pricing policies. The net present value calculation is shown in Equation (1) where  $ir$  represents the interest rate,  $t$  represents the time period;  $IC_t$  represents the capital investment in time period  $t$ ;  $R_t$  represents the product revenue in time  $t$ ;  $O_t$  represents the operating cost in time  $t$ ; and  $C_{CO_2,t}$  represents the carbon dioxide cost in time period  $t$ .

$$NPV = \sum_{t \in T} (1 + ir)^{-t} \{ -(IC_t - IC_{t-1}) + R_t - O_t \pm C_{CO_2,t} \} \quad (1)$$

The power law model in Equation (2) captures the capital costs where  $a_u$  and  $b_u$  are parameters for unit operation  $u$ , and  $x_{u,t}$  represents the cumulatively capacity of unit  $u$  in period  $t$ . The cost of capacity expansion in time period  $t$  is captured as the difference between the cost of a plant with the cumulative capacity and the cost of the plant in the previous expansion period.

$$IC_t = \sum_{u \in U} a_u (x_{u,t})^{b_u}, t > 1 \quad (2)$$

The revenue,  $R_t$ , generated in time period  $t$  by the



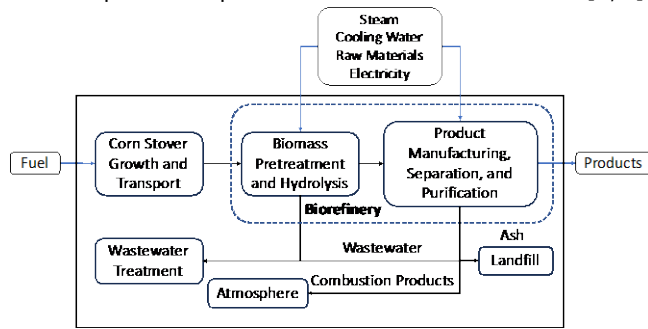
products are captured in Equation (3), where  $b_{i,t}$  is the amount of chemical  $i$  produced in time period  $t$ , and  $C_i$  represents the cost of chemical  $i$ .

$$R_t = \sum_{i \in I} b_{i,t} C_i \quad (3)$$

The operating cost,  $O_t$ , in time period  $t$  is captured in Equation (4), where  $f_{j,t}$  is the operating level of unit operation  $j$  in time period  $t$ ;  $C_j$  is the unit cost of running unit  $j$ ;  $E_{w,t}$  is the energy usage of utility  $w$ ;  $C_w$  is the unit cost of operating utility  $w$ ; and  $fcf$  is the fixed cost factor.

$$O_t = \sum_{j \in J} f_{j,t} C_j + \sum_{w \in W} E_{w,t} C_w + IC_t(fcf) \quad (4)$$

The environment impact calculation is based on a cradle-to-gate life cycle assessment. The system boundary is depicted in Figure 2 and considers biomass transportation, raw material production, utilities, combustion products, landfill, and wastewater treatment. Biomass is assumed to be carbon neutral. The data for the calculations are obtained from the Ecoinvent v.3.8 database, and the Global Warming Potential (GWP) indicator from the ReCipE2016 impact assessment method is used[8, 9].



**Figure 2.** System boundary depicted for environmental impact calculations. Biomass is transported to the biorefinery. Raw materials and utilities are imported into the biorefinery for chemical production. Wastewater is exported for wastewater treatment. Emissions from the combustion of lignin to the atmosphere and the resulting ash sent to the landfill are included.

The environmental impact in time period  $t$ ,  $TE_t$ , is given by Equation (5) where  $GWP_w$  represents the GWP of utility  $w$ ;  $b_{ww,t}$  represents the amount of wastewater generated in time period  $t$ ;  $GWP_{ww}$  represents the GWP of wastewater treatment;  $f_{rm,t}$  represents the amount of raw material  $rm$  used in time period  $t$ ;  $b_{Ash,t}$  represents the amount of ash generated in time period  $t$ ;  $GWP_{Ash}$  represents the GWP of sending the ash to the landfill;  $b_{comb}$  represents the combustion products in time period  $t$ ; and  $GWP_{comb}$  represents the GWP of the combustion product mixture.

$$TE_t = \sum_w E_{w,t} GWP_w + b_{ww,t} GWP_{ww} + \sum_{rm \in RM} f_{rm,t} GWP_{rm} + b_{Ash,t} GWP_{Ash} + b_{comb} GWP_{comb} \quad (5)$$

The second objective function, cumulative emissions (CE), is given by Equation (6) where  $TE_t$  represents the total emissions in time period  $t$ .

$$CE = \sum_t TE_t \quad (6)$$

In this work, it is assumed that the residence time of CO<sub>2</sub> emissions is longer than the planning period and, therefore, that the emissions across the time periods have equal weight.

## Constraints

The processing of materials is described by the molar balance Equation (7) where  $v_{i,j}$  represents the conversion coefficient for compound  $i$  in operation  $j$ ;  $f_{j,t}$  is the extent of process  $j$  in time period  $t$ ; and  $b_{i,t}$  is the amount of chemical  $i$  in time period  $t$ .

$$\sum_j v_{i,j} f_{j,t} = b_{i,t} \quad (7)$$

The unit operation expansion is described by Equations (8a) and (8b) where  $x_{j,t}$  represents the capacity of unit operation  $j$  in time period  $t$ ; where  $x_{j,t}^{exp}$  the additional capacity added to unit operation  $j$  in period  $t$ ; and  $x_j^{init}$  represents the initial capacity built for unit operation  $j$ .

$$x_{j,t} = x_{j,t-1} + x_{j,t-1}^{exp} \quad \forall j; t > 1 \quad (8a)$$

$$x_{j,t} = x_j^{init} \quad \forall j; t = 1 \quad (8b)$$

The capacity expansion is limited in its lower and upper bound as expressed in Equation (9), where  $Y_{j,t}$  is a binary variable that equals 1 when there is capacity expansion; where  $Cap^{Lo}$  represents the minimum possible capacity expansion; and  $Cap^{Up}$  represents the maximum possible capacity expansion.

$$Cap^{Lo} Y_{j,t} \leq x_{j,t}^{exp} \leq Cap^{Up} Y_{j,t} \quad (9)$$

The capacity is limited to a fixed number of expansions represented by Equation (10), where  $Y_{j,t}$  represents expansion in time period  $t$ , and  $E_j$  represents the number of expansions permitted for unit  $j$ .

$$\sum_t Y_{j,t} \leq E_j \quad \forall j \quad (10)$$

The operating level of unit  $j$  is constrained by the maximum capacity of unit  $j$ , which is expressed in Equation (11).  $x_{j,t}$  represents the capacity of unit  $j$  in time period  $t$ ;  $h$  represents the minimum operating ratio; and  $f_{j,t}$  represents the operating level of unit  $j$  in time period  $t$ .

$$hx_{j,t} \leq f_{j,t} \leq x_{j,t} \quad \forall j \in J \forall t \in T \quad (11)$$

The plant size is limited to amount  $m_{bm}$  as expressed in Equation (12) where  $x_{bm,t_{end}}$  represents the biomass feedstock,  $bm$ , processing capacity in the last period.

$$m_{bm} = \sum_{bm \in BM} x_{bm,t_{end}} \quad (12)$$

## Environmental Constraints

The formulations for the carbon policies utilized in this work are presented below. These policies enforce environmental constraints and may affect the NPV calculation.

The carbon cap policy enforces a fixed amount of CO<sub>2</sub> emissions. In this work, we consider the cap to be placed on the aggregated amount of emissions in the time period of one year. The constraint is expressed in Equation (13) where  $TE_t$  represents the amount of CO<sub>2</sub> emitted in time period  $t$ , and  $TE_t^{cap}$  represents the emissions cap in time period  $t$ .

$$TE_t \leq TE_t^{cap} \quad \forall t \in T \quad (13)$$

In the cap-and-trade policy, the governing body allocates a number of allowances to manufacturers. This represents a soft emissions cap, which can be exceeded by purchasing additional allowances from other manufacturers or can be sold for profit. The constraint is expressed in Equations (14-16) where  $P_{CO_2,t}$  represents the price of CO<sub>2</sub> in time period  $t$ ;  $E_t^+$  represents the allowances purchased in time period  $t$ ;  $E_t^-$  represents the allowances sold in time period  $t$ ; and  $TE_t^{cap}$  represents the allowances provided in time period  $t$ .

$$C_{CO_2,t} = P_{CO_2,t}(E_t^+ - E_t^-) \quad (14)$$

$$TE_t \leq TE_t^{cap} + E_t^+ - E_t^- \quad \forall t \in T \quad (15)$$

$$E_t^+ > 0, E_t^- > 0 \quad \forall t \in T \quad (16)$$

Under a carbon tax policy, manufacturers are charged per tCO<sub>2</sub> emitted. The total carbon tax is given by Equation (17), where  $P_{CO_2,t}$  represents the price of one ton of CO<sub>2</sub> emitted in time period  $t$ ;  $TE_t$  represents the amount of CO<sub>2</sub> emitted in time period  $t$ ; and  $C_{CO_2,t}$  represents the carbon tax cost associated with those emissions in period  $t$ .

$$C_{CO_2,t} = P_{CO_2,t}TE_t \quad (17)$$

## CASE STUDY RESULTS AND DISCUSSION

The epsilon constraint method is used to construct a Pareto front for the multi-period biorefinery optimization problem. NPV and CE are the two functions considered in the bi-objective optimization. The nonlinear Equation (2) is reformulated via piecewise linearization to keep

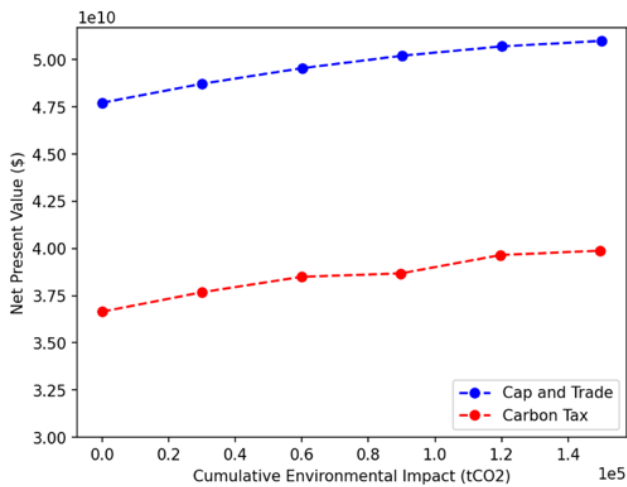
the formulation linear. Consequently, all instances of the model are formulated and solved in GAMS as a MILP using CPLEX solver on an Intel Xeon E-2247G @ 4.00 GHz CPU and 32.0 GB of RAM.

In our case study, a biorefinery is considered in McClean, IL with a plant capacity set at 2984 metric tons per year corresponding to four times the nominal corn stover production in McClean. Additional biomass can be purchased within the five closest counties within McClean. The years 2020 to 2050 are considered to represent a thirty-year time horizon with each time period having a length of one year. The carbon pricing parameters considered correspond with the aggressive scenario set by the Paris Agreement, which aims to maintain global warming below 2°C. Table 1 presents the parameter benchmarks. Linear interpolation is used to determine the parameters in the intermediate years. The carbon tax and cap-and-trade policies are evaluated by constructing pareto fronts to compare economic and environmental trade-offs. The carbon cap policy is analyzed yearly to elucidate the effects of a shrinking carbon cap.

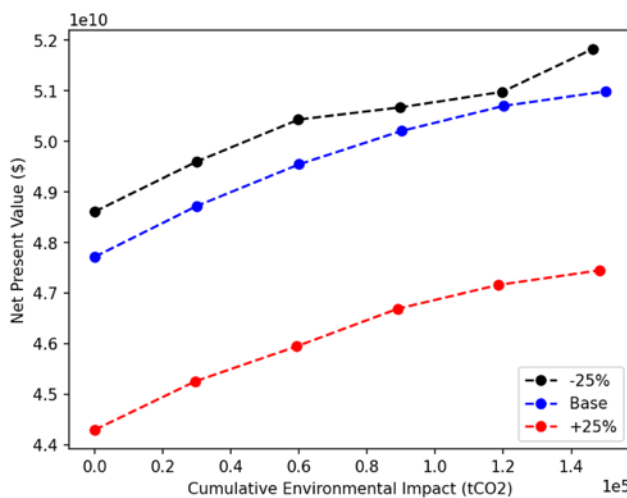
**Table 1:** Carbon pricing policy parameters during milestone years

Year	Carbon Cap (%)	Carbon Tax (\$/tCO <sub>2</sub> )
2020	100	0
2030	45	59
2050	0	295

Figure 3 demonstrates increasing NPV with increasing CE for both cap-and-trade and carbon tax policies. For the cap-and-trade policy, the allowances provided are equal to the carbon cap parameters given in Table 1. Similarly, the carbon prices are set at the carbon tax value in Table 1. At the minimum CE point for both policies, a positive NPV exists. Across all points on the pareto front, the cap-and-trade policy results in a higher NPV than the carbon tax policy. This is a consequence of the allowances that can be sold for a profit when the carbon cap is high as well as the allowances providing tax-free emissions. Additional production incurs a larger financial penalty under the carbon tax policy, resulting in lower overall production. This is noted through the maximum profit point for the carbon tax policy having lower CE than the cap-and-trade policy.



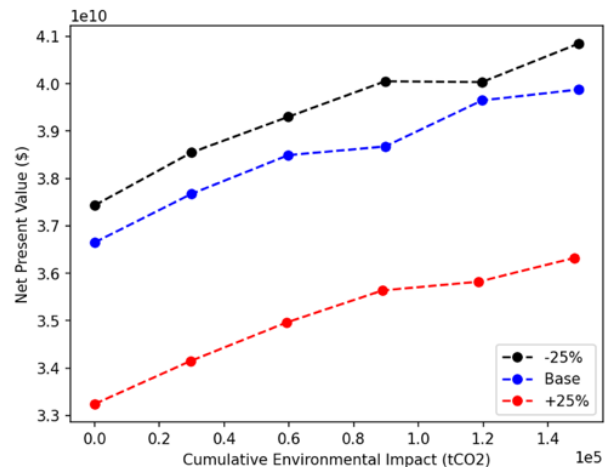
**Figure 3.** NPV and CE pareto curve for cap-and-trade and carbon tax policies



**Figure 4.** Sensitivity analysis for NPV and CE pareto curve for a cap-and-trade policy

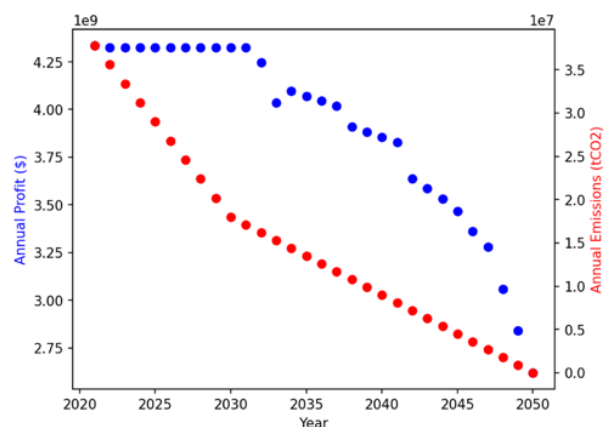
A sensitivity analysis was performed for the cap-and-trade and carbon tax policies. The raw material costs were varied by 25%. Figure 4 and Figure 5 show the changes in the pareto curves for the cap-and-trade and carbon tax policies, respectively. In both policies, increasing the raw material price by 25% has a significantly greater impact than decreasing the raw material price by 25%. For the carbon tax policy, the average relative difference to the base case for the 25% increase and decrease case is 9.0% and 2.2%, respectively. For the cap-and-trade policy, the average relative difference to the base case between the 25% increase case and decrease case is 7.1% and 1.4%, respectively. The large decrease in NPV in the 25% increase case is explained by the change in production relative to the base case. In the 25% increase case, production shifts from ethyl lactate to ethanol production which has higher raw material costs. The average relative differences for the carbon tax policy are greater than those of the cap-and-trade policy because

emissions are more heavily penalized under the carbon tax policy.



**Figure 5.** Sensitivity analysis for NPV and CE pareto curve for carbon tax policy

Figure 6 demonstrates the effect of an increasingly restrictive carbon cap over time. It is clear from the emissions curve that the rate of carbon dioxide reduction is greater between 2020 to 2030 than between 2030 and 2050. Despite the rapid reduction in emissions levels, the yearly profit generated is unaffected until 2029. This is a consequence of decreasing production in high carbon dioxide emitting chemicals that do not significantly contribute to profit. After 2030, emissions cannot continue to decrease without decreases to profit. Every year an investment is made, the slope of the annual profit line changes, reflecting the change in operation regimes as more CCS and CCU is required to maintain policy compliance.



**Figure 6.** Annual profit and emissions with time for carbon cap policy with reduced CCS and CCU costs

Table 2 displays the investment capacities and capacity expansions for both CCS and CCU technologies

under the carbon cap policy. In year 2020, there is a small investment in both CCS and CCU technology during the initial construction of the biorefinery. Initially, annual emissions are reduced by altering operations. In year 2033 and 2047, there is a significant investment in CCU technology. In the year 2041 and 2045, there is an additional investment in CCS technology. There is a higher investment in CCU technology overall due to its lower operating cost, despite its higher capital cost.

**Table 2:** Initial capacity and expansions for carbon capture technologies under a carbon cap policy

Year	Carbon Capture Capacity (tCO <sub>2</sub> /yr)	
	CCS	CCU
2020	1.9E3	1.1E5
2033	0	5.6E6
2037	0	4.9E6
2041	3.5E6	0
2045	2.1E6	0

## CONCLUSIONS

This work formulated a biorefinery process design and capacity expansion problem. A multi-period programming approach was utilized to consider the capacity expansion decisions when carbon pricing increases and carbon caps decrease with time in accordance with benchmarks of the Paris Agreements. The carbon tax, cap-and-trade, and carbon cap policies were formulated as constraints to evaluate their effects on the NPV and EI pareto fronts. The framework allows manufacturers to plan biorefinery product portfolios and future expansion projects considering carbon pricing policies.

The cap-and-trade policy is evaluated to be more profitable compared to the carbon tax policy and includes greater flexibility as a consequence of the purchasing and selling of allowances mechanism as well as the portion of carbon tax free emissions. The carbon cap policy has shown the importance and necessity of reducing the cost of CCUS technologies for chemical plants to adhere to increasingly strict carbon caps over time. The carbon tax policy results in decreased chemical production due to the financial penalty further highlighting the need for low cost CCUS technologies to offset emissions and achieve net zero carbon emissions.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge financial support from the National Science Foundation under award number NSF-2134471.

## REFERENCES

1. CO<sub>2</sub> Emissions. Paris: IEA; 2023.

- Shukla PR, Skea J, Slade R, Al Khourdajie A, Van Diemen R, McCollum D, et al. Climate change 2022: Mitigation of climate change. Contribution of working group III to the sixth assessment report of the Intergovernmental Panel on Climate Change. 2022;10:9781009157926.
- Yusuf A, Giwa A, Mohammed EO, Mohammed O, Al Hajaj A, Abu-Zahra MR. CO<sub>2</sub> utilization from power plant: A comparative techno-economic assessment of soda ash production and scrubbing by monoethanolamine. *Journal of Cleaner Production*. 2019;237:117760.
- Wang R, Ashkanani HE, Li B, Morsi BI. TEA of a Unique Two-Pathways Process for Post-Combustion CO<sub>2</sub> Capture. *Journal of Energy and Power Technology*. 2022;4(04).
- Bank W. State and Trends of Carbon Pricing 2023. Washington, DC: Wolrd Bank; 2023.
- Luo Y, Ierapetritou M. Multifedstock and Multiproduct Process Design Using Neural Network Surrogate Flexibility Constraints. *Ind Eng Chem Res*. 2023;62(5):2067-79.
- Sabet E, Yazdani B, Kian R, Galanakis K. A strategic and global manufacturing capacity management optimisation model: A Scenario-based multi-stage stochastic programming approach. *Omega*. 2020;93:102026.
- Wernet G, Bauer C, Steubing B, Reinhard J, Moreno-Ruiz E, Weidema B. The ecoinvent database version 3 (part I): overview and methodology. *Int J Life Cycle Assess*. 2016;21(9):1218-30.
- Huijbregts MAJ, Steinmann ZJN, Elshout PMF, Stam G, Verones F, Vieira M, et al. ReCiPe2016: a harmonised life cycle impact assessment method at midpoint and endpoint level. *Int J Life Cycle Assess*. 2017;22(2):138-47.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Screening Green Solvents for Multilayer Plastic Films Separation

Ugochukwu M. Ikegwu, Victor M. Zavala, and Reid C. Van Lehn\*

University of Wisconsin-Madison, Department of Chemical and Biological Engineering, Madison, Wisconsin, USA.

\* Corresponding Author: [vanlehn@wisc.edu](mailto:vanlehn@wisc.edu).

---

## ABSTRACT

This paper introduces a computational framework for selecting green solvents to separate multilayer plastic films, particularly those challenging to recycle through mechanical means. The framework prioritizes the selective dissolution of polymers while considering solvent toxicity. Initial screening relies on temperature-solubility dependence, utilizing octanol-water partition coefficients (LogP) to identify non-toxic solvents ( $\text{LogP} \leq 3$ ). Additionally, guidelines from GlaxoSmithKline (GSK), Registration, Evaluation, Authorization, and Restriction of Chemical Regulation (REACH), and the US Environmental Protection Agency (EPA) are employed to screen for green solvents. Molecular-scale models predict temperature-dependent solubilities and LogP values for polymers and solvents. The framework is applied to identify green solvents for separating a multilayer plastic film composed of polyethylene (PE), ethylene vinyl alcohol (EVOH), and polyethylene terephthalate (PET). The case study demonstrates the framework's effectiveness in identifying environmentally friendly solvents and balancing trade-offs between solvent toxicity and solubility. Furthermore, the framework informs process design by screening for suitable green solvents in selective dissolution processes, potentially leading to the development of more sustainable dissolution processes and the identification of easily recyclable polymer blends in multilayer plastic films.

---

**Keywords:** Plastics Recycling, Green Solvents, Process Design, Life Cycle Analysis, Technoeconomic Analysis, Polymer, COSMO-RS.

## INTRODUCTION

Multilayer plastic films have been pivotal in enhancing product safety, security, and handling [1]. These films, mostly comprised of polyethylene (PE), ethylene vinyl alcohol (EVOH), polyethylene terephthalate (PET), polyvinyl acetate (PVA) polymers, etc., serve diverse functions in prolonging the shelf life of food products [2]. Nevertheless, the intricate mix of these polymers complicates the recycling process, resulting in only 5% of multilayer plastic films being recycled in the United States; the rest are disposed of through incineration or in landfills. Various technologies have been utilized for recycling plastic waste, including thermochemical methods such as pyrolysis and thermal oxo-degradation, chemical processes, supercritical fluid application, and mechanical processes, but these processes are often not suitable for multilayer films.

As documented in previous studies, solvent-based separation technologies have exhibited significant potential in recovering virgin resins from multilayer plastic films [3]. An illustrative embodiment of this methodology is the solvent-targeted recovery and precipitation (STRAP™) process introduced by Walker et al. [4]. This approach systematically segregates the polymer constituents of multilayer plastic films by employing selective solvents, capitalizing on the temperature-dependent solubility of polymers in solvents, anti-solvents, or a combination of both, as expounded by Sánchez-Rivera et al. [5]. A critical step in achieving this separation lies in the solvent screening, which serves as the linchpin of the entire procedure. Therefore, prediction of temperature-dependent polymer solubility is a critical requirement. To address this need, molecular-scale models have been harnessed to expediently predict the temperature-dependent solubilities of various plastic components,



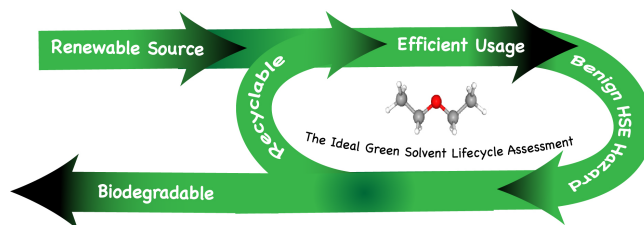
including but not limited to PE, EVOH, PET, and PP, across a spectrum of ~1000 solvents [6]. This was realized by modeling polymers as short oligomers, employing molecular dynamics simulations, and the COnductor-like Screening MOdel for Real Solvents (COSMO-RS), grounded in quantum chemistry principles for solubility determinations.

Furthermore, establishing a predictive framework for polymer solubility, as elucidated by the researchers mentioned above, underscores the influence of selective solvent availability in shaping the sequence of polymer resin recovery. This, in turn, informs the design of the overall process and its associated economic and environmental implications.

In addressing critical environmental challenges, deploying solutions like STRAP™ is essential. However, it is imperative to mitigate the emergence of new toxic waste streams proactively. In dissolution-based recycling methods, solvents typically remain unconsumed. Nonetheless, over repeated usage, minute quantities of these solvents tend to escape into the environment or are deliberately disposed of, potentially creating new waste streams. The ecological impact of such waste streams can be substantial when the solvents in question possess toxic attributes. Moreover, using toxic solvents may pose risks to the safety of operation personnel. Therefore, it is crucial to incorporate process safety and toxicity considerations in selecting solvents, influencing both process design and economic factors [7].

Numerous investigations have sought to elucidate the concept of "green" solvents within the framework of green chemistry principles. Saleh and Koller [8] characterize green solvents as those possessing low reactivity and flammability, exemplified by ionic liquids, per the 12th principle of green chemistry. Choi et al. [9] initially identified natural deep eutectic solvents (NADES) and deemed them environmentally friendly due to their derivation from plant sources.

Considering these different perspectives of green solvents, as discussed in the literature, Figure 1 depicts the ideal green solvent and its life cycle. This ideal green solvent, designed to harmonize sustainability and technical efficiency, should ideally originate from a renewable source, notably plant-derived solvents, and maintain cost-effectiveness as a fundamental criterion. Its utility in dissolution-based procedures should also be distinguished by enhanced solubility and reduced energy consumption, exemplifying its advantageous influence on process efficacy. It should also have benign health, safety, and environmental hazards throughout its life cycle. It should also be easily recyclable and biodegradable for disposal. Although it is necessary to find an ideal green solvent that satisfies all these criteria, it is difficult to achieve these metrics.



**Figure 1.** The Ideal green solvent.

The assessment of solvent greenness is subject to variations in ranking based on the chosen metrics. This was demonstrated by the different perspectives considered by the authors mentioned above to classify solvents as green. Nevertheless, Hessel et al. [10] assert that the principal objective of green solvents should be the mitigation of health, safety, and environmental risks. Therefore, it is vital to adopt a holistic approach in screening solvents for specific purposes, hence this study. To date, no existing study in the literature has formulated a comprehensive framework for the systematic evaluation and selection of environmentally friendly solvents designed explicitly for the separation of multilayer plastic films.

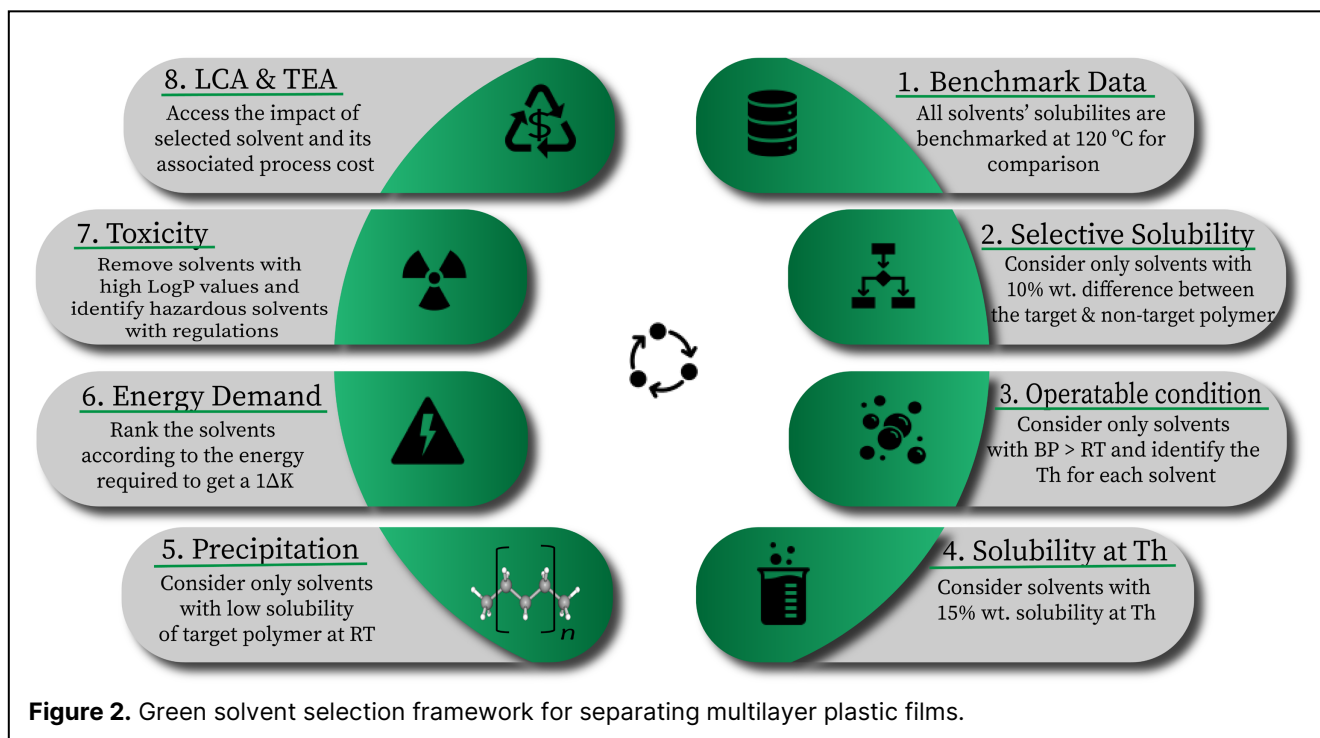
Therefore, in this study, we aim to develop a fast green solvent-screening framework for separating multilayer plastic films into virgin resins that considers solvent solubility and toxicity and how they impact process design and economics. Although this framework was developed in the context of multilayer plastic film separation, it can be extrapolated to any dissolution-based process. This work also elucidates the need to redesign multilayer plastic films, accounting for the ease of recycling food packaging materials. It also motivates the development of new green solvents to separate multilayer plastic films.

## FRAMEWORK DEVELOPMENT

As previously detailed, this study introduces a multistep framework for the fast screening of green solvents in the context of efficiently separating multilayer plastic films into their respective virgin resins, as depicted in Figure 2. The framework starts by predicting solubilities in various solvents at a standard temperature of 120 °C to avoid the thermal degradation of the polymer while providing adequate temperature leeway to achieve optimal solubilities, as per Zhou et al. [6]. While steps (2) through (5) pertain specifically to selective dissolution-based procedures conducted at standard atmospheric pressure, other aspects of this framework are generalizable to solvent-based processes.

### Solubility Prediction and Octanol-Water Partition Coefficient (LogP)

In this investigation, we employed the solubility prediction framework introduced by Zhou et al. [6] to predict the solubilities of target polymers in ~1000 solvents at



**Figure 2.** Green solvent selection framework for separating multilayer plastic films.

120 °C. The target polymers were first modeled as short oligomers in reference solvents using molecular dynamics simulations to obtain different conformational structures. These conformations in the form of trajectories are then optimized and served as inputs into COnductor-like Screening MOdel for Real-Solvents (COSMO-RS) to predict the temperature-dependent solubilities at infinite dilution as described in detail by the authors mentioned above.

The octanol-water partition coefficient (LogP) gives information about a solvent's tendency for bioaccumulation. It is often described as the ratio of concentrations of a chemical in a mixture of octanol (hydrophobic or non-polar phase) and water (hydrophilic or polar phase) in equilibrium. This study employed the COSMO-RS method to predict the LogP values of all solvents.

Solvents with LogP > 0 are considered lipophilic and hence have the possibility of bioaccumulation, while solvents with LogP < 0 are considered hydrophilic. Although solvents with LogP > 0 are generally not regarded as green solvents due to their possibility of bioaccumulating in mammals and aquatic life, this study considered solvents with LogP ≤ 3 as an initial screening process. This choice was to assess the trade-offs among the different selected solvents.

### Fast green solvent screening

Having benchmarked the solubilities of all target polymers in all solvents and computed the required molecular properties described in the preceding sub-section, python-based data analytical techniques were employed to screen for green solvents quickly. To ensure

selective solubility, the solubility difference between the target polymer and other polymers in the multilayer plastic films must be greater than 10 wt.%.

Furthermore, the solvents having boiling points lower than room temperature (RT) were removed. This rule only applies to processes that wish to operate at standard atmospheric pressure. Similarly, the elevated temperature (Th) was set to a degree less than the solvent's boiling point to maintain this operating condition. Solvents with at least 10 – 15 wt.% at Th are then selected. This solubility range has been identified to be reasonable for multilayer plastic film separation [7]. To recover the dissolved layer after hot filtration, the solvent must exhibit low solubility of the target polymer at room temperature for easy precipitation [5]. To assess the first environmental impact of the selected solvents, the minimum energy required by each solvent was determined using Equation 1, and these solvents were ranked accordingly.

$$Q = m C_p \Delta T \quad (1)$$

$m(g)$  denotes the mass of solvent,  $C_p(J/gK)$  denotes the heat capacities of the solvents at the operating temperatures, and  $\Delta T$  is the temperature difference between room temperature and the elevated temperature. Furthermore, solvents with LogP greater than three were removed since these would have a high bioaccumulation tendency. Further solvent screening was subsequently done using data from publicly available industrial solvent selection guidelines. Finally, an LCA and TEA analysis will be done within a process such as STRAP™ to assess the environmental impact and the associated process cost

attached to each selected solvent.

## Solvent toxicity database and regulatory framework

Using a toxic solvent in any solvent-based process runs beyond contributing to the inefficiency of the process. It constitutes the most significant environmental burden to processes resulting from its energy demand and end-of-life fate. This observation has led to different societies and companies developing quantitative toxicity metrics for common solvents as a guide for solvent selection. These guides mostly contain the environmental implications of each solvent's usage. Some of these include guides developed by GlaxoSmithKline (GSK) [11], the ACS Green Chemistry Institute Pharmaceutical Roundtable (ACS GCI-PR) [12], and AstraZeneca (AZ) [13].

The GSK guideline categorized 111 solvents with over ten primary functional groups under six headings. These categories include wastes, environmental impact, health, flammability, reactivity, and life cycle. Scores between 1 and 10 were given to each solvent, with 1 indicating the solvents with the most concern and 10 indicating the solvents with the least concern. Two extra categories were included to classify solvents based on legislation and restriction bans placed by the Environmental Health Services (EHS). Conversely, AZ and ACS GCI-PR guidelines categorized 46 and 63 solvents, with a score of 10 indicating the solvent with the most concern and 1 with the least concern. Although these guidelines contain different metric categories, they all seek to rate solvents based on the Health, Safety, and Environment (HSE) metrics.

This study employed only the GSK solvent guideline because it contains the most solvents, covers the HSE metrics, and includes the legislation and regulatory ban information about some solvents. Although the GSK guide can be used with other guides, the difference in the meaning of the rank assigned to each solvent makes it more difficult. Moreover, using these guides is not the sole toxicity check in this framework, as these guidelines do not contain information about most of the solvents in our database. Hence, other solvent regulatory standards are considered.

The Registration, Evaluation, Authorization, and Restriction of Chemical Regulation (REACH) initiative, a European Union regulation, focuses on limiting the adverse effects of chemicals on human health and the entire ecosystem [14]. This initiative guides the use of solvents both for large-scale processes and our day-to-day lives. This initiative also places restrictions on both common and uncommon chemicals. For example, as of the period of conducting this study, benzene and toluene were enlisted on the list of restricted substances under the REACH [15]. Similarly, the US Environmental Protection

Agency (EPA) aims at securing human lives and the environment from toxic chemicals. The Toxic Substances Control Act (TSCA), established in 1976, provides the EPA with the authority to demand full disclosure of any process relating to chemical substances and mixtures [16].

Furthermore, as of the period this study was conducted, common chemicals like toluene, formamide, and fluorine were included in the list of chemicals covered by the Toxic Release Inventory (TRI) [17]. Processes that use chemicals in this list are generally required to disclose their operations due to their toxicity. Other chemical regulatory frameworks include the Chemical Hazards and Toxic Substances (OCHA) framework, which strives towards similar goals, but rather than enforcing restrictions on chemicals, this framework implements operational standards and handling of chemicals to ensure safe working conditions for personnel.

This study will use the REACH and EPA TSCA solvent list of concern to access the selected solvents from the automated screening process.

## Process Economics and Environmental Impact Assessment

### Techno-economic analysis (TEA)

A techno-economic analysis is done to assess the process cost associated with using a choice solvent within the context of the selected process. This will determine the minimum selling price of the recovered polymer. As illustrated in Figure 1, the ideal solvent must exhibit high efficiency. This translates into a balance across different metrics mentioned in previous sections and process costs. The TEA is done on BioSTEAM, an open-source python-based steady-state process simulator that has been benched marked against existing proprietary software like Aspen Plus [18]. The choice of solvent for multilayer plastic film separation can also inform the dissolution sequence. Hence, the TEA of all possible sequences will be evaluated to consider all scenarios and assess the trade-offs. Although this TEA does not include the process cost associated with the process safety equipment needed for each choice solvent, this can be introduced to this framework in future studies.

### Life Cycle Assessment (LCA)

The impact of using each chosen solvent within the multilayer plastic film separation process on the environment and climate (kg CO<sub>2</sub> eq.) will be measured using an LCA methodology. For simplicity, the functional unit considered in the LCA is the production of 1 kg of multilayer plastic film. Thus, we compare the impact of producing 1 kg of ML film using some choice solvents. BioSTEAM will also be used to perform this assessment. The case study will describe more details of the method employed. This framework does not include accessing the full Life Cycle

Assessment (LCA) of each solvent outside the context of the dissolution process. However, it can be incorporated as an additional step, demonstrating the adaptability of this framework.

## RESULTS AND DISCUSSION

As described earlier, STRAP™, a solvent-based process, has been identified as a state-of-the-art technology for separating complex waste plastics containing multiple polymers as components. Moreover, three different variations of this technology have been developed [5]. STRAP™-A is anti-solvent dependent, whereby solvents and antisolvents dissolve and precipitate the polymer layers. STRAP™-B relies on the solubility dependence on temperature whereby after dissolution occurs at an elevated temperature, the polymer solution is cooled to precipitate the polymer. STRAP™-C employs both anti-solvent and temperature to achieve a similar purpose.

Munguía-López et al. [19] quantified the environmental impact of all three variants using the LCA tool to evaluate the carbon footprint, energy use, water use, and toxicity. It was observed that STRAP™-B (temperature dependent) when applied to a multilayer plastic film manufactured by Amcor, had the best metrics among the three variants. This STRAP™-B variant was also compared to the environmental impact of producing virgin films and outperformed fossil-based virgin films on all metrics. Hence, as this work also focuses on driving a sustainable process, STRAP™-B is adopted as a case study to elucidate the strength of this framework. The STRAP™-B process schematic can be seen in the study mentioned above.

This study delineates two distinct scenarios originating from a systematic, bottom-up approach. The first scenario focuses on selecting a green solvent to separate a two-component system comprising EVOH-PET plastic film. Subsequently, the second scenario pertains to selecting a green solvent for separating a three-component system, specifically involving PE-EVOH-PET plastic film. However, in this paper, only the former scenario will be presented. As depicted in Figure 2, the framework's initial phase involves assessing solubility data for all polymers and solvents at a temperature of 120 °C.

### Two-component system (EVOH-PET)

#### Selective dissolution and precipitation

In alignment with the separation sequence outlined by Sánchez-Rivera et al. [5], our objective is to identify green solvents suitable for the selective dissolution of EVOH (ethylene vinyl alcohol) from a composition of EVOH-PET (polyethylene terephthalate) polymer blend. To ensure the targeted dissolution of EVOH, solvents exhibiting a minimum of 10 %wt. difference in solubility between EVOH and PET at 120 °C were retained. This initial

screening resulted in the identification of 116 potential solvents.

Subsequently, we narrowed our selection to solvents amenable to operation at atmospheric pressure, i.e., those possessing boiling points above room temperature. This criterion led to the exclusion of 19 solvents, resulting in a remaining pool of 97 solvents. We then determined each solvent's elevated temperatures (Th) and selected only those exhibiting a minimum of 15 %wt. solubility at Th to ensure the solubility of EVOH. This refinement further reduced the solvent list to 56 candidates. To facilitate subsequent precipitation processes, we further refined our selection by retaining only those solvents with less than 2% weight solubility of EVOH at room temperature (RT). Consequently, the number of viable solvents was reduced to 45.

#### Energy Demand Evaluation

The minimal energy requirements for each of the 45 solvents were calculated using Equation 1. A representation of some solvents' respective minimum energy requirements is presented in Figure 3. The x-axis of Figure 3 denotes solvents, with each letter signifying a specific solvent and its EVOH solubility at Th in brackets. Although higher energy requirement equals higher carbon emissions, solvents were not screened solely based on energy demand. However, they were considered when quantifying the trade-offs amongst all the green metrics.

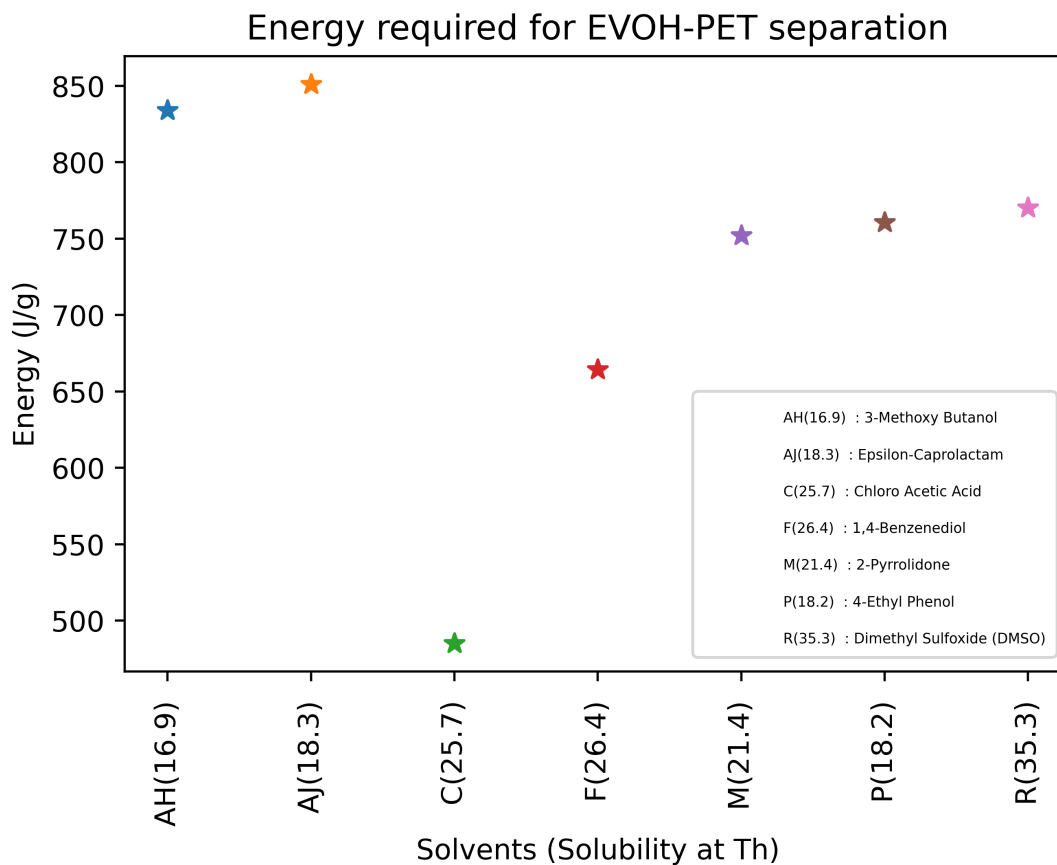
#### Toxicity Evaluation

Partition coefficients were next computed to remove solvents with a LogP greater than 3, indicating a propensity for bioaccumulation. Nevertheless, the solvent count remained at 45, and we performed additional screening steps 6 through 8. Furthermore, we aimed to identify solvents with less environmental concerns and regulations. For example, dimethyl formamide (DMF), ethylene glycol monomethyl ether, N,N-dimethyl acetamide, N-methyl formamide, and N-methyl-2-pyrrolidone all have restrictions placed on their usage by the United States Environment, Health, and Safety (EHS) [11]. However, as earlier mentioned, the GSK solvent guideline does not contain all ~1000 solvents studied in this work. Hence, other solvent guides/restriction lists are queried to screen the selected solvents further.

As shown in Figure 3, dimethyl sulfoxide (DMSO) fulfills the selective solubility criteria of this separation process. This observation agrees with the solvent screening process done by Walker et al. [4], showing that DMSO fulfills the selective solubility criteria without necessarily considering its HSE concerns. However, this study's framework considers solvent toxicity. Furthermore, as

**Table 1:** Selected candidate solvents for EVOH dissolution and its associated hazards

S/N	Solvent Name	Hazard	Flash Point (°C)	Operating Temp (°C)	Melting Temp (°C)	Solubility (% wt.)	LogP (-)	Energy (J/g)	Note
1	4-Ethyl Phenol	Corrosive	104	120	46.0	18.2	2.3	760.5	Used as a flavor in pharmaceuticals.
2	Dimethyl Sulfoxide (DMSO)	Irritant	95	120	18.5	35.3	-1.5	770.3	Used as a pharmaceutical.
3	Epsilon-Caprolactam	Irritant	110	120	70.0	18.3	-0.2	851.1	Used as monomer for plastics and paints.

**Figure 3.** The minimum energy required by selected solvents for EVOH selective dissolution.

human health safety is paramount in any process, solvents that pose human health hazards are strongly discouraged from being considered to avoid trace amounts from contacting food. For example, diethanolamine, Ethylene Glycol Monomethyl Ether, and Hexamethylphosphoramide all pose serious health hazards.

Having considered the various hazards associated with each of the 45 solvents based on the data contained in the GSK solvent guidelines, the PubChem public database, the European Chemical Agency (ECHA), and the Environmental Protection Agency (EPA), three candidate solvents with the least known hazards have been

identified as contained in Table 1.

## CONCLUSION AND FUTURE WORK

This study presented a fast green solvent selection framework for multilayer plastic film separations. This work uses STRAP™-B, a variant of STRAP™, to elucidate the need to employ this holistic framework to guide green solvent selection. This will eliminate the risk of producing new toxic waste streams associated with solvent-based approaches.

The framework was developed by benchmarking



the solubilities of all target polymers in ~1000 solvents. This was done by adopting a similar solubility prediction workflow developed by previous researchers. This framework comprises two key components: selective dissolution and solvent greenness. This selective dissolution was achieved by ensuring a 10 %wt. difference between the target polymers and the non-target polymer(s). Subsequently, for a solvent to be selected, the solubility of the target polymer must be at least 15 %wt. at elevated temperatures, and its boiling point must be higher than room temperature to ensure its operability at standard atmospheric pressure.

The solvent greenness screening involved quantifying the energy requirement of each solvent and predicting the octanol-water partition coefficients (LogP) to eliminate solvents with LogP > 3. Selected candidate solvents were further subjected to the GSK solvent screening guide and regulatory frameworks such as the REACH and the EPA. The framework's final stage encompassed evaluating each candidate solvent's cost and environmental impact in a dissolution-based process. Furthermore, this framework was applied to a case study involving separating a 3-component multilayer plastic film (PE, EVOH, and PET) employing a two-step dissolution process. Additionally, the framework's robustness was demonstrated by its application in informing the process design of dissolution-based processes.

The case study employed a separation sequence of PE-EVOH-PET. Initial consideration focused on selectively dissolving EVOH from an EVOH/PET polymer blend, followed by PE selective dissolution from a PE/EVOH/PET polymer blend. In the first separation (EVOH), only 45 out of ~1000 solvents could dissolve EVOH selectively. Three solvents with the least hazards were identified through a solvent greenness screening process. While DMSO (770 J/g) demonstrated a lower energy requirement than epsilon-caprolactam (851 J/g), both solvents exhibited promising features, including negative LogP values. Additionally, epsilon-caprolactam is a precursor in plastic production.

We have demonstrated the efficacy of the initial seven steps within this framework for identifying environmentally friendly solvents in separating multilayer plastic films. In our immediate future work, we intend to apply this framework to the selective dissolution of PE from a PE/EVOH/PET polymer blend. Additionally, to demonstrate how this framework informs process design, we will apply this framework to select feasible separation sequences based on the availability of green solvents using a similar multilayer plastic film blend.

Furthermore, we intend to evaluate each solvent's Techno-Economic Analysis (TEA) and Life Cycle Assessment (LCA) in the context of the STRAP™ process. This assessment aims to establish the framework as a green solvent selection tool and a valuable resource for

informing process design considerations.

While selective dissolution and LogP screenings offer rapid and automated assessments, solvent selection based on specified guidelines and regulatory frameworks is manual and time-consuming. This underscores the necessity to quantify solvent toxicity numerically, considering the diverse regulations imposed on each solvent. Our future research aims to achieve this, expediting the green solvent selection process.

## ACKNOWLEDGEMENTS

This material is based upon work supported by the US. Department of Energy, Office of Energy Efficiency and Renewable Energy, Bioenergy Technologies Office under the Award number DEEE0009285

## REFERENCES

1. Alias AR, Wan MK, Sarbon NM. Emerging materials and technologies of multi-layer film for food packaging application: A review. *Food Control* 136:108875 (2022).
2. Barry MA. The Science and Technology of Flexible Packaging: Multilayer Films from Resin and Process to End Use. *Elsevier* (2022).
3. Zhao YB, Lv XD, Ni HG. Solvent-based separation and recycling of waste plastics: A review. *Chemosphere* 209:707–720 (2018).
4. Walker TW, Frelka N, Shen Z, Chew AK, Banick J, Grey S, Kim MS, Dumesic JA, Van Lehn RC, Huber GW. Recycling of multilayer plastic packaging materials by solvent-targeted recovery and precipitation. *Sci. Adv.* 6:47 (2020).
5. Sánchez-Rivera KL, Zhou P, Kim MS, González Chávez LD, Grey S, Nelson K, Wang S, Hermans I, Zavala VM, Van Lehn RC, Huber GW. Reducing Antisolvent Use in the STRAP Process by Enabling a Temperature-Controlled Polymer Dissolution and Precipitation for the Recycling of Multilayer Plastic Films. *ChemSusChem* 14:4317–4329 (2021).
6. Zhou P, Yu J, Sánchez-Rivera KL, Huber GW, Van Lehn RC. Large-scale computational polymer solubility predictions and applications to dissolution-based plastic recycling. *Green Chem.* 25:4402–4414 (2023).
7. Winterton N. The green solvent: A critical perspective. *Clean Technol. and Environ. Policy* 23:2499–2522 (2021).
8. Saleh HE-DM and Koller M. Introductory Chapter: Principles of Green Chemistry. *Green Chem.* (2018).
9. Choi YH, Van Spronsen J, Dai Y, Verberne M, Hollmann F, Arends IWCE, Witkamp GJ, Verpoorte R. Are Natural Deep Eutectic Solvents the Missing Link in Understanding Cellular Metabolism and

- Physiology? *Plant Physiol.* 156:1701–1705 (2011).
10. Hessel V, Tran NN, Asrami MR, Tran QD, Van Duc Long N, Escribà-Gelonch M, Tejada JO, Linke S, Sundmacher K. Sustainability of green solvents – review and perspective. *Green Chem.* 24:410–437 (2022).
  11. Henderson RK, Jiménez-González C, Constable DJC, Alston SR, Inglis GGA, Fisher G, Sherwood J, Binks SP, Curzons AD. Expanding GSK’s solvent selection guide – embedding sustainability into solvent selection starting at medicinal chemistry. *Green Chem.* 13:854 (2011).
  12. Prat D, Hayler J, Wells A. (2014). A survey of solvent selection guides. *Green Chem.* 16:4546–4551 (2011).
  13. Diorazio LJ, Hose DRJ, Adlington NK. Toward a More Holistic Framework for Solvent Selection. *Organic Process Research & Development* 20:760–773 (2016).
  14. ECHA. <https://echa.europa.eu/regulations/reach/understanding-reach>.
  15. ECHA. <https://echa.europa.eu/substances-restricted-under-reach>.
  16. EPA. <https://www.epa.gov/enforcement/toxic-substances-control-act-tsca-and-federal-facilities>.
  17. EPA. <https://www.epa.gov/toxics-release-inventory-tri-program/tri-listed-chemicals>.
  18. Cortes-Peña Y, Kumar D, Singh V, Guest JS. BioSTEAM: A Fast and Flexible Platform for the Design, Simulation, and Techno-Economic Analysis of Biorefineries under Uncertainty. *ACS Sustain. Chem. Eng.* 8:3302–3310 (2020).
  19. Munguía-López ADC, Göreke D, Sánchez-Rivera KL, Aguirre-Villegas HA, Avraamidou S, Huber GW, Zavala VM. Quantifying the environmental benefits of a solvent-based separation process for multilayer plastic films. *Green Chem.* 25:1611–1625 (2023).

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Integration of Chemical Looping Reforming and Shift Reactors for Blue H<sub>2</sub> and N<sub>2</sub> Production

Adrian R. Irhamna<sup>a</sup> and George M. Bollas<sup>a\*</sup>

<sup>a</sup> University of Connecticut, Department of Chemical and Biomolecular Engineering, Storrs, CT, USA

\* Corresponding Author: [george.bollas@uconn.edu](mailto:george.bollas@uconn.edu).

## ABSTRACT

Chemical looping Reforming (CLR) is seen as a promising technology for blue hydrogen production. With proper control, CLR in fixed bed reactors has demonstrated the capability to generate blue hydrogen and nitrogen from a single reactor. To enhance efficiency and H<sub>2</sub> purity in the product stream, integration of a CLR reactor with a heat recovery system and a Shift reactor is essential. This study explores the design and control of an integrated CLR-Shift reactors system. The integrated system yields a product stream with 75% H<sub>2</sub> mole fraction during the Reforming step of CLR, and a nitrogen with high purity (98%) during the Oxidation step. In the best-case scenario, the integrated system produces H<sub>2</sub> and N<sub>2</sub> at a molar ratio of 1.26 with H<sub>2</sub> production efficiency of 80.1%.

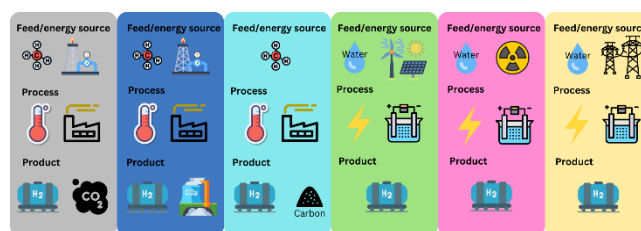
**Keywords:** Chemical-looping reforming, shift reactor, optimal control problem, blue hydrogen and nitrogen production.

## INTRODUCTION

The imperative of decarbonization makes hydrogen one of the top candidates to address the need for future energy carriers. Currently, 95% of hydrogen comes from fossil fuels through conventional steam methane reforming, resulting in emissions of over 12 kg of CO<sub>2</sub> per kg of H<sub>2</sub> [1]. This hydrogen production route is known as grey hydrogen. **Figure 1** provides a visual representation of the hydrogen color spectrum, classifying hydrogen based on production techniques, byproducts, feedstock, and energy source. Green hydrogen, generated through water electrolysis powered by renewable energy sources, is regarded as the most environmentally friendly method of hydrogen production. Despite its environmental advantages, the current cost of green hydrogen is nearly four times that of grey hydrogen [2], posing economic challenges for its utilization. Blue hydrogen emerges as a promising alternative. While it still originates from hydrocarbons, this process requires capturing CO<sub>2</sub>, minimizing its release to the atmosphere. This approach allows for the utilization of established hydrogen production technologies, such as steam methane reforming, with the addition of a CO<sub>2</sub> capture process in the plant. Blue hydrogen is considered a cost-effective

solution during this transitional phase until green hydrogen becomes more economically competitive.

The increase in hydrogen production requires a suitable carrier for large-scale transport and storage. A hydrogen carrier becomes imperative to tackle the inherent challenge of hydrogen's low density. Ammonia is considered as a potential candidate for a hydrogen carrier, offering the ability to store larger quantities of hydrogen in a smaller volume. This route is made possible by capitalizing on the existing global ammonia infrastructure developed for the fertilizer industry [3].



**Figure 1.** Hydrogen color spectrum

Chemical-looping reforming (CLR) is considered as a promising method for blue hydrogen production [4]. This process relies on the capability of oxygen carriers, performing a cycle of reduction, oxidation, and catalytic

reforming reactions. Opting for a fixed bed reactor in CLR proves more advantageous than a fluidized bed reactor due to its simplicity, prolonged oxygen carrier's lifespan, and the absence of a need for gas-solid separation. However, using fixed bed reactors require a dynamic switching operation to complete the CLR cycle. With proper control, CLR in a fixed bed reactor yields an exit stream with a high  $H_2$  concentration during Reforming step, and a high  $N_2$  concentration during Oxidation step [5]. In our previous research, the optimized CLR reactor produced  $H_2$  with an  $H_2/CO$  ratio of 3 during REF and  $N_2$  stream with a purity of 98% during OX [5]. From these results, we identified two key improvements necessary for the CLR reactor. The first enhancement involves the incorporation of a shift reactor to increase the hydrogen concentration in the product stream. The second improvement involves introducing a heat recovery system that utilizes the temperature of the exit gas from CLR as a heat source. This heat recovery system has the potential to substantially improve the overall efficiency of the system.

This study focuses on evaluating the performance of an integrated system comprising a CLR reactor, a shift reactor, and a heat recovery system: a preheater and a steam generator. The main goal is to explore an optimal control strategy for the integrated CLR-Shift reactors system by solving a dynamic optimization problem. The specific objective is to maximize  $H_2$  production while concurrently meeting criteria that guarantee nearly pure  $N_2$  production and ensure the safety and operability of each component within the system. The novelty of this work lies in its integrated approach combining a CLR fixed bed reactor, a heat recovery system, and a shift reactor, facilitating the determination of optimal control parameters and a heat management strategy for highly efficient blue  $H_2$  and  $N_2$ .

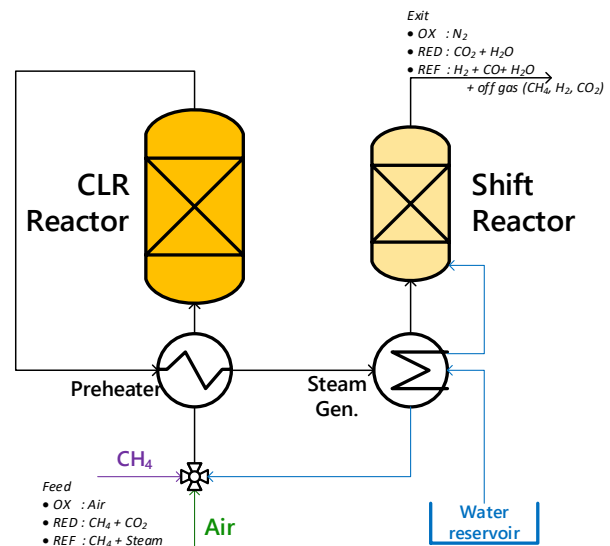
## PROCESS DESCRIPTION

The integrated system consists of a CLR reactor, a shift reactor, a preheater, and a steam generator, as illustrated in the process flow diagram presented in **Figure 2**. In the CLR reactor, CLR cycle takes place performing oxidation (OX), reduction (RED), and reforming (REF) stages alternately. During OX, the oxygen carrier experiences oxidation by air, generating heat within the reactor and leaving  $N_2$  as the primary product in the exit stream. During RED, the oxygen carrier undergoes reduction of the active metal through a gas-solid reaction with the reducing gas (i.e.  $CH_4$ ,  $H_2$ ,  $CO$ ). The exit stream consists of  $CO_2$  and  $H_2O$ , easily separable through condensation for subsequent  $CO_2$  capture and storage. During REF, catalytic reactions occur, resulting in a product stream primarily composed of  $H_2$  and  $CO$ . The heat required for both RED and REF comes from the heat generated during OX. The product stream of CLR reactor in each stage is

directed to the heat recovery system.

The heat recovery system utilizes the high temperature of the CLR product streams as the heat source. Initially, each hot stream enters the preheater and subsequently progresses to the steam generator. Both the preheater and the steam generator operation are synchronized with the ongoing stage in the CLR reactor. In the preheater, the cold feed gas stream of CLR is heated to the target temperature, around  $600\text{ }^\circ\text{C}$ . After the preheater, the product stream is used in the steam generator. In the steam generator, the quantity of steam generated in each CLR stage is tuned based on the enthalpy and temperature difference remaining in the hot stream. At the exit of the steam generator, the temperature of the hot stream is set to be around  $200\text{ }^\circ\text{C}$ .

After the steam generator, the product stream is fed to the Shift reactor. In this reactor, the steam and  $CO$  in the product stream undergo conversion to  $H_2$  through the water gas shift reaction. The reactor bed warms up due to the exothermic nature of the water gas shift reaction. During OX and RED, no reactions occur in the Shift reactor, as the stream cool down the reactor bed, keeping it within the active temperature range of  $200\text{--}350\text{ }^\circ\text{C}$ .



**Figure 2.** The flow diagram integrated system of the reactor

## MODELING

The CLR reactor uses a 1D heterogenous dynamic model of the mass and energy balances for the solid and fluid phases. A summary of the partial differential equations (PDEs) governing the fixed-bed reactor model is presented in **Table 1**. Ni-based oxygen carriers are used in the CLR reactor with reaction kinetics, developed in prior work [6]–[8]. **Table 2** outlines the reactions used in the model, while their corresponding kinetic expressions are provided in [9].

**Table 1:** Summary of the governing equations of the fixed

bed reactors for CLR and Shift and heat recovery system model of the preheater and the steam generator.

### Fixed bed Reactor

#### Fluid phase

Mass Balance:

$$\epsilon_b \frac{\partial C_i}{\partial t} + \frac{\partial u C_i}{\partial z} = \epsilon_b \frac{\partial}{\partial z} \left( D_{ax,i} \epsilon_b \frac{\partial C_i}{\partial z} \right) + k_{c,i} a_v (C_{c,i}|_{Rp} - C_i)$$

Energy Balance:

$$\begin{aligned} \epsilon_b C_{p,f} C_T \frac{\partial T}{\partial t} + C_{p,f} C_T \frac{\partial u T}{\partial z} \\ = \epsilon_b \frac{\partial}{\partial z} \left( \lambda_{ax} \frac{\partial T}{\partial z} \right) + h_f a_v (T_c|_{Rp} - T) \end{aligned}$$

Momentum Balance:

$$\frac{dP}{dz} = - \left( \frac{1 - \epsilon_b}{\epsilon_b^3} \right) \left( \frac{\rho u_o^2}{D_p} \right) \left( \frac{150}{Re_p} + 1.75 \right)$$

Boundary conditions:

$$\epsilon_b D_{ax,i} \frac{\partial C_i}{\partial z} \Big|_{z=0} = u_{in} (C_i|_{z=0} - C_{i,in}),$$

$$\epsilon_b \lambda_{ax} \frac{\partial T}{\partial z} \Big|_{z=0} = u_{in} C_T (T|_{z=0} - T_{in}),$$

$$\frac{\partial C_i}{\partial z} \Big|_{z=L} = \frac{\partial T}{\partial z} \Big|_{z=L} = 0, P|_{z=L} = P_{out},$$

#### Solid phase

Mass Balance:

$$\epsilon_c \frac{\partial C_i}{\partial t} + \frac{1}{r_c^2} \frac{\partial}{\partial r_c} (r_c^2 J_i) = \rho_s \sum R_i$$

Energy Balance:

$$\begin{aligned} \left( (1 - \epsilon_c) \rho_s C_{p,s} + \epsilon_c C_{p,c} C_{T,c} \right) \frac{\partial T_c}{\partial t} \\ = \frac{\lambda_s}{r_c^2} \frac{\partial}{\partial r_c} \left( r_c^2 \frac{\partial T_c}{\partial r_c} \right) + \rho_s \sum (-\Delta H_n) (R_n) \end{aligned}$$

Dusty gas model:

$$-\frac{\partial C_{c,i}}{\partial r_c} = \sum_{j=1}^N \frac{1}{D_{ij}^e} (y_{kj} J_i - y_{ij} J_k) + \frac{J_i}{D_{ik}^e}$$

$$J_i|_{r_c=R_p} = k_{c,i} (C_{c,i}|_{r_c=R_p} - C_i),$$

$$-\lambda_s \left( \frac{\partial T_c}{\partial r_c} \right) \Big|_{r_c=R_p} = h_f (T_c|_{r_c=R_p} - T).$$

### Heat Recovery System

Streams in each side

$$\dot{Q} = m_c c_{p,c} (T_{c,out} - T_{c,in}) = m_h c_{p,h} (T_{h,in} - T_{h,out}),$$

Heat exchanger model

$$\dot{Q} = UA \Delta T_{LMTD}, \Delta T_{LMTD} = \frac{\Delta T_a - \Delta T_b}{\ln(\Delta T_a / \Delta T_b)},$$

Co-current flow

$$\Delta T_a = T_{h,in} - T_{c,in}, \quad \Delta T_b = T_{h,out} - T_{c,out},$$

Counter-current flow

$$\Delta T_a = T_{h,in} - T_{c,out}, \quad \Delta T_b = T_{h,out} - T_{c,in}.$$

Like the CLR reactor, the Shift reactor is modeled using a fixed bed reactor with a 1D heterogenous dynamic model. The corresponding PDEs for the Shift reactor are presented in **Table 1**. Operating at low temperatures (200–350°C), the Shift reactor uses a Cu/ZnO/Al<sub>2</sub>O<sub>3</sub> catalyst for the water-gas shift reaction. The kinetic

correlation for this reaction is adopted from [10], [11] and presented in Eq. (1), where  $r_{s,WGS}$  is the reaction rate of the water gas shift reaction,  $d_{cat}$  and  $\epsilon$  are the diameter and the porosity of the catalyst, respectively,  $P_s$  is the pressure in atm,  $y$  is the species mole fraction,  $K_{eq}$  is the reaction equilibrium described in Eq. (2),  $R$  is the gas constant, and  $T_s$  is the gas temperature in the Shift reactor in K. The feed of the Shift reactor is the product stream from the CLR reactor, which has been cooled to 200°C.

**Table 2:** List of feasible reactions in the reduction, reforming, and oxidation using Ni oxygen carriers.

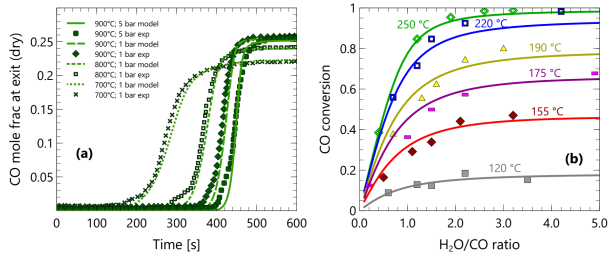
Index	Reactions
(R1)	$H_2 + NiO \rightarrow Ni + H_2O$
(R2)	$CO + NiO \rightarrow Ni + CO_2$
(R3)	$CH_4 + NiO \rightarrow Ni + 2H_2 + CO$
(R4)	$CH_4 + H_2O \leftrightarrow 3H_2 + CO$
(R5)	$CO + H_2O \leftrightarrow H_2 + CO_2$
(R6)	$CH_4 + CO_2 \leftrightarrow 2CO + 2H_2$
(R7)	$CH_4 \leftrightarrow 2H_2 + C$
(R8)	$C + H_2O \leftrightarrow CO + H_2$
(R9)	$C + CO_2 \leftrightarrow 2CO$
(R10)	$O_2 + 2Ni \rightarrow 2NiO$
(R11)	$O_2 + C \rightarrow CO_2$
(R12)	$O_2 + C \rightarrow 2CO$
(R13)	$O_2 + 2CO \rightarrow 2CO_2$

$$\begin{aligned} r_{s,WGS} = d_{cat} (1 - \epsilon) P_s^{(0.5 - P_s/250)} \left( 2.96 \right. \\ \left. \times 10^5 \frac{\text{mol}}{\text{g} \cdot \text{hr}} \right) \exp \left( \frac{-47400 \text{ J/mol}}{RT_s} \right) \\ \times \left( y_{CO} y_{H_2O} - \frac{y_{CO_2} y_{H_2}}{K_{eq}} \right) \end{aligned} \quad (1)$$

$$\begin{aligned} K_{eq} = \exp \left( \frac{5693.5}{T_s} + 1.077 \ln(T_s) + 5.44 \times 10^{-4} T_s \right. \\ \left. - 1.125 \times 10^{-7} T_s - \frac{49170}{T_s^2} - 13.148 \right) \end{aligned} \quad (2)$$

The energy balance in the preheater and the steam generator are summarized in **Table 1**. The momentum balance, pressure drop, and heat loss in both the preheater and the steam generator were neglected. The models for the CLR and Shift reactors were validated with available experimental data and are illustrated in **Figure 3**.





**Figure 3.** Model validation of CLR (a) and Shift (b) reactors using experimental data reported in [12] and [11] respectively

## OPTIMIZATION

The operational challenges associated with a fixed-bed reactor stem mainly from its semi-batch operation. In the fixed bed reactor, the oxygen carrier remains in the reactor while the feed alternates serving CLR stages and complete the cycle. These reactions are kinetically controlled and exhibit different reaction enthalpies. Consequently, the bed temperature and the state of the oxygen carrier experience significant changes during the transition between CLR stages. Each CLR stage produces different gas products, resulting in downstream discontinuities. For  $H_2$  production, the steam-to-methane ratio during REF plays an important role in determining the production of  $H_2$  and in managing carbon deposition. The integration with the heat recovery system and the Shift reactor requires an effective control strategy to ensure the heat generated during OX is appropriately distributed for reactions occurring in the CLR, the heat recovery system, and the Shift reactor. The primary challenge lies in maintaining a high  $H_2$  production while consistently retaining sufficient heat within the integrated system and meeting constraints related to fuel conversion, high  $N_2$  concentration, the energy-temperature difference required for feed gas preheating, steam generation, and the Shift reactor.

To achieve this objective, we formulated an optimal control strategy for the decision variables for the integrated CLR-Shift system. The focus was on maximizing hydrogen production while consistently meeting the dynamic constraints of the system. The decision variables chosen were the time interval of the CLR stages: RED, REF, and OX ( $\tau_{OX}$ ,  $\tau_{RED}$ , and  $\tau_{REF}$ ), feed gas temperatures, steam-to-methane ratio, the percentage of active metal in oxygen carrier, and the steam added to the Shift reactor, as summarized in **Table 3**. The control profile for the feed gas was modeled using piecewise constant functions, denoted as  $\mathbf{u}(\tau_i) = \mathbf{u}$ , where  $\mathbf{u}$  is the vector of temperature, flow, and composition of the gas stream, while  $\tau_i$  is the time duration of the  $i$ -th CLR step, i.e., OX, RED, and REF. The design vector,  $\phi$ , summarizes the set of control variables, and was constrained by upper and lower limits defined within the design space,  $\Phi$ , as presented in Eq. (3).

**Table 3:** Design variables for the optimal control of the integrated CLR-Shift reactors system

Control variables	Notation
Feed gas temperature	$\mathbf{u}_i$
Steam-to-methane ratio @REF	
Methane-to- $CO_2$ ratio @RED	
Steam added into the Shift reactor @REF	
Time interval of OX	$\tau_{OX}$
Time interval of RED	$\tau_{RED}$
Time interval of REF	$\tau_{REF}$
Active metal content in OC	$\omega$

The objective function of the optimal control problem is to maximize hydrogen production during the cyclic steady-state operation of the integrated system. The objective function uses a metric called the hydrogen production efficiency,  $\eta$ , shown in Eq. (4), where  $F_{S-out,H_2}$  is the hydrogen flowrate at the exit of Shift reactor,  $F_{R-in,CH_4}$  is the methane flowrate at the inlet of CLR Reactor,  $t_0$  and  $t_f$  are the initial time and final time of integration, with  $t_f = \tau_{OX} + \tau_{RED} + \tau_{REF}$ .

$$\phi = [\mathbf{u}_i, \tau_{OX}, \tau_{RED}, \tau_{REF}, \omega, F_{S-in,H_2O}] \in \Phi. \quad (3)$$

$$\eta = \frac{\int_{t_0}^{t_f} F_{S-out,H_2}(t) dt}{\int_{t_0}^{t_f} F_{R-in,CH_4}(t) dt}, \quad (4)$$

The optimization of the integrated system is subject to the following constraints. A minimum methane conversion ( $X_{CH_4}$ ) of 96% in the CLR reactor must be achieved during both RED and REF (Eq. (5)). During OX, the CLR reactor must yield a  $N_2$  stream with 98% minimum mole fraction, as shown in Eq. (6), where  $F_{R-out,N_2}$  is the  $N_2$  molar flowrate at the exit of CLR reactor and  $F_{R-out,T}$  is the total flowrate at the exit of CLR reactor. Since the product stream of the CLR reactor is used in the heat recovery system, the total enthalpy carried by the product stream must exceed the enthalpy required for gas preheating and steam generation, Eq. (7), where,  $F_{R-in,T}$  is the total flowrate at the inlet of the CLR reactor,  $c_p$  is the heat capacity of the inlet/outlet streams,  $h_{fg,H_2O}$  is the enthalpy of evaporation of water,  $T_{R-in}$  and  $T_{R-out}$  are the inlet and outlet temperatures at the CLR reactor, respectively. In Eq. (7), it is assumed that in the heat recovery system product streams are cooled to 200 °C, while feed streams are heated up from 25°C.  $F_{R-in,H_2O}$  and  $F_{S-in,H_2O}$  are the molar flowrate of steam fed into CLR reactor and Shift reactor, respectively. To prevent agglomeration and sintering of the Ni-based oxygen carrier, the maximum temperature allowed in CLR reactor,  $T_{R-bed}$ , at any axial location  $z$ , must not exceed 1100°C, (Eq. (8)). The preheater must be able to heat up the cold stream,  $T_{PR-c,out}$ , to between 500 and 700°C, the desired temperature for feed gas of the CLR reactor (Eq. (9)). In the steam

generator, the exit temperature of the hot stream,  $T_{SG-h,out}$ , must be between 200 and 250°C, close to the operation temperature of the Shift reactor as in Eq. (10).

The complete optimal control problem can be formulated as shown in Eq. (11), where  $\mathbf{f}$  is the set of differential-algebraic equations (DAEs) that describe the model of the reactors in the integrated CLR-Shift system, comprising mass, energy, and momentum balance and reaction kinetics,  $\mathbf{x}$  is the vector of state variables (*i.e.*, mass, temperature, and pressure), and  $\dot{\mathbf{x}}$  is time derivatives of  $\mathbf{x}$ . Cyclic steady-state conditions are typically achieved after more than two cycles. Therefore, the optimization time horizon was set to at least two times the  $\tau_{\text{cycle}}$ .

$$X_{\text{CH}_4}(t) = 1 - \frac{\int_{t_0}^{t_f} F_{R-out,\text{CH}_4}(t) dt}{\int_{t_0}^{t_f} F_{R-in,\text{CH}_4}(t) dt} \geq 96\%, \quad (5)$$

$$\beta(t) = \frac{\int_{t_0}^{t_f} F_{R-out,\text{N}_2}(t) dt}{\int_{t_0}^{t_f} F_{R-out,T}(t) dt} \geq 98\%, \quad (6)$$

$$\begin{aligned} \gamma(t) = & \int_{t_0}^{t_f} F_{R-out,T}(t) c_{p,out}(T_{R-out}(t) - 200^\circ\text{C}) dt \\ & - \int_{t_0}^{t_f} F_{R-in,T}(t) c_{p,in}(T_{R-in}(t) - 25^\circ\text{C}) dt \\ & - \int_{t_0}^{\tau_{\text{REF}}} (F_{R-in,\text{H}_2\text{O}}(t) + F_{S-in,\text{H}_2\text{O}}(t)) h_{fg,\text{H}_2\text{O}} dt \geq 0, \end{aligned} \quad (7)$$

$$T_{R-bed}(z) \leq 1100^\circ\text{C}, \quad (8)$$

$$500^\circ\text{C} \leq T_{PR-c,out}(t) \leq 700^\circ\text{C}, \quad (9)$$

$$200^\circ\text{C} \leq T_{SG-h,out}(t) \leq 250^\circ\text{C}, \quad (10)$$

$$\begin{aligned} & \max_{\phi \in \Phi} \eta \\ & \text{subject to:} \\ & \mathbf{f}(\dot{\mathbf{x}}(t), \mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}, t) = 0, \\ & \mathbf{f}_0(\dot{\mathbf{x}}(t_0), \mathbf{x}(t_0), \mathbf{u}(t_0), \boldsymbol{\theta}, t_0) = 0, \quad (11) \\ & \text{Eqs. (5) - (10),} \\ & \mathbf{x}^{\min} \leq \mathbf{x}(t) \leq \mathbf{x}^{\max}, \\ & \mathbf{u}_i^{\min} \leq \mathbf{u}_i(t) \leq \mathbf{u}_i^{\max}, \\ & \boldsymbol{\tau}_i^{\min} \leq \boldsymbol{\tau}_i(t) \leq \boldsymbol{\tau}_i^{\max}, \\ & \boldsymbol{\omega}^{\min} \leq \boldsymbol{\omega} \leq \boldsymbol{\omega}^{\max}. \end{aligned}$$

The system of DAEs were built and solved in the equation-based process modeling platform gPROMS Model Builder, 7.0.9 [13]. For the spatial discretization, the backward finite difference method was used, dividing the reactor's axial direction into 21 nodes and its solid particle radial direction into 11 nodes. The system of DAEs was then solved with the DAEBDF solver. The non-linear optimization was solved with control vector parameterization with single shooting (CVP\_SS) solver available in gPROMS. With control vector parameterization, the control variables were discretized as piecewise constant over a specified time interval, while with the single shooting approach, the control variables were fixed during the entire respective time horizon in each iteration.

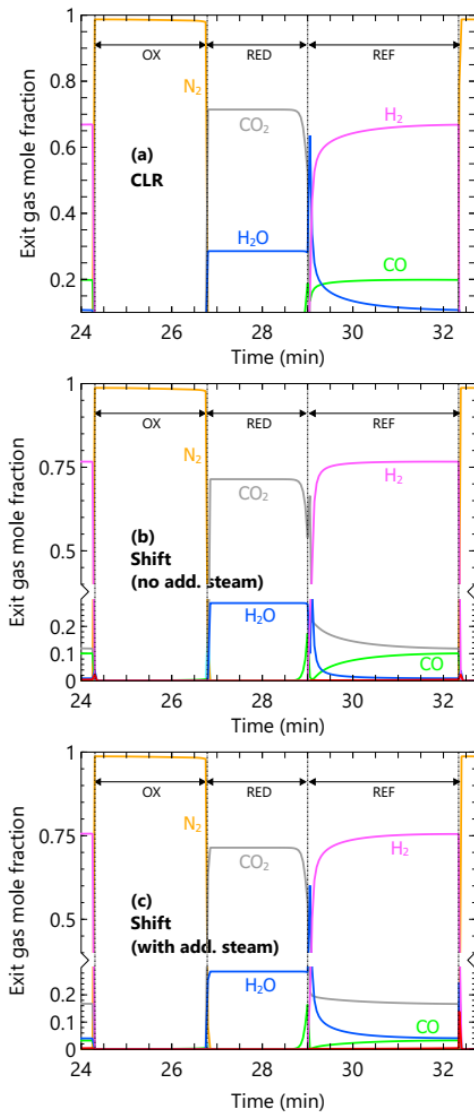
## RESULTS AND DISCUSSION

The integrated CLR-Shift system was used to study its performance for small-scale blue hydrogen and nitrogen production. The system was configured to run in parallel with multiple CLR reactors, collectively achieving a production rate of 300 kg-H<sub>2</sub> per day. **Table 4** (top rows) reports the reactor design and operating parameters, including reactor diameter, length, and operating pressure. Eq. (11) was solved, resulting in the optimized control variables reported in **Table 4** (bottom rows). These parameters were then used in simulations to evaluate the performance of the integrated CLR-Shift system.

**Table 4:** Parameters used in the integrated CLR-Shift system. The top rows are the common parameters, while the bottom rows show the optimized parameters.

Parameters	Values		
	OX	RED	REF
CLR Reactor			
- length, [m]		1.5	
- diameter, [m]		0.6	
Shift Reactor			
- length, [m]		0.6	
- diameter, [m]		0.6	
Operating pressure, [bar]		1.5	
Feed vol. flow, [NCMH]	360	36	180
CO <sub>2</sub> -to-CH <sub>4</sub> ratio	--	4.0	--
Steam-to-CH <sub>4</sub> ratio	--	--	1.5
Feed temp., [°C]	600	600	600
CLR stages interval, [s]	150	135	200
Active metal in OC, [%]		5	
Add. steam @ Shift, [mol/s]	--	--	0.81

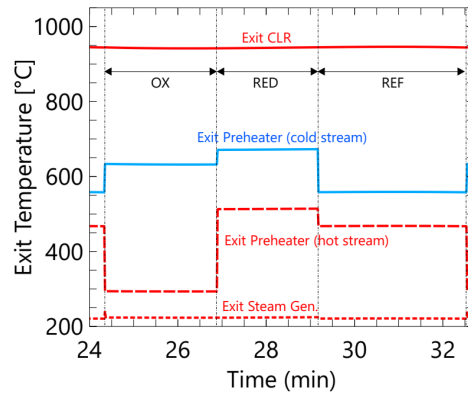
**Figure 4** presents the exit mole fractions of CLR and Shift reactors at cyclic steady state. At the optimal operating conditions, the exit stream of the CLR reactor (**Figure 4a**) exhibits compositions of 98% N<sub>2</sub> during OX, 70% CO<sub>2</sub> and 30% steam during RED, and 66% H<sub>2</sub> during REF. The N<sub>2</sub> stream with high purity during OX is promising for the further integration with an ammonia plant. In RED, the exit stream, consist of CO<sub>2</sub> and H<sub>2</sub>O, can be separated by a condensing unit before captured or recirculated for plant utility. The 66% H<sub>2</sub> during REF showcases the feasibility of CLR as one promising route for blue hydrogen production. In REF, there is also smaller gas constituents consisting of CO (20%), H<sub>2</sub>O (10%), and CO<sub>2</sub> (2%) that need further purification.



**Figure 4.** The performance of CLR and Shift reactor during cyclic steady-state, simulated at optimized parameters (Table 4). (a) Exit mole fraction at CLR Reactor; (b) exit mole fraction at Shift reactor without additional steam (Case 1); and (c) exit mole fraction at Shift reactor with additional steam (Case 2).

The product stream from CLR goes to the heat recovery system and serves as a heat source. **Figure 5** shows the temperature of the CLR product stream decreasing as it passes through the preheater and the steam generator, respectively. From the CLR reactor, the temperature of product stream is around 950°C. As the product stream exits the preheater, the temperature drops to 300–500°C. The exit temperature at the preheater varies in each CLR stage depending on the enthalpy remaining and temperature difference between hot and cold streams. Nonetheless, the CLR feed gas is successfully preheated to 500–700°C. The product

stream then enters the steam generator producing steam needed for the reforming reactions. The steam generator operates in each CLR stage, using the remaining available enthalpy and temperature in the product stream as a heat source. The portion of the steam generated in each stage was tuned so that the exit temperature of the product stream in the steam generator was close to 200 °C (**Figure 5**).



**Figure 5.** Exit gas temperature in the preheating and steam generating systems. The exit stream of CLR reactor is used as the heat source for the preheating and steam generating systems. The exit preheater cold stream feeds the CLR, while the hot stream from the preheater is used for steam generation.

We explored two scenarios for the Shift reactor: 1) Shift reactor without additional steam (Case 1), and 2) Shift reactor with additional steam (Case 2). In Case 1, the feed of the Shift reactor is the CLR product stream that has been cooled in the heat recovery system. In Case 2, an optimized amount of steam (**Table 4**) is added to the feed during REF. **Figures 4b** and **4c** present the result of Case 1 and Case 2, respectively. In both cases, the Shift reactor increased H<sub>2</sub> mole fraction from 66% to 75–76%. In Case 1, CO conversion was low, only 49.2%, due to the limited steam available for water gas shift reaction. At the exit, around 10% of CO remain unconverted and the H<sub>2</sub> production rate was 3 mol/s. In Case 2, the steam addition increased the extent of the water gas shift reaction and increased CO conversion to 82.7%. At the exit, H<sub>2</sub> production rate was 3.2 mol/s (75.6% of the total stream), with smaller gas constituents consisting of: CO<sub>2</sub> (16.7%), H<sub>2</sub>O (4%), and CO (3.2%).

Overall, the integrated CLR-Shift reactors system successfully generated H<sub>2</sub> and N<sub>2</sub>, with high concentration and efficiency. In Case 1, H<sub>2</sub> and N<sub>2</sub> were produced with a ratio of 1.16 and an H<sub>2</sub> production efficiency of 74.9%. Meanwhile, Case 2 yielded H<sub>2</sub> and N<sub>2</sub> with a ratio of 1.26 and an H<sub>2</sub> production efficiency of 80.1%. Furthermore, the dynamic operation of this integrated system suggest the possibility of integration with an ammonia synthesis process reported in [14].

## CONCLUSIONS

Integration of the CLR reactor, Shift reactor, and heat recovery system demonstrated promising results for blue hydrogen and nitrogen production. An optimized control scheme is crucial to ensure high H<sub>2</sub> production that satisfies the numerous system constraints. Under the optimized control scenario, the integrated CLR-Shift reactors system produced H<sub>2</sub> with a 75% mole fraction during REF, N<sub>2</sub> with 98% purity during OX, and a potential for CO<sub>2</sub> capture during RED. The control scenario is also essential in distributing the heat within the integrated system, meeting the heat requirements for overall reactions, the feed gas preheating and steam generation. The addition of steam in the Shift reactor is essential to maintain steam availability for the water gas shift reaction. Further purification in the downstream product, such as pressure-swing adsorption, must be considered to achieve higher H<sub>2</sub> purity. A future study on multiple reforming reactor system and reactor scheduling will target a continuous flow of hydrogen and nitrogen products.

## ACKNOWLEDGEMENTS

This work was supported by a DIKTI-funded Fulbright Fellowship and the Pratt & Whitney Institute of Advanced Systems Engineering (P&W-IASE) of the University of Connecticut. Any opinions expressed herein are those of the author and do not represent those of the sponsor.

## REFERENCES

1. IEA, "Towards hydrogen definitions based on their emissions intensity," 2023. [Online]. Available: [www.iea.org](http://www.iea.org).
2. J. F. George, V. P. Müller, J. Winkler, and M. Ragwitz, "Is blue hydrogen a bridging technology? - The limits of a CO<sub>2</sub> price and the role of state-induced price components for green hydrogen production in Germany," *Energy Policy*, vol. 167, Aug. 2022, doi: 10.1016/j.enpol.2022.113072.
3. J. Andersson and S. Grönkvist, "Large-scale storage of hydrogen," *Int. J. Hydrogen Energy*, vol. 44, no. 23, pp. 11901–11919, 2019, doi: 10.1016/j.ijhydene.2019.03.063.
4. C. A. Del Pozo, S. Cloete, Á. J. Álvaro, F. Donat, and S. Amini, "The potential of gas switching partial oxidation using advanced oxygen carriers for efficient h<sub>2</sub> production with inherent co<sub>2</sub> capture," *Appl. Sci.*, vol. 11, no. 10, May 2021, doi: 10.3390/app11104713.
5. A. R. Irhamna and G. M. Bolas, "Process intensification in a fixed bed reactor for a small-scale process in the stranded assets," in *33rd*

*European Symposium on Computer Aided Process Engineering*, 2023, pp. 3043–3048.

6. Z. Zhou, L. Han, and G. M. Bolas, "Model-based analysis of bench-scale fixed-bed units for chemical-looping combustion," *Chem. Eng. J.*, vol. 233, pp. 331–348, 2013, doi: 10.1016/j.cej.2013.08.025.
7. Z. Zhou, L. Han, O. Nordness, and G. M. Bolas, "Continuous regime of chemical-looping combustion (CLC) and chemical-looping with oxygen uncoupling (CLOU) reactivity of CuO oxygen carriers," *Appl. Catal. B Environ.*, vol. 166–167, pp. 132–144, 2015, doi: 10.1016/j.apcatb.2014.10.067.
8. O. Nordness, L. Han, Z. Zhou, and G. M. Bolas, "High-Pressure Chemical-Looping of Methane and Synthesis Gas with Ni and Cu Oxygen Carriers," *Energy and Fuels*, vol. 30, no. 1, pp. 504–514, 2016, doi: 10.1021/acs.energyfuels.5b01986.
9. A. R. Irhamna and G. M. Bolas, "Intensified reactor for lean methane emissions treatment," *AIChE J.*, pp. 1–40, 2023, doi: 10.1002/aic.18040.
10. T. A. Adams and P. I. Barton, "A dynamic two-dimensional heterogeneous model for water gas shift reactors," *Int. J. Hydrogen Energy*, vol. 34, no. 21, pp. 8877–8891, 2009, doi: 10.1016/j.ijhydene.2009.08.045.
11. Y. Choi and H. G. Stenger, "Water gas shift reaction kinetics and reactor modeling for fuel cell grade hydrogen," *J. Power Sources*, vol. 124, no. 2, pp. 432–439, 2003, doi: 10.1016/S0378-7753(03)00614-1.
12. P. Alexandros Argyris *et al.*, "Chemical looping reforming for syngas generation at real process conditions in packed bed reactors: An experimental demonstration," *Chem. Eng. J.*, vol. 435, May 2022, doi: 10.1016/j.cej.2022.134883.
13. Process Systems Enterprise, "gPROMS." 2021, [Online]. Available: [www.psententerprise.com/products/gproms](http://www.psententerprise.com/products/gproms).
14. L. Burrows, P. X. Gao, and G. M. Bolas, "Thermodynamic feasibility analysis of distributed chemical looping ammonia synthesis," *Chem. Eng. J.*, vol. 426, no. May, p. 131421, 2021, doi: 10.1016/j.cej.2021.131421.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Model assessment for Design of Future Manufacturing systems using Digital Twins: A case study on a single-scale pharmaceutical manufacturing unit

Prem Jagadeesan<sup>a</sup> and Shweta Singh<sup>a,b,c\*</sup>

<sup>a</sup> Department of Agricultural and Biological Engineering, Purdue University, West Lafayette, IN, US

<sup>b</sup> Division of Ecological and Environmental Engineering, Purdue University, West Lafayette, IN, US

<sup>c</sup> Davidson School of Chemical Engineering (By Courtesy), Purdue University, West Lafayette, IN, US

\* Corresponding Author: [singh294@purdue.edu](mailto:singh294@purdue.edu).

## ABSTRACT

Designing a digital twin will be crucial in developing automation-based future manufacturing systems. The design of digital twins involves data-driven modelling of individual manufacturing units and interactions between the various entities. The goals of future manufacturing units such as zero waste at the plant scale can be formulated as a model-based optimal control problem by identifying the necessary state, control inputs, and manipulated variables. The fundamental assumption of any model-based control scheme is the availability of a “reasonable model”, and hence, assessing the goodness of the model in terms of stability and sensitivity around the optimal parameter value becomes imperative. This work analyses the data-driven model of an acetaminophen production plant obtained from SINDy, a nonlinear system identification algorithm using sparse identification techniques. Initially, we linearize the system around optimal parameter values and use local stability analysis to assess the stability of the identified model. Further, we use what is known as a conditional sloppiness analysis to identify the sensitivity of the parameters around the optimal parameter values to non-infinitesimal perturbations. The conditional sloppiness analysis will reveal the geometry of the parameter space around the optimal parameter values. This analysis eventually gives valuable information on the robustness of the predictions to the changes in the parameter values. We also identify sensitive and insensitive parameter direction. Finally, we show using numerical simulations that the linearized SINDy model is not good enough for control system design. The pole-placement controller is not robust, and with high probability, the control system becomes unstable to very minimum parameter uncertainty in the gain matrix.

**Keywords:** Dynamic Modelling, System Identification, Stability, Sloppiness, Identifiability

## INTRODUCTION

Design of future manufacturing systems will benefit from building digital twins that can inform the controller design at plant scale. Several future manufacturing goals such as planning and scheduling for zero waste at the plant scale can be formulated as a control system design problem, where an optimal control problem can be formulated and solved.

The design of optimal control mandates a reasonable model. Developing a mechanistic model at the plant scale might involve hundreds of state variables and parameters; hence we adopt a data-driven nonlinear

system identification approach known as SINDy to identify a reduced-order parsimonious model in this work. SINDy uses the idea of sparse identification to discover the underlying governing equations [2]. The SINDy algorithm has already been used to design an entire algal biodiesel industrial network for sustainable design of carbon capture and utilization technologies [3]. In addition to this, SINDy has also been used to identify governing equations of unit operations in a plant and natural systems [7]. In previous work, the authors have used the SINDy algorithm to identify a reduced order dynamical model for the distillation column [6].

The models developed in the above works were



satisfactory in replicating the system dynamics; however, the model has to satisfy additional requirements for designing control systems. The critical factors that affect the control system design are stability, model and parameter uncertainty. The seminal work of Bhattacharyya et al [5] showed that while designing robust and optimal controllers such as  $H_2$  and  $H_\infty$ ; a very small parameter uncertainty in the controller parameters will result in an unstable control system. They argue that the fragility of the controller is a result of the parameter sensitivity of the plant that is transferred to the controller. Hence, it becomes imperative to assess the sensitivity of the plant before the controller design; otherwise, uncertainty in the controller output will impact the overall performance metrics.

In this work, we propose a method to analyze the control relevance of the SINDy model; even though this method is applied to models developed from SINDy, in general, this proposed method can be used for any surrogate dynamical model. The novelty in the present work is proposing a new method to assess the robustness of a pole-placement control system design by assessing the model's parameter space by introducing the concept of conditional sloppiness. In general, it is a novel method to assess whether the estimated model is control-relevant. As a first step, we analyze the stability of the linearized model around the operating point. We use sloppiness analysis as a next step to characterize the model's behaviour around the nominal parameter set. Sloppiness is a phenomenon where there are regions in the parameter space over which model predictions are nearly identical. The role of sloppiness in system identification has been extensively studied in the past two decades [1,4]. Here, we use what is known as conditional sloppiness to assess the model sensitivity. Together with stability and sloppiness analysis we provide directions to refine the estimated model. When the model sensitivity index is very high in the vicinity of the parameter space, the model is unsuitable for controller design. This has been demonstrated using a simple pole-placement controller design.

The rest of the paper is organized as follows: Section 2 presents the preliminary concepts, Section 3 illustrates the detailed methodology for assessing the model structure, we assess the goodness of the model of a pharmaceutical node (single plant) in Section 4, and the paper ends with some concluding remarks in Section 5.

## PRILIMINARIES

### Local Stability Analysis

Analysing the stability of the model plays a crucial role in designing the controllers. The dynamics of industrial systems are predominantly represented by a set of first-order nonlinear differential equations, these are known as state-space models. The set of solution to

these nonlinear differential equations are known as state trajectories. A system is completely characterized by the values of the state variables at any given instance of time. The generic nonlinear model representation is given below

$$M: \begin{cases} \frac{dx(t)}{dt} = f(x(t), u(t), \theta) \\ y(t) = g(x(t), u(t), \theta) \end{cases} \quad (1)$$

where,  $x(t)$  is the state vector,  $u(t)$  is the input vector and  $\theta$  is the parameter vector. Even though the dynamics is nonlinear, in most cases it can be approximated to a linear dynamics around the operating region. Hence, in this work we use linear stability analysis on the linearized state space model of Eqn 1. The linearized model is given in Eqn 2.

$$\tilde{M}: \begin{cases} \dot{x} = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases} \quad (2)$$

where A, B, C and D are state-space matrices obtained by evaluating Jacobin of Eq (1) at the operating point  $(x^*, u^*)$ . The eigenvalues of the A matrix will indicate the stability of the linear perturbation system (2).

### Sloppiness

In models with nonlinear predictors, often there are large regions in the parameter space over which the model predictions are nearly identical, this is known as sloppiness or model sloppiness. For infinitesimal perturbations sloppiness is quantified by the inverse of the condition number of the Hessian of the cost-function. The Hessian of the cost function can be approximated as

$$H_{ij} = \frac{1}{N} \sum_{n=1}^N \frac{\partial y}{\partial \log \theta_i} \frac{\partial y}{\partial \log \theta_j} \quad (3)$$

where  $y$  denotes model output. More formally, for non-infinitesimal perturbations, sloppiness can be conditioned on the experiment space known as conditional sloppiness. A model  $\mathcal{M}$  is conditionally  $(\epsilon, \delta)$  sloppy with respect to an experiment space  $\mathcal{Z}_{\mathcal{M}}$  at  $\theta^* \in \mathcal{J} \subset \mathcal{D}_{\mathcal{M}}$ , if

$$\|\theta^* - \theta_1\|_2 > \delta \quad \forall \theta \in \mathcal{S} \subset \mathcal{J} \quad (4)$$

$$\|y(\theta^*, t) - y(\theta_1, t)\|_2^2 < \epsilon \quad \forall u \in \mathcal{Z} \subset \mathcal{Z}_{\mathcal{M}} \quad (5)$$

for every  $(\theta_1, \theta^*)$  satisfying (4) and (5).  $\epsilon$  is arbitrarily small.  $\delta \gg \epsilon$ .

The role of sloppiness in the system identification has been extensively studied in. Sloppiness often affects the uncertainty in the parameter estimates. it is well known the uncertainty in the model structure and parameter estimates affects the robustness of the controller and hence, assessing the parameter uncertainty of the model becomes imperative to a satisfactory control system design. In this work, we use

conditional sloppiness analysis to assess the insensitivity in the parameter directions in SINDy models.

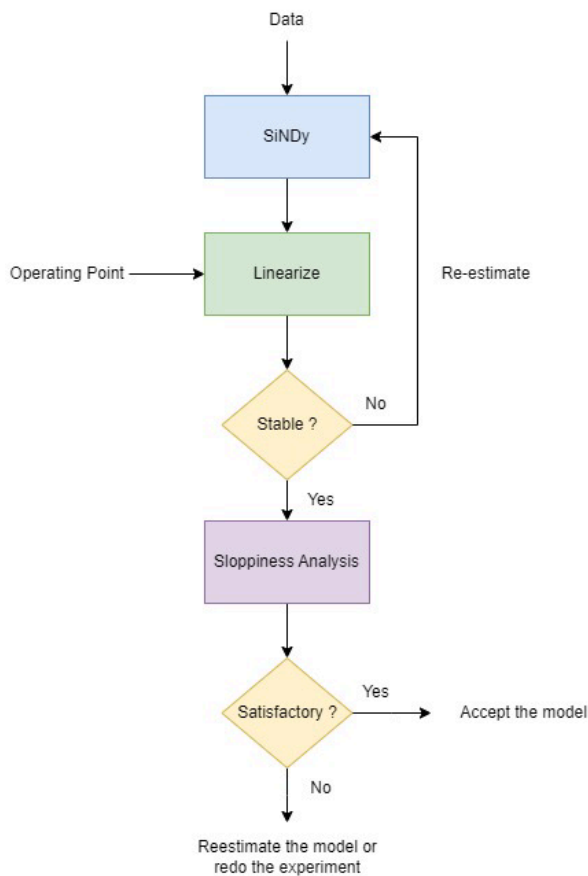


Figure 1: Algorithm for model assessment.

## METHODOLOGY

This section illustrates the method proposed for assessing the goodness of the model estimated using the SINDy algorithm. However, the method proposed, in general, can be used to evaluate any model. The proposed method has two significant steps: Firstly, assessing the stability of the local perturbation model (Linearized) and model sloppiness. As a next step, we construct what is known as a  $(\delta - \gamma)$  plot as proposed in [1] for conditional sloppiness analysis around the optimal parameter estimated from the SINDy algorithm. The central idea of this analysis is to characterize the model's behaviour around the point of interest in the parameter space; this is done by constructing an n-ball and evaluating the change in the model's output for all the parameters inside the n-ball with respect to the reference point. The procedure for constructing the  $(\delta - \gamma)$  plot is in Figure 2.

One of the main advantages of the proposed method is its ability to identify the parameters that are likely to be estimated with poor precision. When those parameters belong to stiff region/sensitive region, then with high probability, the controller system designed may become fragile. In the next section, we demonstrate the working of the proposed method in a dynamical model of an industrial node developed from the SINDy algorithm.

## Numerical Results

In this section, we demonstrate the working of the proposed method in a model developed for acetaminophen production plant.

### A linearized state-space model of acetaminophen production network

To demonstrate the working of the proposed method, we consider a process where Para-

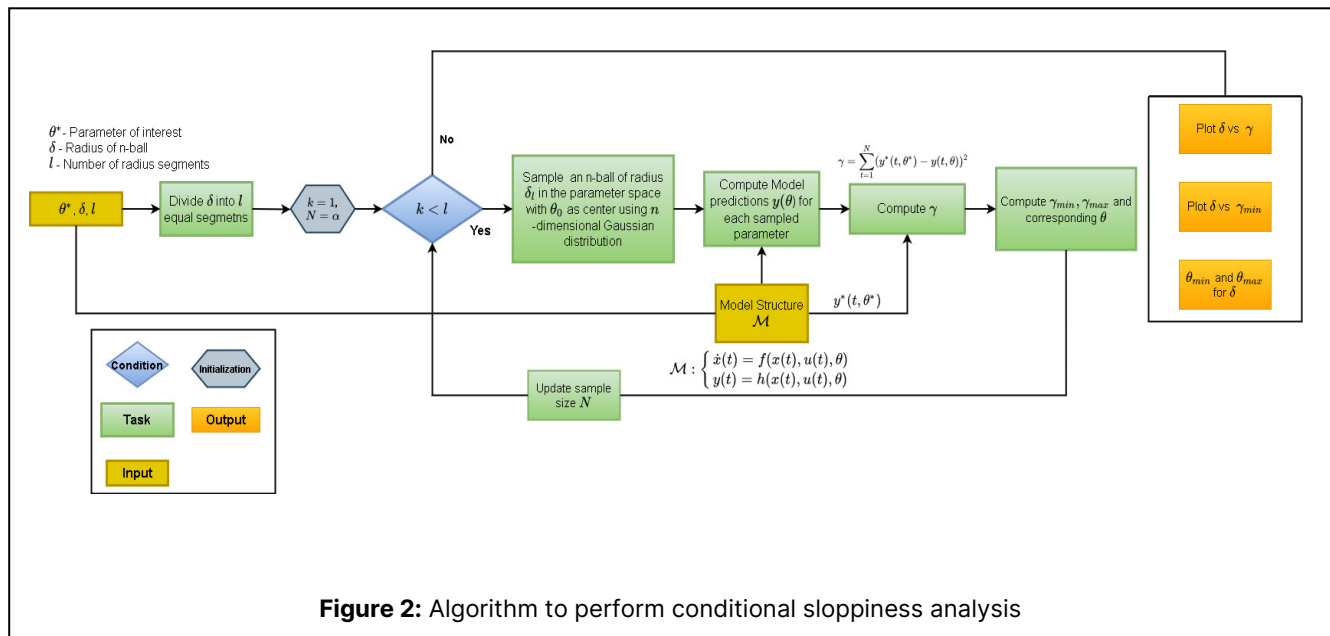


Figure 2: Algorithm to perform conditional sloppiness analysis

Aminophenol(PAP) reacts with Acetic Anhydride to produce Acetaminophen (A-PAP). Equation 7 is the linearized model obtained from the non-linear SINDy model. The detailed state variables and input variables are given in Table 1. The nonlinear ODE model identified using SINDy algorithm is linearized around the operating point. The Operating point and the linearized model is given in the Table 1 and equation 4

$$\dot{x}(t) = \begin{bmatrix} -16.34 & -5.06 & -15.83 & 9.04 \\ 12.22 & -3.41 & 10.72 & 3.43 \\ -0.07 & 7.85 & -1.10 & -5.63 \\ 3.70 & 0.006 & 2.97 & -3.90 \end{bmatrix} x(t) +$$

$$\begin{bmatrix} 11.15 & 3.21 & 4.39 & -7.56 \\ 2.84 & -0.62 & 2.38 & 7.58 \\ 0.76 & 2.78 & -1.91 & -3.52 \\ -0.03 & 4.80 & 4.36 & -0.96 \end{bmatrix} u(t) \quad (6)$$

$$y(t) = [1 \ 0 \ 0 \ 0]x(t) \quad (7)$$

where  $x_1(t)$  is APAP,  $x_2(t)$ ,  $x_3(t)$  and  $x_4(t)$  are the waste materials produced and  $u_1(t)$  is PAP.Fm,  $u_2(t)$  is Acetic Anhydride.Fm,  $u_3(t)$  is water and  $u_4(t)$  is water

The Eigen values of the system matrix A are given below

$$\lambda(A) = [-19.27 \quad -5.28 \quad -0.1 + 3.43i \quad -0.1 - 3.43i]$$

The system has both real and complex poles with negative real parts, which guarantees that the system is capable of damped oscillations. It is also worth noting that the complex poles have real parts that are close to zero value, which says that the system is on the verge of instability. However, the system is in the verge of instability. In the next section we perform conditional sloppiness analysis to study the sensitivity of the system to the parameter perturbations.

### Conditional Sloppiness Analysis

In this section, we do the conditional sloppiness analysis for the unforced system, i.e. the input is turned off in the Eq (6) and the  $x(1)$  is measured as the output. The model output is generated with the same initial conditions for  $t=0$  to  $t=10$  seconds. In this work, we analyzed the system for  $\delta = 10^{-5}$ , an infinitesimal distance from its optimal parameter. The visual plots for the model assessment are given below.

From Fig. 3 it is clear that the model is extremely sensitive in the vicinity of the operating initial condition. The system clearly becomes unstable. In addition to that, it is observed that the system is sensitive is almost identical in all the directions. The Model sensitivity index shows that the model is locally unidentifiable, ie there exist several parameter sets in the vicinity that results in an unstable model. In summary, the model identified by SINDy is extremely sensitive to parameter perturbations and a very small perturbation leads to instability. In the next section we demonstrate the role of this insensitivity

in the controller design.

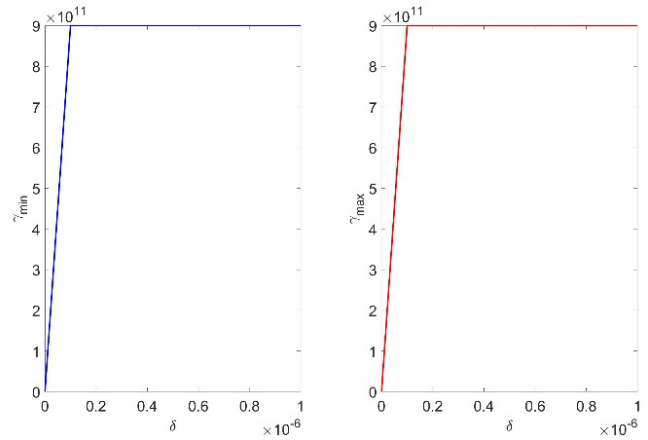


Figure 3: Minimum and Maximum sensitivity plot

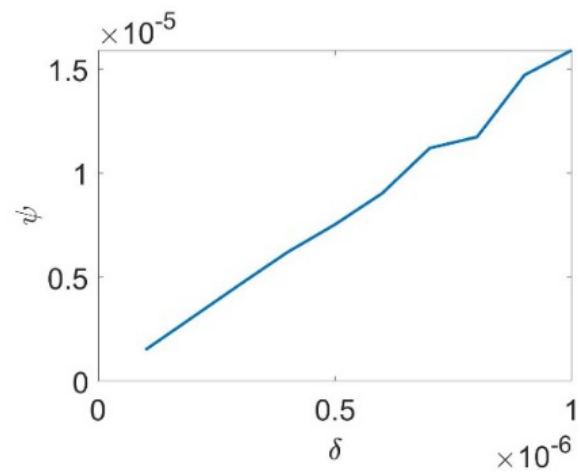


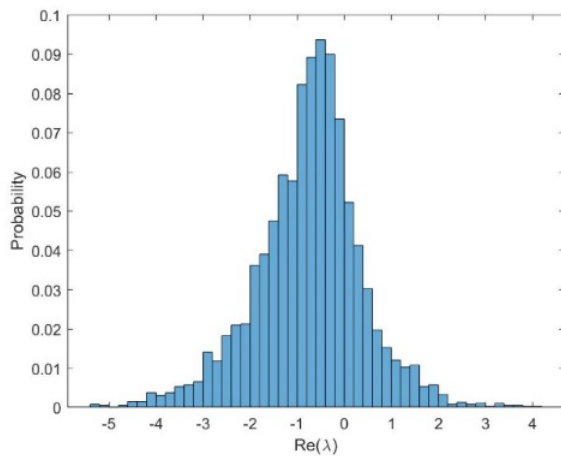
Figure 4: Model sensitivity index

### Analysis of controller design

In the previous section, using conditional sloppiness analysis, we showed that the model is very sensitive in the vicinity of the operating region, and an infinitesimal perturbation from the equilibrium position has destabilized the system. In this section, we show that designing the controller for such a system leads to a fragile controller, in the sense a very small perturbation in the controller parameters will destabilize the system. To demonstrate this, we adopt a pole-placement controller design. We place the poles on the following location to stabilize the oscillations in the system  $p = [-0.5 \ -1 \ -1 \ -0.5]$ .

The gain matrix is obtained for the given A, B and p. The gain matrix is given below. We add a Gaussian random matrix with the controller gain matrix to analyse the controller fragility. We generate a hundred such matrices and compute eigenvalues of the closed-loop system matrix.  $A_k = A - BK$  and plot the histogram of the real part of the eigenvalues in Figure 5. It is evident that the controller design is not robust with respect to very small uncertainty in the controller parameters as the

distributions shows that are positive real eigen values with significantly high probability. This analysis shows that the presence of sloppy and stiff directions in the parameter space of the plant affects controller robustness.



**Figure 5:** Histogram of real part of the closed loop system matrix.

## CONCLUSION

In this work, we proposed a novel method to assess the goodness of data driven surrogate dynamical models developed using the SINDy algorithm for control system design. We analyzed the model generated from the acetaminophen production plant. The original model is a nonlinear ODE. The model is then linearized for the stabilized operating conditions. The analysis revealed that the linearized model is on the verge of instability, and the sloppiness analysis revealed that a tiny perturbation leads to instability. To assess the role of sensitivity in the controller design, we designed a pole placement controller to stabilize the oscillations. We added a small amount of noise to the controller parameters. The Monte Carlo simulations revealed that the closed-loop system would become unstable with a significant probability of minimal uncertainty in the controller parameters. This re-emphasizes that the model identified using the SINDy algorithm is not good enough for control system design. As a logical extension to this work, we propose to formulate and solve a multi-objective robust-optimal control problem to ensure stability, achieve zero waste and maximize the productivity of acetaminophen. This study opens up new avenues in the controller design for sloppy systems.

## ACKNOWLEDGEMENTS

The Authors acknowledge Abhimanyu Raj Shekhar for proving the SINDy model for the acetaminophen plant.

We are also grateful for the support from the U.S. National Science Foundation CBET FMRG ECO-2229250

## REFERENCES

1. Jagadeesan P, Raman K, Tangirala A K (2023) Sloppiness: Fundamental study, new formalism and its application in model assessment. *PLOS ONE* 18(3).
2. Brunton, S. L., Proctor, J. L., & Kutz, J. N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15), 3932–3937. doi:10.1073/pnas.1517384113
3. Shekhar, A. R., Moar, R. R., & Singh, S. (2023). A hybrid mechanistic machine learning approach to model industrial network dynamics for sustainable design of emerging carbon capture and utilization technologies. *Sustainable Energy Fuels*, 7, 5129–5146.
4. Gutenkunst, R. N., Waterfall, J. J., Casey, F. P., Brown, K. S., Myers, C. R., & Sethna, J. P. (2007). Universally Sloppy Parameter Sensitivities in Systems Biology Models. *PLOS Computational Biology*, 3(10), 1–8.
5. L. H. Keel and S. P. Bhattacharyya, "Robust, fragile, or optimal?," in *IEEE Transactions on Automatic Control*, vol. 42, no. 8, pp. 1098-1105, Aug. 1997.
6. Farlessyost, W., Singh, S. Reduced order dynamical models for complex dynamics in manufacturing and natural systems using machine learning. *Nonlinear Dyn* 110, 1613–1631 (2022).
7. Subramanian, R., Moar, R. R., & Singh, S. (2021). White-box Machine learning approaches to identify governing equations for overall dynamics of manufacturing systems: A case study on distillation column. *Machine Learning with Applications*, 3, 100014.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Resource Integration Across Processing Clusters: Designing a Cluster of Clusters

Mohammad Lameh, Dhabia Al-Mohannadi, and Patrick Linke\*

Texas A&M University at Qatar, Department of Chemical Engineering, Education City, Doha, Qatar

\* Corresponding Author: [patrick.linke@gatar.tamu.edu](mailto:patrick.linke@gatar.tamu.edu).

## ABSTRACT

Achieving worldwide sustainable development is a practical challenge that demands an efficient management of resources across their entire value chains. This practical task requires the optimal selection of pathways for extracting, processing, and transporting resources to meet the demands in different geographic regions at minimal economic cost and environmental impact. This work addresses the challenge by proposing a systematic framework for designing resource-processing networks that can be applied to resource management problems. The framework considers the integration and resource exchange within and across multiple processing clusters. It allows for the life cycle assessment of the environmental and economic impacts of the defined value chains, and design accordingly the different processing and transport systems from extraction to final use. The proposed representation and optimization model are demonstrated in a case study to assess the impact of energy transition under decarbonization constraints on long-distance energy supply chains. The objective is to identify optimal cluster designs and interconnecting transportation networks for decarbonized energy supply between energy exporters and importers.

**Keywords:** Optimization, Supply Chain, Energy, Carbon Dioxide, Life Cycle Analysis

## INTRODUCTION

Population growth is associated with increased demand for resources (water, energy, food, products...), fuel combustion, emissions levels, and waste production. Addressing these challenges requires effective resource management to ensure economic prosperity and avoid environmental disasters. Achieving the goals of sustainable development globally requires a life cycle view of the resource management problem.

Process systems engineering has provided different methods and approaches to manage material and energy resources through optimal integration. Heat [1] and mass [2] integration have been implemented at the chemical process scale to minimize feedstock intake and waste generation while meeting a defined production level. The scope of integration has been extended through total site analysis [3] to include exchanging resources between different plants within close vicinity.

Process Integration has been the basis for designing integrated systems, where different plants can exchange specified resources to achieve defined objectives

(economic or environmental). Many problems in this field are developed as mathematical programming models whose solutions yield the optimal outcomes for the defined problems. For example, Almansoori and Shah [4] proposed a mixed integer linear programming (MILP) model for optimizing the design and operation of H<sub>2</sub> supply chains. Alnouri et al. [5] proposed an optimization model for water integration and optimal brine management. Al-Mohannadi et al. [6] proposed a mixed integer linear programming model for integrating CO<sub>2</sub> across sources and sinks to optimize the design of CO<sub>2</sub> capture, utilization, and sequestration (CCUS) systems. These are some applications among many of the Process Integration approaches that have been widely described and reviewed [7]. Many of these approaches focus on a specific number of resources, which limits the scope of such approaches. Expanding the scope would then require developing new models that account for the missing opportunities.

To address this gap, new methods have been proposed that consider more holistic sets of resources by incorporating more materials. For example, the C-H-O



approach balances carbon, hydrogen, and oxygen elements in integrated processing clusters [8]. Different energy streams have been considered in multi-sectoral energy systems design through simulations and linear programming [9]. Although such approaches provide more generic frameworks than the direct applications described earlier, they still lack the comprehensiveness of integrated designs needed when considering different raw materials, intermediates, waste, and emissions streams. Ahmed et al. [10] proposed a resource integration representation of a processing cluster and a MILP model for maximizing the profit of the integrated system while abiding by environmental constraints. The flexible representation allows a complete accounting of all material and energy resources. The framework considers the specifications and conditions at which resources can be exchanged, with an inclusive accounting for the economics of processes and flows. However, this resource integration representation allows resource extraction, production, use, and reuse only within a defined system, which is limiting when it comes to life-cycle assessment considerations or representing the interactions across geographic locations.

Some of the energy systems modeling tools have considered the spatial aspect of the integration problem [11]. In such approaches, the transportation network is an infrastructure dedicated to one of the energy vectors (like a pipeline that transports  $H_2$  from location A to location B). Besides the fact that such models are limited to the energy systems they present (by the resource and process selection), the representation of the transportation network does not allow the consideration of multiple options that may impact the processing in different locations. This work proposes a framework for simultaneous resource integration within and across multiple processing clusters. The novel approach builds on Ahmed et al.'s [4] representation and expands the scope to include multiple systems with supply chains for exchanging resources. The transportation networks and processing clusters have similar generic representation which allows the consideration of multiple technologies for processing and transportation of the resources. A MILP model is proposed to minimize the integrated system's costs. This is the first generic framework for resource integration that yields optimal value chains under environmental limitations while considering transportation networks as processing clusters. An illustration of the model is presented by considering the problem of decarbonizing the energy supply chain between an energy importer and an energy exporter, including long-distance shipping of energy carriers and  $CO_2$ . The case study considers the emissions across the energy value chain in the different systems, as well as different decarbonization pathways, to propose cost-optimal strategies for achieving decarbonization targets at the global level.

## PROBLEM STATEMENT

Given is a system characterized by multiple processing clusters and different resources that can be processed and exchanged within and across the clusters. Each cluster is characterized by a set of processing options that can consume, produce, and convert the resources accessible by the cluster. The aim is to find the optimal capacities of the processing options that result in minimizing the net cost of the overall system while meeting the defined demands of resources in each cluster and abiding by defined footprint limitations.

The proposed representation allows for considering generic systems representing resources' value chains. Each processing cluster can be attributed to a geographic location or a transportation network. This allows the consideration of resource or process availability that may vary across geographic regions. Supply chains or transportation networks can access resources in different geographic locations, which allows for considering different modes of transport, each of which can be characterized as a processing option within the supply chain cluster. Integrating supply chains into the design of processing networks allows accounting for all possible opportunities to determine optimal configurations that meet defined demands at constrained footprint impacts. For example, when designing energy vectors for long-distance shipping, while some pathways may be considered better for shipping (due to high energy density, for example), they might produce more emissions or require higher costs for processing with respect to other less energy-dense carriers. Hence, considering the different processing and transportation options of the various resources simultaneously allows the identification of the synergies and trade-offs to support efficient decisions.

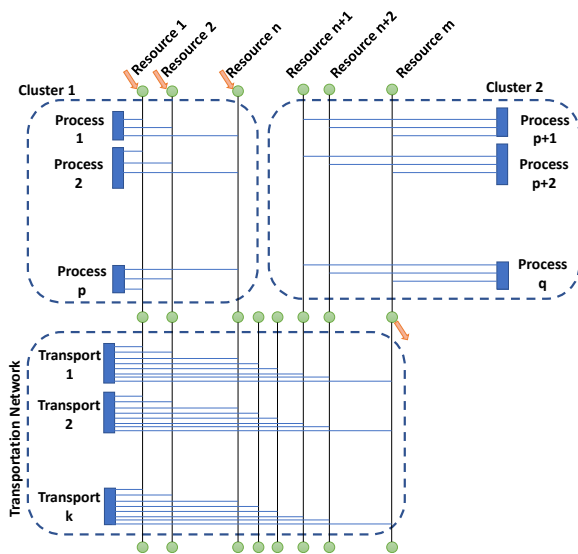
This work proposes a generic approach for considering integrated value chains, giving the user flexibility in defining the specifics of the problem. The aim is to determine the optimal processing options, their location, and their capacity by designing low-cost, low-footprint value chains.

## CLUSTER-OF-CLUSTERS REPRESENTATION

Figure 1 illustrates an example of a superstructure considered in the proposed framework. Each resource that can be consumed, produced, or exchanged within each system is characterized and tracked along the corresponding resource line. Each resource line is characterized by its physical properties (temperature, pressure, composition, AC/DC power, frequency, etc.) and geographic location. The characterization builds upon common infrastructures for energy and material transmission (like power, feedstock, waste, etc.).

Different resources can be imported to each system as fresh feedstock or energy/power inputs, or exported as products, by-products, waste streams, etc. The processes shown in the representation act as conversion units with connections to each resource line. These conversion processes account for any processing a resource line must go through to change its location or specifications. The units are characterized by their costs and capacity limitations, allowing full economic accounting across the entire systems.

The proposed methodology provides a comprehensive scope for considering the integration and resource exchange within and across multiple processing clusters. Each cluster is defined based on accessibility to the available processes and resources. This incorporates the spatial dimensions into the planning framework. The proposed representation considers the supply chains as systems with different transport modes that can exchange resources between different geographic locations. By considering the exchange of resources between multiple clusters, the entire value chain can be assessed, and the impact of transportation modes on resource production and utilization in different systems can be evaluated.



**Figure 1.** Resource integration across multiple clusters.

## OPTIMIZATION MODEL

Given are a set of resources  $R = \{r_1, r_2, \dots, r_n\}$ , a set of processes  $P = \{p_1, p_2, \dots, p_n\}$ , a set of components  $C = \{c_1, c_2, \dots, c_n\}$ , and a set of systems (or clusters)  $S = \{s_1, s_2, \dots, s_n\}$ . Note that the transportation units are considered in the proposed representation as “processes” since they “convert” resource lines from one location to another. Cost, capacity limits, and environmental impact characterize each processing and transportation unit. The proposed framework uses MILP to select the

processing and transportation units to minimize the system's cost under defined constraints on resource imports and exports (resource availability, demand, environmental constraints, etc.). The optimization variables are: the capacities of the different processes (and transportation options) in each system, the resources imports and exports flow rates into and out of each system, the flows of resources exchanged between systems, and the binary variables that activate each of the processes in each system.

The parameters used for characterizing the processes and resource lines are as follows: the conversion parameters, which are defined as the ratio of the flow rate of resource  $r$  produced by process  $p$  in system  $s$  relative to  $F_{p,s}$ , the composition of each component in each resource line in each system, the minimum and maximum capacity of each process in each cluster, the minimum and maximum resource import and export flowrate from each system, and the minimum and maximum resource flow rates of resources exchanged across the different systems.

The following parameters are considered in establishing the economic model to formulate the objective function: the capital expenditure of the processing units in each system is represented as the ratio of the annualized capital costs relative to the capacity of the units, the operating costs of the processing units in each system are represented as the ratio of the operating costs relative to the capacity of the units, and the selling and buying prices of each resource at each system. Economies of scale may result in cost advantages at high production capacities, leading to a non-linear relationship between the capital cost and the capacity. However, when the capacities are high enough, the impact of the economies of scale diminishes, which justifies the linear relation assumed in this work.

The model is set by defining the following constraints that ensure the applicability of the solution and its abidance to mass and energy balances and technical and environmental constraints defined by the user:

- Resource balance: resource balances are defined for each resource line to ensure mass and energy conservation across exports, imports, exchanges, and process conversions.
- Capacity limits: the lower and upper limits on the capacity of each considered process are introduced to ensure the applicability of the solutions based on the process's technical limitations.
- Flowrates limits: the limits on the flowrates of resource imports, exports, and exchanges between the different clusters are introduced to allow the user to define resource demands, consider resource availability, and account for any

limitation on the exchanges between systems.

- Components limits: the limits on the components in the resource exports and imports are defined to allow for the consideration of environmental limitations (such as emissions reduction targets, waste discharges, toxicity limits, etc.).

The total net cost of the overall system (the cluster of all clusters) is determined by accounting for the costs of installing and operating the processing (and transportation) units, the costs of acquiring the resources in the different clusters, and the revenues from selling the resources in each system. The objective function is developed by defining the net cost based on all the accounted costs and revenues to minimize it.

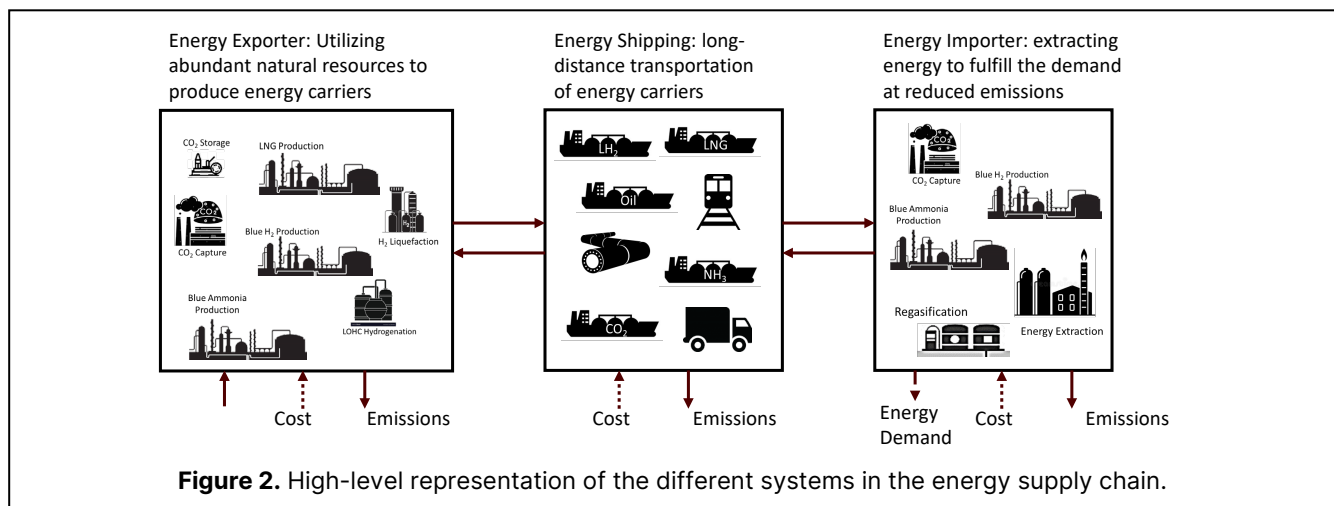
Considering the spatial dimension coupled with the ability to exchange resources across the clusters allows a holistic vision of the interactions between the systems, where the decisions in one system affect the entire value chain. The MILP model can efficiently screen through all possible combinations to give real-time solutions and suggestions on the optimal technology mix of the whole supply chain (from production to transportation, to final use). Note that all the optimization runs performed in the case study (the following section) gave solutions in less than 1 second.

## CASE STUDY: DECARBONIZING LONG-DISTANCE ENERGY SUPPLY CHAIN

The framework applies to analyzing energy supply chains, which can be represented by three systems: an energy importer, an energy exporter, and energy shipping. Figure 2 shows a high-level representation of the considered systems. The energy exporter utilizes abundant natural resources to produce energy carriers that an energy importer can store, transport, and use to supply a defined demand. The case study presented here is based on exporting energy between an energy importer and an

energy exporter 8000 nautical miles apart. The base case corresponds to exporting Liquefied Natural Gas (LNG). This case establishes a performance benchmark for supply chain costs and emissions against which decarbonization is assessed. The emissions from the supply chain can be reduced either through CO<sub>2</sub> capture, shipping, and storage to decarbonize the LNG-based energy supply chain, or by replacing natural gas with alternative hydrogen fuels and carriers. Applying the proposed framework to the described problem allows determining the processes to be implemented in each system (importer, exporter, and shipping) and their capacities, and the flowrates of the different resources imported, exported, and exchanged across the systems. These variables directly relate to key decisions for the optimal supply chain design, which include the following: What are the cost-optimal supply chain decarbonization pathways? Where can decarbonization technologies be applied to reduce targeted supply chain emissions? Which energy carriers should be produced and shipped? Where to produce alternative fuels? How will the cost-optimal decarbonization affect the direct costs and emissions incurred by each system? Considering all the systems constituting the supply chain simultaneously in this design process allows accounting for the impact each of the considered pathways has on the others. For example, transitioning to H<sub>2</sub> based fuels at the importer would decrease the LNG production level at the exporter, given that the energy demand to be supplied at the importer is fixed. This would affect the scale of CO<sub>2</sub> capture and storage from the LNG production process, and the shape of the required shipping fleet (type and number of ships).

The different decisions are considered by defining energy carriers' production processes at the exporter and the corresponding shipping and processing units in the remaining systems. The energy carriers considered are LNG (base case), H<sub>2</sub>, which can either be directly shipped as a liquid or loaded on a liquid organic H<sub>2</sub> carrier (LOHC), and Ammonia. An additional case is investigated



considering LNG as an energy carrier (EC) and decarbonizing the supply chain by capturing the CO<sub>2</sub> at the exporter and the importer. The captured emissions at the importer are shipped back to the exporter to be stored, as CO<sub>2</sub> sequestration may result in safety concerns in some locations [12]. An option for decarbonizing the ships using the shipped alternative fuels (H<sub>2</sub> and NH<sub>3</sub>) as fuel for shipping instead of heavy fuel oil (HFO) is

considered.

Defining the energy supply chain problem in the cluster-of-clusters framework requires introducing the following resources: energy carriers (H<sub>2</sub>, NH<sub>3</sub>, LNG), the processed H<sub>2</sub> conditioned for shipping (liquid or loaded on LOHC), forms of direct energy, CO<sub>2</sub> footprint (from each of the processes), and CO<sub>2</sub> streams in the supply chain (captured, compressed, and liquified CO<sub>2</sub>).

**Table 1:** Key parameters characterizing energy carriers' production and processing units at the exporter.

Process	Cost (\$/GJ)	Natural Gas Intake (GJ NG/GJ)	Power Intake (MWh/GJ)	CO <sub>2</sub> emissions (kgCO <sub>2</sub> /GJ)
LNG production	2.61	1.17	0	17.8
Blue H <sub>2</sub> Production	4.54	1.46	0	10
H <sub>2</sub> Liquefaction	5.82	0	59	0
LOHC Hydrogenation	0.40	0	0.58	0
Blue NH <sub>3</sub> production	13.90	1.52	24	2.2

**Table 2:** Key parameters characterizing energy carriers' processing units at the importer.

Process	Cost (\$/GJ)	EC Usage for heat (%)	Power Intake (kWh/GJ)
LNG Regasification	0.34	1.5%	0.10
H <sub>2</sub> Regasification	0.14	0%	1.2
LOHC Dehydrogenation	1.28	21%	2.3

**Table 3:** Key parameters characterizing energy carriers' ships in the shipping system.

Process	Cost (\$/GJ)	EC loss (%)	CO <sub>2</sub> emissions (kgCO <sub>2</sub> /GJ)
LNG Ships	1.66	6.21%	3.13
H <sub>2</sub> Ships	7.59	7.31%	10.7
LOHC Ships	5.15	0.00%	12.4
NH <sub>3</sub> Ships	3.23	0.15%	7.22
H <sub>2</sub> -fueled LOHC ships	3.44	16.3%	0
NH <sub>3</sub> -fueled NH <sub>3</sub> ships	2.25	9.64%	0

**Table 4:** Key parameters characterizing the different units across the CO<sub>2</sub> supply chain.

System	Process	Cost (\$/tCO <sub>2</sub> )	Natural Gas Intake (GJ/tCO <sub>2</sub> )	Power Intake (kWh/tCO <sub>2</sub> )	CO <sub>2</sub> emissions (tCO <sub>2</sub> /tCO <sub>2</sub> )
Exporter	CO <sub>2</sub> capture and compression	60	3	108	0.11
	CO <sub>2</sub> Processing	3.26	0	4	0
	CO <sub>2</sub> Storage	10	0	0	0
Shipping	CO <sub>2</sub> Ships	70	0	-	0.12
	LNG Fueled CO <sub>2</sub> ships	61	1.05	-	0.09
Importer	CO <sub>2</sub> Capture	48	3	28.7	0.11
	CO <sub>2</sub> Liquefaction	8.5	0	83.1	0

The system contains a three clusters, 24 resources, and 25 processes. The resources and processes are defined within each cluster based on availability, cost, energy requirements, and CO<sub>2</sub> emissions levels. The systems incorporate two major supply chains for shipping energy and CO<sub>2</sub>. The optimization problem aims to find the optimal configuration that supplies the energy demand for the importer at minimum cost while abiding by a defined supply chain emissions level. The energy demand at the importer is set to 140 million GJ/y.

Tables 1 through 4 summarize the data used in setting up the case study. The data on LNG production is based on Katebah et al. [13] and Raj et al. [14]. LNG regasification parameters are estimated based on Hafner and Luciani [15] and Park et al. [16]. The data on blue H<sub>2</sub> production is based on the integrated process reported by Katebah et al. [17]. The parameters characterizing the blue ammonia production process are based on the description of the ammonia process in Ullmann's Encyclopedia [18] and on data reported by Wang et al. [19] and Pfromm [20] on the ammonia process and CO<sub>2</sub> capture [21]. The studies reported by Raab et al. [22] and Niermann et al. [23] were used to estimate the data for H<sub>2</sub> liquefaction and LOHC hydrogenation and dehydrogenation. CO<sub>2</sub> capture data are based on the parameters of CO<sub>2</sub> capture from a natural gas combustion flue gas [24]. The costs of CO<sub>2</sub> sequestration are based on a report by the Global CCS Institute [25], assuming 10 \$/tCO<sub>2</sub> storage and transportation costs. The parameters characterizing CO<sub>2</sub> liquefaction, processing, and CO<sub>2</sub> ships are based on a report published by Element Energy Limited [26] in collaboration with Brevik Engineering, Polarconsult, TNO, and SINTEF. Note that the parameters shown in the tables are used to define the conversions of the different processes based on the intake and production levels of the different resources with respect to their capacities. This allows defining clear boundaries for each process to avoid double counting. For example, the power requirements parameter is defined as zero for some processes. This means that the process does not require importing the resource "power" from the integrated cluster. Power may be produced and consumed within the process boundaries, which may impact the conversion factors of the resources like cost and CO<sub>2</sub> emissions (which are also stated in the tables).

Table 1 shows the key parameters that define the energy carrier's production (from natural gas) and processing units in the cluster representing the energy exporter. The parameters are reported based on LHV energy content in the fuels produced and processed by each unit. Note that the reported costs do not include the cost of natural gas or power, which are considered as resources imported into the cluster as fresh feeds. The power is assumed to be obtained from natural gas-fired power plants with an emissions factor of 0.54

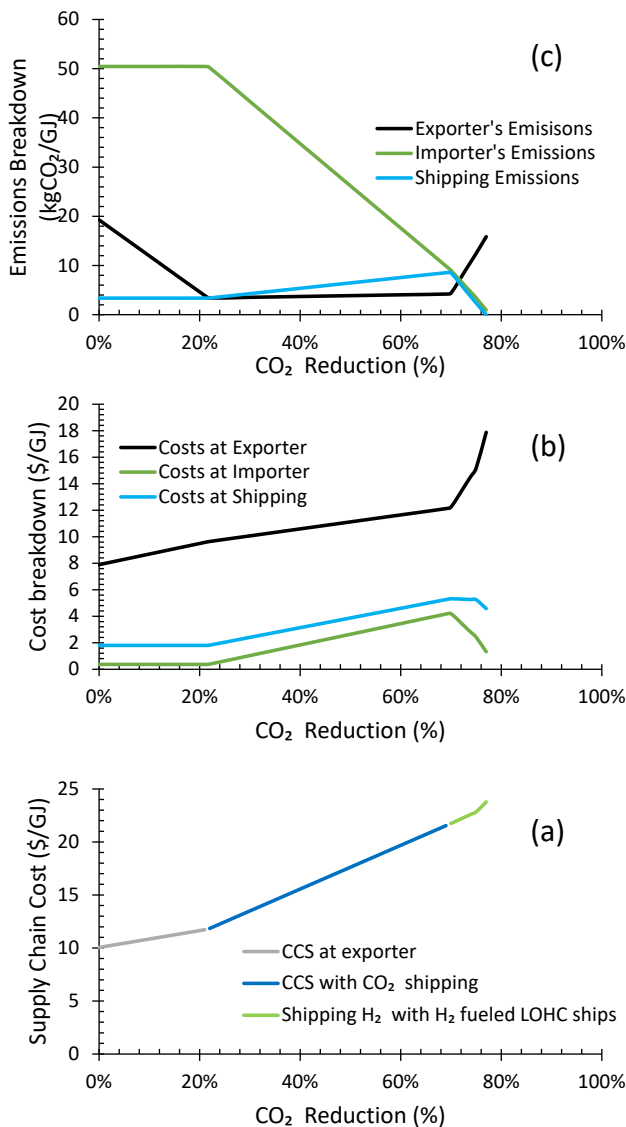
kgCO<sub>2</sub>/kWh. The power price at the exporter is assumed to be 0.032 \$/kWh, and the natural gas cost is 4 \$/GJ. Table 2 shows the parameters used to characterize the processing units at the importer. These units treat the received energy carriers from the ships to be at conditions suitable for energy extraction (through combustion). The heat requirements of these units are covered by the corresponding energy carriers, which is accounted for through the inefficiency or the EC loss parameter. LOHC dehydrogenation is an endothermic process that results in high energy requirements covered by the shipped H<sub>2</sub>. The power is assumed to be covered by a natural gas power plant at a price of 0.19 \$/kWh. Natural gas combustion at the importer (from the case of shipping LNG) produces 50.3 kgCO<sub>2</sub>/GJ. Table 3 presents the parameters characterizing the shipping fleets. The reported costs include heavy fuel oil (HFO), used as the fuel for H<sub>2</sub> ships, LOHC ships, and NH<sub>3</sub> ships. Energy carriers' losses from the ships are due to EC evaporation (LNG and H<sub>2</sub> and NH<sub>3</sub>) or due to utilizing the shipped EC as shipping fuels (H<sub>2</sub>-fueled and NH<sub>3</sub>-fueled ships). The evaporated energy carriers are recovered and used onboard to supply power to the ship's engine (with HFO). CO<sub>2</sub> emissions from the ships are due to the combustion of the shipping fuels (HFO and LNG). The costs of the decarbonized ships are determined by assuming a 20% increase in the ship's capital costs. The cost reduction is due to not utilizing HFO on the decarbonized ships. The trade-off is between avoiding the HFO cost or reducing the energy efficiency of the supply chain by using the energy carriers as fuels. Table 4 shows the parameters that characterize the different units along the CO<sub>2</sub> supply chain. CO<sub>2</sub> is captured from the LNG production, compressed, and stored at the exporter. The CO<sub>2</sub> captured from natural gas combustion at the importer is treated and shipped back to the exporter, where it is processed and stored. The heating requirements for CO<sub>2</sub> capture are covered by burning the natural gas in the supply chain. Two options for fueling the CO<sub>2</sub> ships are considered: HFO or LNG (produced at the exporter) as shipping fuels. Natural gas power is assumed to cover the power requirements of CO<sub>2</sub> processing units.

## RESULTS

The optimization model, based on the defined system, contains 870 constraints and 654 variables. The model is set up in Python and solved using Gurobi. The optimal scenario with the highest emissions flow rate was obtained by running the optimization model at relaxed emissions constraint. This scenario corresponded to producing LNG at the exporter and exporting it on LNG carriers to the importer, where it is regasified, and natural gas is combusted to cover the 140 million GJ/y energy demand. The corresponding emissions level in that



scenario is 10.23 MMtCO<sub>2</sub>/y (26% from the exporter, 69% from the importer, and 5% from the ships).



**Figure 3.** The total supply chain cost (a) and the breakdown of costs (b) and emissions (c) across the three systems at different CO<sub>2</sub> reduction levels.

The emissions limit constraint was then introduced by setting an upper limit on the CO<sub>2</sub> exports across all the systems. The CO<sub>2</sub> emission limit was reduced from 10.23 MMtCO<sub>2</sub>/y to 2.35 MMtCO<sub>2</sub>/y. Further CO<sub>2</sub> reduction was not attainable considering the described technologies. The supply chain cost and the costs and CO<sub>2</sub> emissions of each considered system are tracked. Figure 3 shows the results obtained at varying CO<sub>2</sub> emissions reduction. Three different system configurations were identified as the CO<sub>2</sub> emissions reduction target increased. Low emissions reduction levels (up to 21%) could be achieved by capturing the emissions from LNG production and storing

them at the exporter. This results in maintaining the cost and emissions level in the shipping system and at the importer, as no changes in the designs are required. The emissions at the exporter drop while the cost increases due to the investment in CO<sub>2</sub> capture and storage (CCS).

Emissions reduction between 21% and 69% can be achieved by capturing the emissions of the exporter and the importer and storing them at the importer. The captured emissions from the importer are shipped back to the exporter on CO<sub>2</sub> ships. This configuration increased the costs for all three systems due to the investments required in CO<sub>2</sub> capture, shipping, and storage. The emissions at the exporter were maintained at the same level achieved in the previous configuration as all the emissions were captured. The importer's emissions level dropped significantly, while the level of emissions from shipping increased. The rise in the shipping energy level is due to introducing a new shipping fleet that burns fuel onboard. Going beyond 69% emissions reduction requires a transition in the energy system at the importer from LNG to H<sub>2</sub>. The optimization solution showed that blue H<sub>2</sub> would be produced and loaded on LOHC at the exporter and then shipped to the importer on oil tankers fueled by H<sub>2</sub>. This scenario reduces the cost and emissions level in the shipping system and at the importer due to burning a fuel mix with less carbon content in both systems. The cost and emissions level increase at the exporter. This is because of the higher cost of blue H<sub>2</sub> production relative to LNG production. Moreover, a high amount of CO<sub>2</sub> is captured from the H<sub>2</sub> production, resulting in high emissions due to the capture inefficiency. The results indicate a variation in the performance of the systems as the emissions reduction target increases. An option that achieves global emissions reduction may increase emissions at a singular system level. At the same time, an option that results in global cost reduction may require an unevenly distributed investments. This indicates the necessity for fair crediting of CO<sub>2</sub> reduction to promote efficient implementation globally.

## CONCLUSIONS

This work presented a methodology for cost-optimal integration of different resources within and across multiple systems. The representation allows tracking energy and material streams based on their defined specifications. The generic representation of resources and processes gives the user flexibility in defining the problem by considering multiple material and energy resources that can be integrated across different processing units. The model allows full economic accounting by considering the cost of processes and transportation pathways, the cost of imported resources, and revenues from selling the exports. A comprehensive environmental accounting is allowed by considering multiple footprints.

Considering the design of multiple processing systems simultaneously results in insightful decision-making that accounts for the impact of systems interactions. The method can be applied to a wide range of problems like circular value chains, industrial symbiosis beyond eco-industrial parks, and the net-zero transition. This study presented an application of the framework for decarbonizing the energy supply chain by defining three systems: an energy carrier, an energy importer, and an energy exporter. Such a problem has been addressed in literature by analyzing one or two supply chains, but no work has considered multiple options simultaneously in an optimization framework. In this work, different technologies are considered for decarbonizing the different systems, and the interconnectivity of the decision-making was accounted for in the proposed method. The results showed that cost-optimal decarbonization starts with CO<sub>2</sub> capture and storage at the exporter (reducing 21% of the emissions), followed by CO<sub>2</sub> capture at the importer (with CO<sub>2</sub> shipping, which reduces up to 69% of the emissions), and energy transition to H<sub>2</sub> is activated at high CO<sub>2</sub> emissions reduction targets, which allows for 77% emissions reduction. The CO<sub>2</sub> marginal abatement cost increases from 110 \$/tCO<sub>2</sub> to 250 \$/tCO<sub>2</sub> as the reduction targets rise. The analysis showed that the environmental and economic impact of the decarbonization transition on the different systems varies with the targeted reduction levels. Hence, the implementation of an optimal transition calls for fair crediting of emissions reduction (environmentally and economically) across the systems.

## REFERENCES

- Linnhoff B, Mason DR, and Wardle I. Understanding heat exchanger networks. *Computers & Chemical Engineering* 3(1-4): p. 295-302 (1979)
- El-Halwagi MM and Manousiouthakis V. Synthesis of mass exchange networks. *AIChE Journal* 35(8): p. 1233-1244 (1989)
- Smith R and Delaby O. Targeting flue gas emissions. *Chemical Engineering Research and Design*; (United Kingdom) 69(A6) (1991)
- Almansoori A and Shah N. Design and operation of a future hydrogen supply chain: Snapshot model. *Chemical Engineering Research and Design* 84(6): p. 423-438 (2006)
- Alnouri SY, Linke P, and El-Halwagi MM. Accounting for central and distributed zero liquid discharge options in interplant water network design. *Journal of Cleaner Production* 171: p. 644-661 (2018)
- Al-Mohannadi DM, Kwak G, and Linke P. Identification of optimal transitions towards climate footprint reduction targets using a linear multi-period carbon integration approach. *Computers & Chemical Engineering* 140: p. 106907 (2020)
- Klemeš JJ, *1 - process integration: An introduction*, in *Handbook of process integration (pi) (second edition)*, J.J. Klemeš, Editor. 2023, Woodhead Publishing. p. 1-24.
- Noureldin MM and El-Halwagi MM. Synthesis of c-h-o symbiosis networks. *AIChE Journal* 61(4): p. 1242-1262 (2015)
- Limpens G, Moret S, Jeanmart H, and Maréchal F. Energyscope td: A novel open-source model for regional energy systems. *Applied Energy* 255: p. 113729 (2019)
- Ahmed R, Shehab S, Al-Mohannadi DM, and Linke P. Synthesis of integrated processing clusters. *Chemical Engineering Science* 227: p. 115922 (2020)
- Kakodkar R, He G, Demirhan CD, Arbabzadeh M, Baratsas SG, Avraamidou S, Mallapragada D, Miller I, Allen RC, Gençer E, and Pistikopoulos EN. A review of analytical and optimization methodologies for transitions in multi-scale energy systems. *Renewable and Sustainable Energy Reviews* 160: p. 112277 (2022)
- Al Ghafri SZS, Revell C, Di Lorenzo M, Xiao G, Buckley CE, May EF, and Johns M. Techno-economic and environmental assessment of lng export for hydrogen production. *International Journal of Hydrogen Energy* 48(23): p. 8343-8369 (2023)
- Katebah MA, Hussein MM, Shazed A, Bouabidi Z, and Al-musleh EI. Rigorous simulation, energy, and environmental analysis of an actual baseload LNG supply chain. *Computers & Chemical Engineering* 141: p. 106993 (2020)
- Raj R, Suman R, Ghandehariun S, Kumar A, and Tiwari MK. A techno-economic assessment of the liquefied natural gas (lng) production facilities in western canada. *Sustainable Energy Technologies and Assessments* 18: p. 140-152 (2016)
- Hafner M and Luciani G, *The palgrave handbook of international energy economics*. 2022: Springer Nature.
- Park J, Lee I, You F, and Moon I. Economic process selection of liquefied natural gas regasification: Power generation and energy storage applications. *Industrial & Engineering Chemistry Research* 58(12): p. 4946-4956 (2019)
- Katebah M, Al-Rawashdeh Mm, and Linke P. Analysis of hydrogen production costs in steam-methane reforming considering integration with electrolysis and co2 capture. *Cleaner Engineering and Technology* 10: p. 100552 (2022)
- Appl M, *Ammonia, 3. Production plants*, in *Ullmann's encyclopedia of industrial chemistry*.
- Wang M, Khan MA, Mohsin I, Wicks J, Ip AH, Sumon

KZ, Dinh C-T, Sargent EH, Gates ID, and Kibria MG. Can sustainable ammonia synthesis pathways compete with fossil-fuel based haber–bosch processes? *Energy & Environmental Science* 14(5): p. 2535-2548 (2021)

20. Pfromm PH. Towards sustainable agriculture: Fossil-free ammonia. *Journal of Renewable and Sustainable Energy* 9(3) (2017)
21. IEAGHG. Techno-economic evaluation of hyco plant integrated to ammonia/urea or methanol production with ccs. IEA Greenhouse Gas R&D Programme: Cheltenham, UK (2017)
22. Raab M, Maier S, and Dietrich R-U. Comparative techno-economic assessment of a large-scale hydrogen transport via liquid transport media. *International Journal of Hydrogen Energy* 46(21): p. 11956-11968 (2021)
23. Niermann M, Drünert S, Kaltschmitt M, and Bonhoff K. Liquid organic hydrogen carriers (lohcs)–techno-economic analysis of lohcs in a defined process chain. *Energy & Environmental Science* 12(1): p. 290-307 (2019)
24. Rubin ES, Davison JE, and Herzog HJ. The cost of co2 capture and storage. *International Journal of Greenhouse Gas Control* 40: p. 378-400 (2015)
25. ZEP. The costs of co<sub>2</sub> capture, transport and storage, in *Post demonstration CCS in the EU*. European Technology Platform for Zero Emissions Fossil Fuel Power Plants (2011)
26. Durusut E and Joos M. Shipping co<sub>2</sub> - uk cost estimation study. Element Energy Limited: Cambridge, UK (2018)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Evaluating Circularity and Sustainability in Plastic Waste Recycling: Open and Closed-Loop Technologies

Wafaa N. Majzoub<sup>a</sup> and Dhabia M. Al-Mohannadi<sup>a\*</sup>

<sup>a</sup> Texas A&M University at Qatar, Department of Chemical Engineering, Doha, Qatar

\* Corresponding Author: [dhabia.al-mohannadi@qatar.tamu.edu](mailto:dhabia.al-mohannadi@qatar.tamu.edu).

---

## ABSTRACT

In a world grappling with mounting plastic waste, the pursuit of sustainable plastic waste management has become pivotal in aligning with Circular Economy (CE) goals, with a strong emphasis on resource conservation, product durability, and carbon footprint reduction. The strategic implementation of recycling methods serves as a stepping stone for transitioning from linear to circular models. This work delves into plastic waste recycling technologies, specifically focusing on open and closed-loop approaches, providing a comprehensive evaluation anchored on economic, environmental, and circularity criteria. Different recycling techniques are thoroughly examined, with particular attention given to chemical recycling methods such as pyrolysis and gasification. This work introduces a comprehensive screening model driven by a new proposed circularity metric validated through a case study to assess these recycling pathways. The results reveal the substantial potential of chemical recycling technologies compared to conventional incineration for energy recovery. Pyrolysis refinery and methanol production from plastic waste demonstrate triple and double the profitability of incineration while significantly enhancing the overall contribution of CE. This work emphasizes the imperative of a sustainable approach to plastic waste management by balancing different metrics considerations.

---

**Keywords:** Plastic waste, Chemical recycling, Circular economy, Circularity

## INTRODUCTION

Plastic materials have become increasingly prevalent as lightweight, robust, and cost-effective alternatives. The annual production of plastics is estimated to reach 430 million metric tons by 2023, of which 139 million metric tons will become waste [5]. The accumulation of plastic waste pollutants poses a pressing global challenge due to its degrading effects on ecosystems and threats to living and nonliving systems [6]. Therefore, sustainable plastic waste management methods are urgently needed to fulfill the requirements of the Circular Economy (CE) [7]. Recycling strategies are imperative for the shift from linear to circular models [8]. Recycling techniques are divided into material and chemical recycling. Material recycling could involve mechanical processes to produce regranulates and degraded plastic quality [9] or physical recycling that relies on solvent-based separation methods [5]. On the other hand, chemical recycling, or molecular recycling, involves breaking

down polymer chains into oligomers, monomers, or other basic chemicals through high-temperature processes. It includes pyrolysis, gasification, and depolymerization. Pyrolysis is a method that thermally degrades long-chain polymer molecules into smaller, less complex molecules, producing valuable products like oil, gas, and char [10]. Gasification is a thermo-chemical process transforming carbon-based materials through partial oxidation into synthetic gas [11]. Depolymerization, or Chemolysis, uses solvents to break polymer chains, making it ideal for homogeneous plastics.

Recycling methods are categorized into polymer, monomer, and molecular loops. Material recycling belongs to the polymer loop, producing purified plastic waste identical to the input. Chemical recycling includes depolymerization of the monomer loop, whereas pyrolysis and gasification are classified as molecular loops. Closed-loop recycling, which allows for the potential production of virgin-quality plastic from recycled materials, is commonly linked to monomer and molecular loops [5].

The potential of chemical recycling technologies comes with the challenge of identifying the most sustainable options, given the numerous possibilities and the early stages of research and development. This approach, which transforms waste into a resource, demands an efficient screening method to explore all possibilities [12]. Moreover, exploring plastic waste management pathways is a pivotal driver toward a CE model, emphasizing sustainability through waste minimization and resource optimization with the potential to mitigate pollution resulting from landfilling and direct combustion [13].

## BACKGROUND

Various state-of-the-art and critical literature studies have contributed significantly to our understanding of sustainable waste management practices. [14] proposed a method integrating mass flow analysis, techno-economic assessment, and life cycle assessment to evaluate primary plastics production and sorting. [15] compared thermochemical depolymerization technologies for LDPE waste, while [16] analyzed closed-loop recycling technologies for common consumer polymers. [17] conducted simulations of recycling processes for PP, and [18] compared thermochemical technologies for MSW treatment. Additionally, [19] and [20] assessed energy recovery options from MSW in Brazilian cities and southern Spain, respectively, providing valuable insights for sustainable waste management practices.

Nevertheless, limited research has focused on building optimization models to assess waste management pathways and enhance CE objectives. For example, [12] proposed a framework utilizing mixed integer linear programming to optimize waste recovery processes, aligning with circular economy principles. [13] introduced a mixed-integer nonlinear programming model for sorting and recycling mixed plastic waste, emphasizing economic feasibility and environmental impact. [6] presented a comprehensive superstructure and a multi-objective mixed integer nonlinear fractional programming model for HDPE recycling. [21] expanded on this work with a multi-objective MINLP model to optimize plastic waste sorting and recycling processes. [22] developed a systematic framework that integrates life cycle assessment and optimization methods for analyzing circular systems and synthesizing processes. Most of these studies have focused on evaluating the viability of diverse waste management technologies for transforming plastic waste into value-added products. However, comparatively less attention has been given to utilizing the primary products derived from chemical recycling processes, such as pyrolysis oil and synthesis gas. Despite their potential as excellent raw materials, they can enable both closed-loop (plastic to plastic) recycling, such as olefin production, as well as open-loop (plastic to product) recycling

options, such as hydrogen, ammonia, and Fisher-Tropsch synthesis. Although most literature studies focus on economic and environmental aspects, they overlook crucial circular economy factors such as material and energy utilization, product recyclability, quality, and durability. A comprehensive evaluation of these aspects is essential for adhering to circular economy principles in open and closed-loop recycling approaches. It is crucial to explore these pathways' potential thoroughly to advance these principles and establish a more sustainable approach to plastic waste management. Furthermore, addressing a common misconception suggests every product and material should undergo maximum recycling is essential. However, it is crucial to recognize that specific disposal methods may be more environmentally friendly for certain products. This distinction is vital because recycling processes for some products can be energy and carbon-intensive [23, 24]. Therefore, a well-balanced and comprehensive evaluation is essential to overcome this trade-off.

Expanding research on plastic waste treatment, mainly through chemical recycling, is crucial to address existing limitations. This work introduces significant contributions to the field by presenting a screening model and decision-making framework for evaluating plastic waste recycling technologies, with a particular focus on chemical recycling and various upgrading and processing technologies within both closed-loop and open-loop pathways to achieve holistic, sustainable waste management strategies, facilitating sophisticated decision-making by balancing economic, environmental, and circularity aspects. Moreover, the model introduces a novel circularity indicator integrating various metrics to assess circularity comprehensively. The introduced measure considers factors such as material consumption, energy demand, and economic viability, thereby enhancing sustainability in plastic waste management. It emphasizes not solely evaluating one factor but focusing on multiple factors simultaneously. This approach ensures a comprehensive assessment of the downstream implications for resulting products, contributing to a more holistic evaluation of plastic waste management strategies.

## METHODS

This work focuses on investigating various plastic waste recycling technologies. Figure 1 provides an overview of potential pathways for plastic waste management and the final products. The approach to screening different plastic waste treatment methods involved conducting a thorough literature review and collecting data from real-life plants and simulations. A structured database covering various technologies was analyzed for techno-economic and environmental performance. A screening model formulated using mixed integer linear



programming aimed to maximize the overall net profit while considering circularity indicators. Validation through a comprehensive case study confirmed the model's effectiveness. The model's primary objective is represented by Eq (1) and subjected to multiple equality and inequality constraints illustrated by Eq (2) and Eq (3). These constraints include material and energy balance, capacity limits, economics considerations, and the degree of circularity metric (DCM) constraints. The net profit of the recycling route is determined by subtracting annual income from total production costs. Capital expenditure (Capex) information is estimated through regression modeling with piecewise linearization to handle nonlinear data. Circularity is measured through multiple factors including Material Indicator (MI), Energy Indicator (EI), Water Indicator (WI), Solid-Waste Indicator (SWI), Global Warming Potential Indicator (GWPI), Economic Indicator (ECI), Co-product Utilization Indicator (CPI), Recyclability Indicator (RI), the product's Quality Indicator (QI), and Technology Readiness Level Indicator (TRLI). The individual indicators are classified into two categories: "Higher the better," where higher values emphasize better circular economy contribution, and the second category, "Lower the better," where lower values are more useful for overall circularity performance. Both categories have been normalized on a scale of 0 to 1 Max-Min normalization method [25], where 1 represents the

best-case scenario, and 0 is the worst-case scenario for both. Simultaneously, higher values of DCM are desirable, ensuring better technology sustainability. All these factors collectively form the total DCM, as shown in Eq (4). Moreover, each indicator in both categories is given a specific weight in the total circularity indicator using the Simple Additive Weighting (SAW) method [25]—more details about how each indicator and the total DCM are illustrated in the supplementary materials. A case study was conducted and evaluated to validate the model's efficacy. The proposed MILP and optimization problem have been solved with Python 3.10.2 - Pyomo 6.4.0 and Gurobi solver 10.0.1. Optimal solutions have been consistently obtained within a few seconds.

$$\text{MAXIMIZE (Net profit)} = f(x) \quad (1)$$

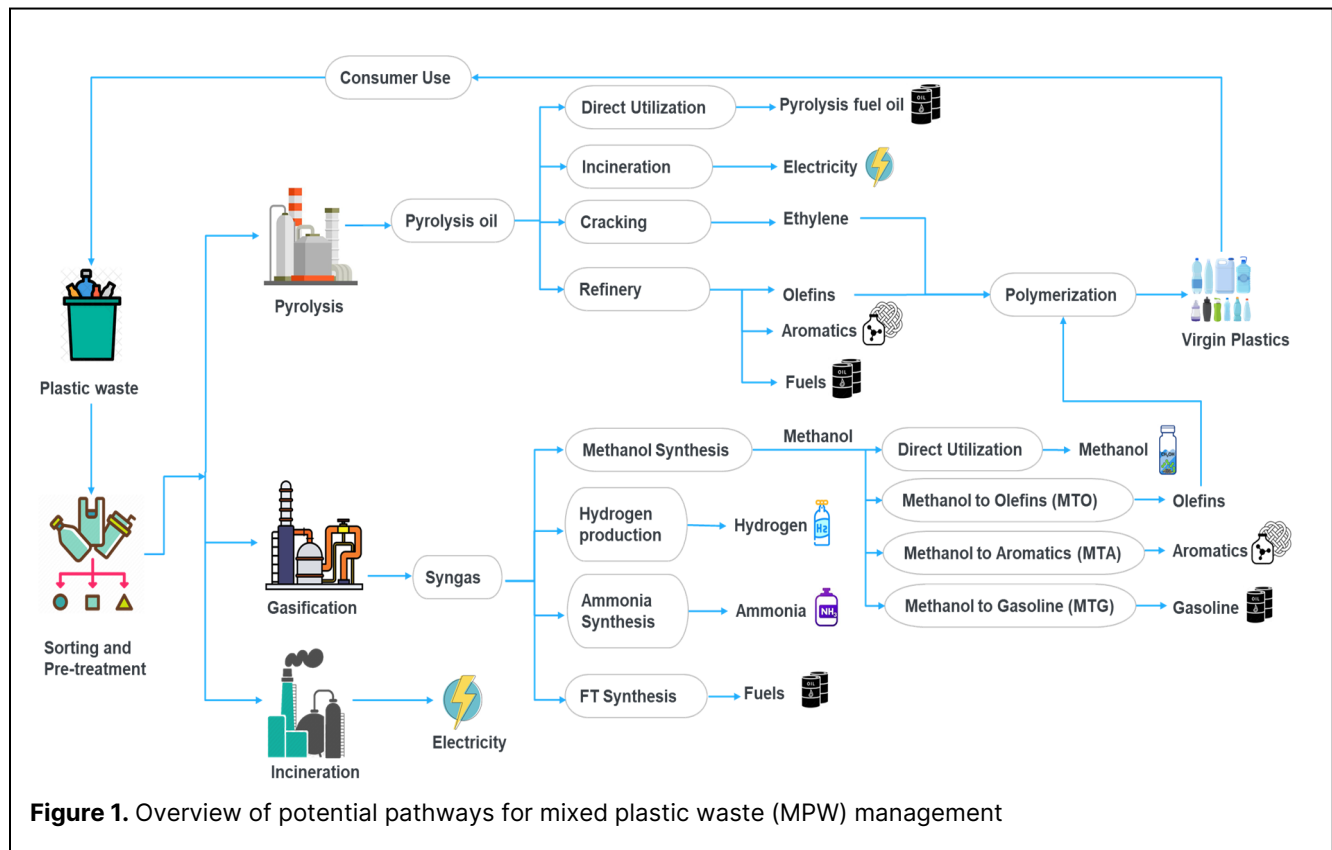
Subject to:

$$g_i(x) \leq 0, \quad i = 1, \dots, m \quad (2)$$

$$h_j(x) = 0, \quad j = 1, \dots, n \quad (3)$$

$$\text{DCM}_k = MI_k + EI_k + WI_k + SWI_k + GWPI_k + ECI_k + CPI_k + RI_k + QI_k + TRLI_k, \quad k = 1, \dots, p \quad (4)$$

Where  $k$  is the set of plastic waste recycling technology alternatives, and  $p$  represents the total count of technology alternatives.



**Figure 1.** Overview of potential pathways for mixed plastic waste (MPW) management

**Table 2.** Required materials and utilities in kg/ tons of plastic waste input and MWh/ton waste for electricity.

Name	NaCl	Sulfolane	H <sub>2</sub>	Bauxite	Limestone	O <sub>2</sub>	He	Water	Solid waste	Energy	Reference
GTM	0	0	0	0	0	790	0	986	6	2.4	[1]
PR	0	0.3	16	0	0	0	0.04	0.2	47	1.6	[2, 3]
IER	6	0	0	6	9	0	0	3830	98	0.3	[4]

## CASE STUDY

The case study considers three pathways: gasification, pyrolysis represented by methanol production, and pyrolysis refinery, both closed-loop pathways capable of producing new plastic, and incineration for energy recovery serving as an open-loop base case scenario. The model aimed at achieving maximum economic profit and maximum contribution to the circular economy through the higher values of DCM. The case study is scaled based on sources that generate 100,000 tons of mixed plastic waste annually [26]. Detailed technical information about each scenario, such as the yield, material consumption requirements, energy, utility requirements, waste generation, and emissions, are provided in Table 1 and Table 2.

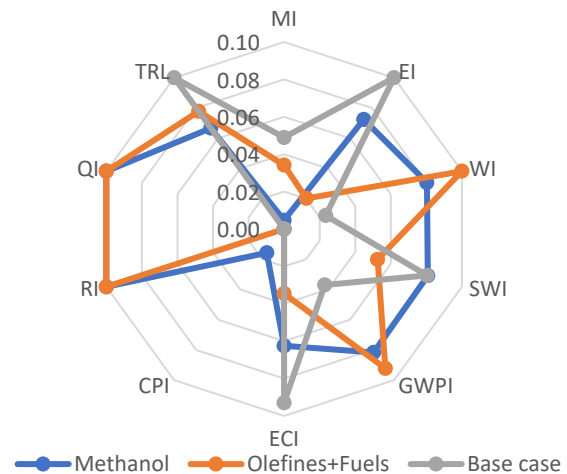
**Table 1.** Pathways considered in the case study.

Name	Main products	Yield %	Ref.
Gasification to Methanol Synthesis (GTM)	Methanol CO <sub>2</sub> (carbon credit)	1.23 0.1	[1]
Pyrolysis+ Refinery (PR)	Ethylene Propylene Aromatics mix. LMWHC (Gasoline) HMWHC (Diesel)	0.20 0.13 0.04 0.54 0.05	[2, 3]
Incineration for energy recovery (IER)	Electricity (Gas, LHV=40.5 kJ/kg)	0.90	[4, 27]

## RESULTS AND DISCUSSION

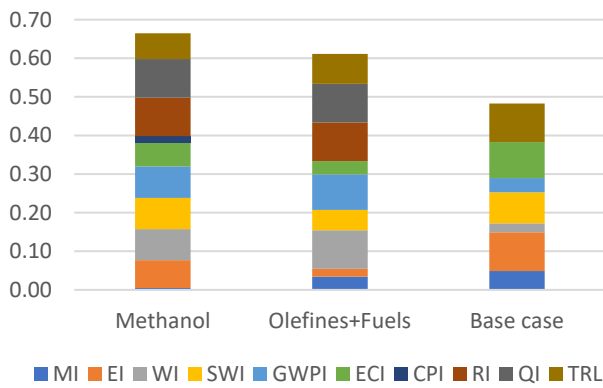
Figure 2 illustrates the individual circularity indicators that contribute to the DCM. It is important to note that an assumption of equal weight distribution has been given to individual metrics impacting the final DCM value, set at 10% each. In the case of methanol production, the MI is notably low at 0.005, indicating inefficient material

utilization and the requirement of extensive materials for recycling the given plastic waste quantity. However, the EI is 0.07, demonstrating good energy efficiency with minimal energy demands. The WI is 0.08, reflecting efficient water use, while the SWI is also 0.08, suggesting minimal solid waste generation. The GWPI stands at 0.08, highlighting efficient waste management practices. The ECI is 0.06, indicating economic feasibility. In contrast, the CPI is very low at 0.02, signifying minimal co-product generation within the process. The RI and QI are at 0.100, emphasizing full recyclability and high product quality. Lastly, TRLI is 0.07, indicating a mature technology. Low MI, EI, ECI, and CPI values signify inefficient resource usage and high annualized costs in the pyrolysis refinery process. Conversely, high WI and GWPI values indicate effective water conservation, reduced emissions, and moderate waste generation. The process excels in recyclability and product quality, underscoring its commitment to sustainability and circular economy goals, as RI and QI values indicate.



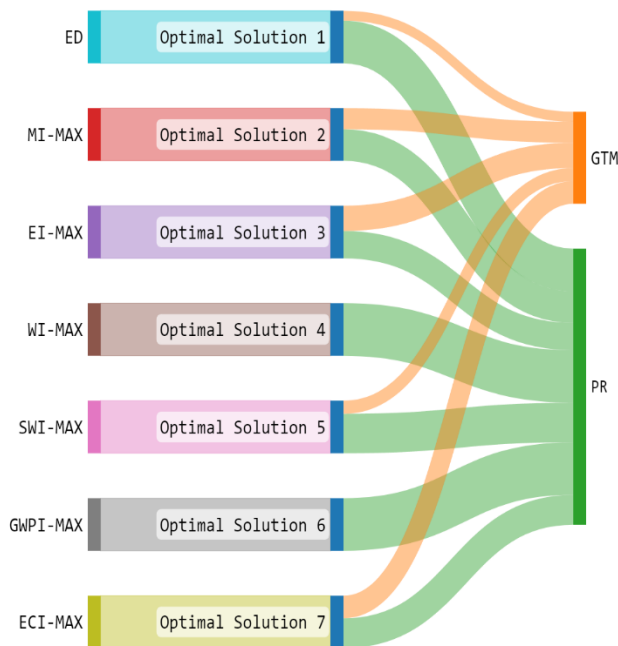
**Figure 2.** Individual circularity indicators for all pathways.

The high TRL reflects a mature technology that efficiently utilizes energy (high EI) and exhibits economic viability (high ECI) for the incineration energy recovery pathway. However, low MI, WI, SWI, GWPI, CPI, RI, and QI values raise environmental concerns.



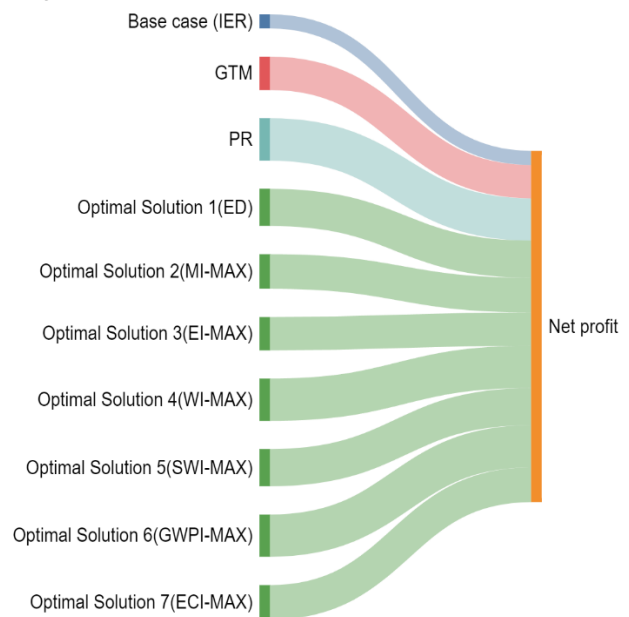
**Figure 3.** DCM Breakdown.

This pathway consumes substantial materials and water, generates significant waste, emits pollutants, produces non-recyclable products, and lacks co-products. The DCM values are represented in Figure 3, where higher values indicate higher circularity levels. The incineration base case achieves the lowest DCM of 0.48, while methanol synthesis contributes to circular economy principles, boasting a DCM of 0.67, nearly 40% better than the base case. Pyrolysis refinery follows closely with a DCM of 0.61, exceeding the base case by 27%. Regarding profitability, all technologies are economically viable and generate positive economic impacts. The base case yields an annual profit of 11 M USD. In contrast, methanol synthesis and pyrolysis refinery outperform, with net earnings of 26 M USD and 33 M USD per year, respectively, doubling and tripling the base case's net profit.



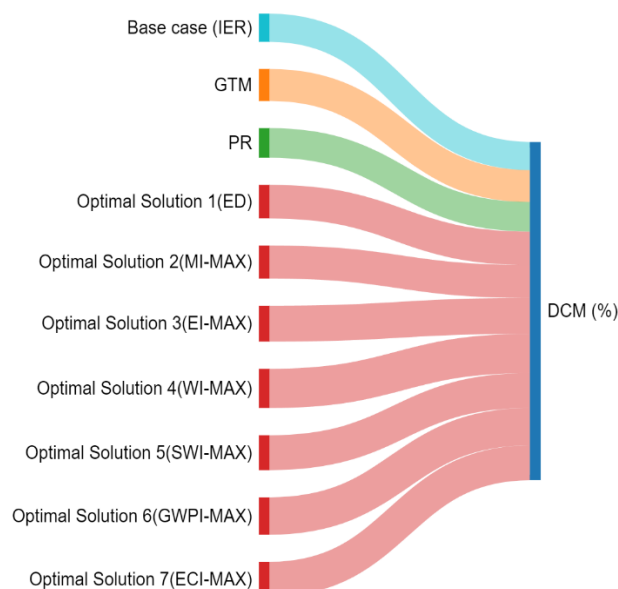
**Figure 4.** Capacity allocation of the optimal solution using different indicator weights.

After conducting a multi-criteria analysis, the optimization model was tested with an objective function aimed at maximizing the net profit of the technology while pursuing the maximum possible circularity. The first case was solved with equal weight distribution (ED) of the individual metrics comprising the DCM. However, the specific preferences for each indicator depended on the context and objectives of circularity within the technology or system. Sometimes, prioritizing one indicator was essential based on geographical location, system needs, and regulatory perspectives. Therefore, the analysis included a second case of assigning a high weight of 55% to the indicator under focus, while other indicators received a minimum equal contribution of 5% each. This part of the analysis explored the optimal solution when one indicator was significant, with the remaining indicators being critical. Both cases were solved using suitable constraints for the same objective function, resulting in the optimal allocation of the proposed capacity depicted in Figure 4.



**Figure 5.** Profitability (M USD/y) of the optimal solutions in comparison to the base case, GTM, and PR processes.

Furthermore, the net profit and DCM values, as illustrated in Figure 5, and Figure 6 respectively, show that the optimal solutions will always have better profitability and circularity contribution compared to the base case. Maximizing the value of the weight of MI, EI, WI, SWI, GWPI, and ECI in the total DCM will not significantly affect the overall economic feasibility of the optimal solutions as reflected by the values of net profit of 27, 26, 33, 29, 33, and 27 M USD/y respectively. Moreover, the DCM faces an apparent variation with values of 68%, 74%, 80%, 72%, 77%, and 72% respectively.



**Figure 6.** Degree of Circularity Metric DCM (%) of the optimal solutions in comparison to the base case, GTM, and PR processes.

## CONCLUSIONS

Given the projected annual increase in plastic waste, sustainable plastic waste management has become necessary. This work delves into open and closed-loop plastic waste recycling technologies. The novel screening model introduced here is a robust tool for evaluating these recycling pathways. The case study centered around 100,000 tons of mixed plastic waste generated annually, considers three pathways: gasification, pyrolysis represented by methanol production, and pyrolysis refinery, both closed-loop pathways capable of producing new plastic, and incineration for energy recovery serving as an open-loop base case scenario. The results reveal chemical recycling technologies' considerable potential. Methanol synthesis and pyrolysis refinery emerge as the stars, generating net profits of 26 M USD and 33 M USD per year, respectively, substantially exceeding the base case's annual profit of 11 M USD. While economic viability is crucial, this analysis emphasizes circularity considerations. The incineration base case records the lowest DCM at 0.48, while methanol synthesis shines as a circular economy champion, boasting a DCM of 0.67, nearly 40% better than the base case. The pyrolysis refinery has a DCM of 0.61, surpassing the base case by 27%. Moreover, the analysis also investigated the impact of different indicator weights on the optimal solutions, allowing for tailored solutions based on specific system needs and preferences. The optimal solution varied accordingly, highlighting the importance of selecting relevant indicators for the given context of the system. In conclusion, the presented approach provides valuable

insights for decision-makers and stakeholders, enabling them to choose appropriate recycling technologies based on their objectives and priorities. The results emphasize the crucial role of circularity in improving both environmental sustainability and economic viability. Future research can build upon these findings, exploring additional factors and technology options to enhance plastic waste management further and contribute to advancing the circular economy concept.

## ACKNOWLEDGEMENTS

We want to acknowledge Dr. Ma'moun Al-Rawashdeh, Dr. Ahmed Abdala, and Dr. Mohamed Nounou, Professors of Chemical Engineering at Texas A&M University at Qatar, for their invaluable guidance and support.

## REFERENCES

1. Prifti K, Galeazzi A, Manenti F. Design and simulation of a plastic waste to methanol process: Yields and economics. *Ind. Eng. Chem. Res.* 62:5083-5096 (2023)
2. Gracida-Alvarez UR, Winjobi O, Sacramento-Rivero JC, Shonnard DR. System analyses of high-value chemicals and fuels from a waste high-density polyethylene refinery. Part 2: Carbon footprint analysis and regional electricity effects. *ACS Sustainable Chem Eng* 7:18267-18278 (2019)
3. Gracida-Alvarez UR, Winjobi O, Sacramento-Rivero JC, Shonnard DR. System analyses of high-value chemicals and fuels from a waste high-density polyethylene refinery. Part 1: Conceptual design and techno-economic assessment. *ACS Sustainable Chem Eng* 7:18254-18266 (2019)
4. Havukainen J, Zhan M, Dong J, Liikanen M, Deviatkin I, Li X, Horttanainen M. Environmental impact assessment of municipal solid waste management incorporating mechanical treatment of waste and incineration in Hangzhou, China. *J Clean Prod* 141:453-461 (2017)
5. Garcia-Gutierrez P, Amadei AM, Klenert D, Nessi S, Tonini D, Tosches D, Ardente F, Saveyn H. Environmental and economic assessment of plastic waste recycling - A comparison of mechanical, physical, chemical recycling, and energy recovery of plastic waste. *OP, Luxembourg.* JRC132067 (2023)
6. Zhao X, You F. Waste high-density polyethylene recycling process systems for mitigating plastic pollution through a sustainable design and synthesis paradigm. *AIChE J* 67:4 (2021)
7. Sakthipriya N. Plastic waste management: A roadmap to achieve circular economy and recent innovations in pyrolysis. *Sci Total Environ.* 809:151160 (2022)

8. Shamsuyeva M, Endres HJ. Plastics in the context of the circular economy and sustainable plastics recycling: Comprehensive review on research development, standardization, and market. *Composites Part C: Open Access* 6:100168 (2021)
9. Schyns ZOG, Shaver MP. Mechanical recycling of packaging plastics: A review. *Macromol Rapid Commun* 42:2000415 (2021)
10. Anuar Sharuddin SD, Abnisa F, Wan Daud WMA, Aroua MK. Energy recovery from pyrolysis of plastic waste: Study on non-recycled plastics (NRP) data as the real measure of plastic waste. *Energy Convers. Manage.* 148:925-934 (2017)
11. Arena U. Process and technological aspects of municipal solid waste gasification: A review. *Waste Manag.* 32:625-639 (2012)
12. Somoza-Tornos A, Pozo C, Graells M, Espuña A, Puigjaner L. Process screening framework for the synthesis of process networks from a circular economy perspective. *Resour. Conserv. Recycl.* 164:105147 (2021)
13. Lim J, Ahn Y, Kim J. Optimal sorting and recycling of plastic waste as a renewable energy resource considering economic feasibility and environmental pollution. *Process Saf Environ Prot.* 169:685-696 (2022)
14. Volk R, Stallkamp C, Steins JJ, Yogish SP, Müller RC, Stapf D, Schultmann F. Techno-economic assessment and comparison of different plastic recycling pathways: A German case study. *J Ind Ecol* 255:1318-1337 (2021)
15. Hernández B, Kots P, Selvam E, Vlachos DG, Ierapetritou MG. Techno-economic and life cycle analyses of thermochemical upcycling technologies of low-density polyethylene waste. *ACS Sustainable Chem Eng* 11:7170-7181(2023)
16. Uekert T, Singh A, DesVeaux JS, Ghosh T, Bhatt A, Yadav G, Afzal S, Walzberg J, Knauer KM, Nicholson SR, et al. Technical, economic, and environmental comparison of closed-loop recycling technologies for common plastics. *ACS Sustainable Chem Eng* 11:965-978 (2023)
17. Bora RR, Wang R, You F. Waste polypropylene plastic recycling toward climate change mitigation and circular economy: Energy, environmental, and techno-economic perspectives. *ACS Sustainable Chem Eng* 8:16350-16363 (2020)
18. Gabbar HA, Aboughaly M. Conceptual process design, energy and economic analysis of solid waste to hydrocarbon fuels via thermochemical processes. *Processes* 9:2149 (2021)
19. Leme MMV, Rocha MH, Lora EES, Venturini OJ, Lopes BM, Ferreira CH. Techno-economic analysis and environmental impact assessment of energy recovery from municipal solid waste (MSW) in Brazil. *Resour. Conserv. Recycl.* 87:8-20 (2014)
20. Fernández-González JM, Grindlay AL, Serrano-Bernardo F, Rodríguez-Rojas MI, Zamorano M. Economic and environmental review of waste-to-energy systems for municipal solid waste management in medium and small municipalities. *Waste Manag.* 67:360-374 (2017)
21. Lee J, Lim J, Joo C, Ahn Y, Cho H, Kim J. Multiobjective optimization of plastic waste sorting and recycling processes considering economic profit and CO<sub>2</sub> emissions using nondominated sorting genetic algorithm II. *ACS Sustainable Chem Eng* 10:13325-13334 (2022)
22. Thakker V, Bakshi BR. Toward sustainable circular economies: A computational framework for assessment and design. *J Clean Prod* 295:126353 (2021)
23. Rocchi L, Paolotti L, Cortina C, Fagioli FF, Boggia A. Measuring circularity: An application of modified Material Circularity Indicator to agricultural systems. *Agric Food Econ* 9:1-13 (2021)
24. Supply Chain S. Measuring progress towards a circular economy: Indicators and metrics tool guidance. *Supply Chain Sustainability School* (2019)
25. Mukhametzhanov IZ. Normalization Of Multidimensional Data for Multi-Criteria Decision-Making Problems: Inversion, Displacement, Asymmetry. Springer (2023)
26. Nikiema J, Asiedu Z. A review of the cost and effectiveness of solutions to address plastic pollution. *Environ. Sci. Pollut. Res.* 1:3-3 (2022)
27. Miandad R, Barakat MA, Aburiazaiza AS, Rehan M, Nizami AS. Catalytic pyrolysis of plastic waste: A review. *Process Saf. Environ. Prot.* 102:822-838 (2016)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Uncertainty and Complexity Considerations in Food-Energy-Water Nexus Problems

Marcello Di Martino<sup>a,b</sup>, Patrick Linke<sup>c</sup>, and Efstratios N. Pistikopoulos<sup>a,b,\*</sup>

<sup>a</sup> Artie McFerrin Department of Chemical Engineering, Texas A&M University, 3122 TAMU, College Station, TX 77843, USA

<sup>b</sup> Texas A&M Energy Institute, Texas A&M University, 1617 Research Pkwy, College Station, TX 77845, USA

<sup>c</sup> Department of Chemical Engineering, Texas A&M University at Qatar, Texas A&M Engineering Building, Education City, PO Box 23874, Doha, Qatar

\* Corresponding Author: [stratos@tamu.edu](mailto:stratos@tamu.edu).

---

## ABSTRACT

The food-energy-water nexus (FEWN) has been receiving increasing interest in the open literature as a framework to address the widening gap between natural resource availability and demand, towards more sustainable and cost-competitive solutions. The FEWN aims at holistically integrating the three interconnected subsystems of food, energy and water, into a single representative network. However, such an integration poses formidable challenges due to the complexity and multi-scale nature of the three subsystems and their respective interconnections. Additionally, the significant input data uncertainty and variability, such as energy prices and demands, or the evaluation of emerging technologies, contribute to the system's inherent complexity. In this work, we revisit the FEWN problem in an attempt to elucidate and address in a systematic way issues related to its multi-scale complexity, uncertainty and variability. In particular, we provide a classification of the sources of data and technology uncertainty from historic data, forecasting and process parameters, and propose ways to quantify their impact on the integrated system analysis. To effectively tame the FEWN's multi-scale complexity, we distinguish between the introduced error of approximation and optimization of employed surrogate models. In turn, it is possible to characterize their impact on optimal FEWN decision-making based on the quantification of the introduced errors at all levels. Thus, we present strategies to systematically characterize FEWN process systems modeling and optimization. Ultimately, this facilitates translating obtained solutions into actionable knowledge by quantifying the level of confidence one can have in the derived process model and optimal results.

---

**Keywords:** Food & Agricultural Processes, Energy, Water, Design Under Uncertainty, Surrogate Model, Environment

## INTRODUCTION

The food-energy-water nexus (FEWN) has been identified in the open literature as an effective methodology to address the increasing disparity between natural food, energy and water resource availability and the respective resource demands sustainably. Based on a detailed analysis of the interconnections of the food, energy, and water subsystems, the FEWN aims to facilitate synergies among the three resource network subsystems, whereas competition is to be avoided to ultimately uncover trade-off solution strategies [1]. This proposed sustainable systems integration approach poses

formidable challenges from a process systems engineering (PSE) perspective, including but not limited to multi-scale uncertainty, appropriate system boundary definition, together with multiple and often conflicting stakeholder perspectives [2]. Furthermore, it has been shown that FEWN research does not produce sufficient actionable knowledge [3], meaning that obtained PSE FEWN frameworks and results are not translated into real world impact. In turn, the achievable synergy level of obtained optimal FEWN solutions is not only limited by the data availability, but also by the utilized modeling and complexity reduction techniques due to the introduced uncertainties and errors [4, 5].

Therefore, this work revisits the FEWN problem statement to systematically address its multi-scale complexity and uncertainty. More specifically, the scope of this work is to present methods to quantify introduced errors across all scales stemming from the necessity to overcome uncertainty, modeling, and computational complexity. While characterizing data uncertainty is well established in the literature, this work presents a novel distinction among the errors of approximation and optimization to fully characterize modeling complexity based on a representative surrogate energy-water supply model. Bridging the computational complexity of the surrogate model is evaluated via clustering of operational decisions of the water supply system. In turn, this work presents a framework to characterize introduced errors at all levels to provide decisionmakers, such as stakeholder or policymakers, with a quantifiable level of confidence one can have in the derived process model and optimal results.

The similarities among data uncertainty and complexity mitigation strategies are a central narrative of this work, the basis of which is elucidated via a representative energy-water supply system in the section FEWN Modeling. For impactful FEWN system modeling and optimization relevant system parameters have to be defined according to historic data which inherently introduces errors due to data uncertainty. Regarding data uncertainty the major challenges can be identified as (i) data variability, (ii) data aggregation, and (iii) data forecasting. All of this is discussed as part of the FEWN Modeling section. Due to the multi-scale complexity of the FEWN an interconnected system model cannot be optimized directly. In turn, modeling complexity mitigation via surrogate modeling, such as feedforward artificial neural networks (ANNs) or linear regression, and computational complexity reduction via algorithmic strategies, such as clustering or rolling horizon approaches, are discussed in the section System Complexity. It is important to note that the observations discussed in this work can readily be extended from FEWN approaches to general PSE problem statements tackling interconnected process systems and are essential for deriving actionable knowledge by quantifying the introduced errors stemming from data uncertainty and variability, as well as from the selected approximation and optimization methods.

## FEWN MODELING

The assumed system boundary inherently influences the FEWN system's modeling and solution generation. Accordingly, it has to be identified based on the defined research goals for impactful decision-making of FEWN frameworks [6]. This work investigates the quantification of complexity and data uncertainty mitigation errors across all scales of FEWN systems for the

derivation of actionable knowledge. Accordingly, a representative FEWN process system, based on benchmark resource supply systems, is selected as the relevant system boundary.

A renewable energy supply system for the identification of the cost-optimal mix of renewable energy generation and storage system has been selected as a representative energy supply system due to its reduced greenhouse gas emissions and therefore key role in the energy transition [7].

$$P_D(t) \leq \sum_t (P(t) - EC(t) + \sum_{t,k} [P_{rel,k}(t) - P_{stor,k}(t)]) \quad (1)$$

$$P(t) = \sum_{Tech} n_{Tech} \cdot p_{Tech}(t) \quad (2)$$

$$P_k^{bat}(t) = P_k^{bat}(t-1) + P_{stor,k}(t) - P_{rel,k}(t) \quad \forall k \in K \quad (3)$$

The governing equations of the energy supply system are summarized for every time point  $t \in T$  from Eq. (1) to (3). The energy demand  $P_D(t)$  has to be satisfied by generating renewable energy from different technologies such as biomass, wind turbines, single axis tracking or fixed angle solar panels (see Eq. (2)) or releasing energy from available energy storage systems ( $P_{rel,k}(t) \quad k \in K$ ). In turn, a surplus of renewable energy availability can be used to store energy ( $P_{stor,k}(t) \quad k \in K$ ) or satisfy the energy consumption of the water supply system ( $EC(t)$ ).

Reverse osmosis (RO) desalination water supply systems have already been identified as benchmark water supply systems in the literature and are accordingly selected for this study [8].

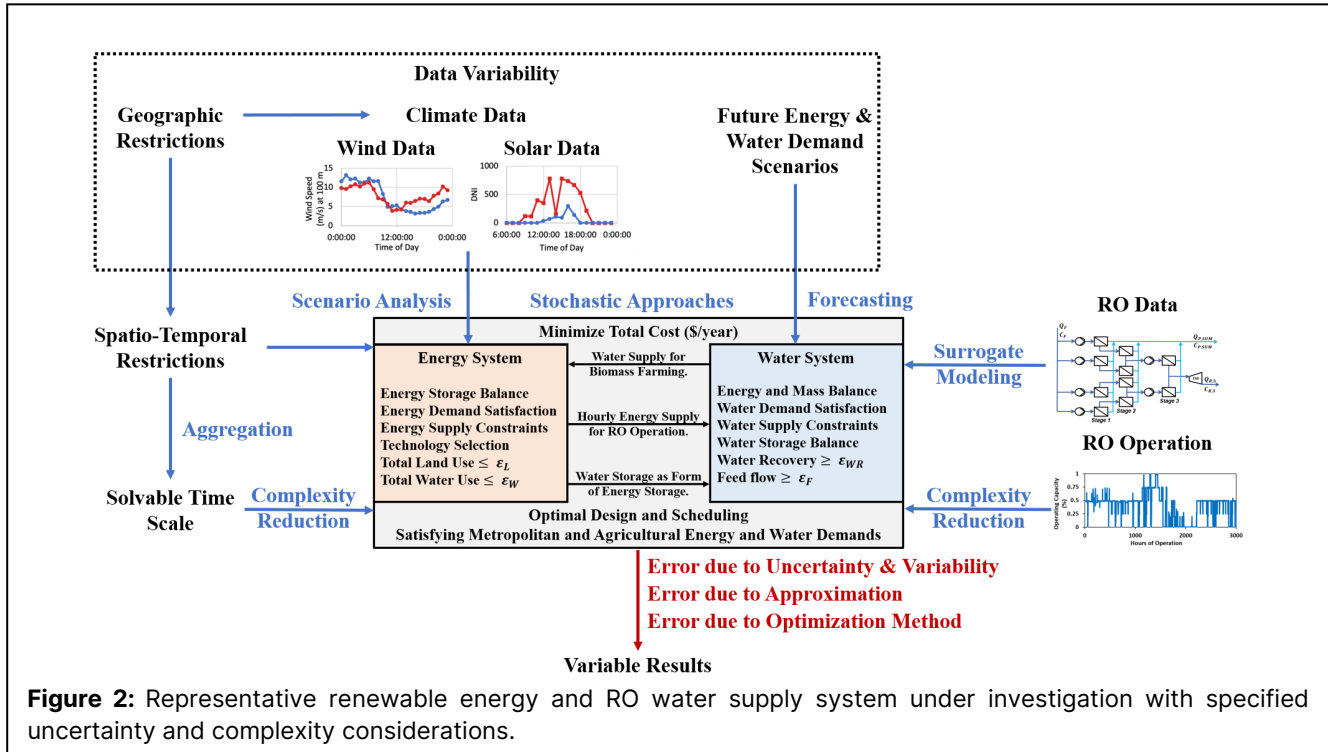
$$SEC(t) = \frac{1}{WR^{sys}(t)} \cdot \left( \frac{1}{4 \cdot \eta_1} \sum_{j=1}^4 P_{f,1,j}(t) + \dots \frac{(1-WR_1(t)) \cdot (1-WR_2(t))}{\eta_1} \left[ \frac{1}{2} \sum_{j=1}^2 P_{f,3,j}(t) - \frac{1}{4} \sum_{j=1}^4 P_{r,2,j}(t) \right] - \dots \frac{(1-WR_1(t)) \cdot (1-WR_2(t)) \cdot (1-WR_3(t)) \cdot \eta_2}{2} \sum_{j=1}^2 \Delta P_{ERD,j}(t) \right) \quad (4)$$

$$EC(t) = SEC(t) \cdot Q_p(t) \quad (5)$$

$$Q_p(t) = WR^{sys}(t) \cdot Q_f(t) \quad (6)$$

$$WR^{sys}(t) = WR_1(t) + (1 - WR_1(t)) \cdot WR_2(t) + (1 - \dots WR_1(t)) \cdot (1 - WR_2(t)) \cdot WR_3(t) \quad (7)$$

Eq. (4) to (7) summarize the governing RO desalination system equations. The specific energy consumption can be calculated based on the feed pressures ( $P_{f,i,j}(t)$ ) of every stage ( $i$ ) and parallel flow ( $j$ ), as well as water recoveries per stage ( $WR_i(t)$ ) and of the overall system ( $WR^{sys}(t)$ ), together with the pressure difference across the energy recovery device ( $\Delta P_{ERD,j}(t)$ ). The remainder of the RO system is characterized by the system's energy consumption ( $EC(t)$ ), together with its produced permeate ( $Q_p(t)$ ) and required feed flow ( $Q_f(t)$ ).  $Q_p(t)$  can then be used to satisfy the municipal water demand and the water demand for biomass farming of the renewable energy supply system. In addition, a water storage balance analogous to Eq. (3) can be added to the RO desalination system.



**Figure 2:** Representative renewable energy and RO water supply system under investigation with specified uncertainty and complexity considerations.

For semi-arid and arid water scarce regions with high temperatures and harsh external conditions, outdoor agriculture is neither a viable nor a sustainable option. Therefore, greenhouse farming systems providing a controlled environment for secure plant growth are selected as a representative food supply system [9].

To reduce the overall system complexity only the renewable energy supply system and the RO water supply system are considered in detail. The greenhouse system is considered via water and energy demand fulfillment constraints in addition to metropolitan energy and water demand constraints [10]. Accordingly, optimal solutions specify the design and operation of an interconnected renewable energy supply system and RO water supply system with energy and water storage consideration for the dynamic water and energy demand fulfillment as specified by the demands of an urban center and a greenhouse farming system.

The overall representative energy-water supply system for metropolitan and agricultural ecosystems is summarized together with all encountered data uncertainties and modeling complexities in Figure 2. The detailed mathematical model of the renewable energy supply system can be found in [7], of the RO water supply system in [8,11] and their modeled interconnections in [10]. The novelty of this work consists of revisiting the FEWN problem statement by systematically discussing mitigation strategies overcoming data uncertainty and variability, as well as system modeling and computational complexity. Dependent on the complexity mitigation strategy not only the feasible region of the optimization

model, but also the obtained optimal results in the form of the objective function and optimization variable values can potentially be altered [11]. To maintain the practicability of derived FEWN frameworks and results, it is of the utmost importance to quantify the introduced errors. This represents a certified level of confidence decision-makers can have in the proposed FEWN process systems solutions.

The identification of relevant data of reduced order compared to the historically available data remains one of the main challenges of solving FEWN problems toward practical impact. To operationalize the FEWN researchers have to not only mitigate data variability, but also derive representative data from a plethora of historical data (aggregation) and extrapolate currently available data to future trends (forecasting), as depicted in Figure 2. For all of these data uncertainty challenges, a sensitivity analysis of derived results can be leveraged as a proxy of the introduced error. One of the most direct ways to quantify the sensitivity of optimal solutions dependent on changing parameter values is multi-period or what-if scenario analysis. These approaches attempt to integrate uncertainty into policymaking through iterative exploration of possible optimal solutions [12]. Here, this can be understood as varying assumed input climate data profiles, such as solar direct normal irradiance (DNI) or wind speeds, or potential future energy and water demand profiles and prices. Other strategies to capture data variability include stochastic approaches which leverage probability distributions to approximate the variable input

data [13]. In this instance, the original deterministic optimization model becomes stochastic and can be solved via either Monte Carlo simulation, chance-constrained or robust programming.

To extrapolate patterns and behavior from historic data into the future forecasting techniques are utilized. Here, the introduced error of uncertainty stems from the quality of representative data utilized for forecasting and how accurate future trends can be captured based on this training data set. While there exist long term projections of future water demands, this is not the case for future energy demands and prices. To address this challenge, Baratsas et al. [14] established a quantitative framework to evaluate the integrated price of energy, the Energy Price Index (EPIC), which, influenced by both the demand and prices of energy products, signifies the average monthly price of energy for end-use consumers. To predict the present value for the demand and price of the energy products, a rolling horizon methodology is utilized since the availability of data for the demand and prices for certain energy products is often lagged by one to three months [15].

The interested reader is advised to consult [16] for a general review of advantages, disadvantages, and novelties of time series forecasting techniques applied to the energy sector. It is important to note that the degree of uncertainty significantly increases with the projected time scale of the forecast. This is for example illustrated by the divergence of long-term projections for industry sector energy consumptions due to varying base and final year assumptions and if non-monetary drivers are used to project material demand [17]. Even if long-term projections are generated via a rolling horizon approach, the introduced uncertainty increases with every forecasted time step due to the possibility of compounding the forecasting error.

Aggregation aims at reducing comprehensive historic data into a manageable representative data set which is subsequently used to replace the original historic data. This is for example applied to hourly or daily wind speed or solar direct normal irradiance profiles for renewable energy systems optimization or available subland clusters for land-use optimization. Accordingly, an error of uncertainty is introduced in the process system due to the scale reduction. However, this scale reduction also reduces the computational complexity which simultaneously introduces an error of optimization. Here, one can appreciate the interconnectedness of complexity reduction and data uncertainty. To address the computational burden associated with hourly time discretization and extensive time horizons in optimization models, Tso et al. [18] proposed a decomposition algorithm based on agglomerative hierarchical clustering (AHC), moving away from traditional k-means clustering, to preserve the time chronology of the input data. To select the optimal

number of clusters, a cut-off threshold, showing the percentage of slope decrease in within-cluster variance, is applied, which can be understood as the allowable aggregation error.

## SYSTEM COMPLEXITY

The individual renewable energy supply system and the RO water supply system are highly complex to model and optimize due to their nonlinear behavior. In fact, dependent on the selected time scale the interconnected model is computationally intractable [10]. However, solely optimizing each subsystem individually is not sufficient to uncover synergies among the supply systems [1]. In turn, researchers employ surrogate models together with advanced algorithmic strategies to overcome modeling and computational complexity [19]. While these are effective approaches to reduce system complexity, these methods introduce errors due to the leveraged approximation and optimization strategies.

The error of approximation is well studied in the literature and can readily be quantified based on the coefficient of determination ( $R^2$ ) or the mean squared error (MSE). On the contrary, the introduced error due to the selected optimization method can be understood as the introduced change in objective function value and degrees of freedom (DOF) due to the utilization of a surrogate model or advanced algorithmic strategy. The variation of DOF and objective value ( $Err_{opt}$ ) can be quantified by calculating the  $R^2$  (or MSE) of the "original" optimization model ( $x_i$ ) and the surrogate model optimization results ( $\hat{x}_i$ ), as shown in Eq. (8). In this case,  $R^2$  can be understood as the amount of original model variability explained by the surrogate model and optimization method which should ideally be one. To summarize all introduced errors,  $R^2$  for each approximation or optimization approach can be calculated and subsequently be multiplied to obtain a final value for the introduced error due to the necessity of system complexity mitigation.

$$Err_{opt} = 1 - R^2 = 1 - \frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (8)$$

$$Err_{obj} = 1 - \frac{\hat{y}}{y} \quad \hat{y}, y \geq 0 \quad (9)$$

Further, the change in objective function value can be quantified based on Eq. (9), where  $Err_{obj}$  denotes the introduced error based on the derived objective function values. In this case,  $\hat{y}$  denotes the optimal solution obtained via the surrogate optimization model, whereas  $y$  represents the originally obtained optimal solution. Based on this definition we can conclude that if

- $Err_{obj} > 0$ , then  $\hat{y}$  is a lower bound for  $y$ .
- $Err_{obj} < 0$ , then  $\hat{y}$  is an upper bound for  $y$ .

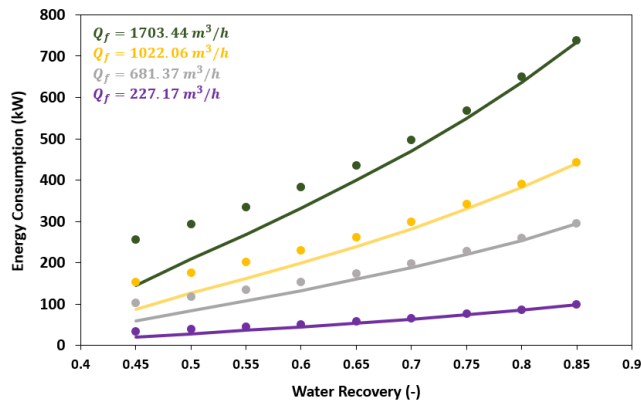
This deduction is only valid if the objective function values are positive. For negative objective function values,

the reversible observations are applicable.

## Modeling Complexity

As previously mentioned, the RO water supply system is selected to illustrate the influence of surrogate modeling on optimal decision-making. We selected simple linear regression, first-order Taylor approximation, multiple linear regression, and feedforward artificial neural networks (ANNs) with rectified linear units (ReLU) as activation functions as potential surrogate models. Shallow ANNs with a limited number of neurons utilizing ReLUs can be interpreted as a multiple variable piecewise linear regression. If one of these models, in the presented sequence, satisfied a set quality of approximation criterion ( $R^2 \geq 0.95$ ) the corresponding model has been selected to balance simplicity and accuracy.

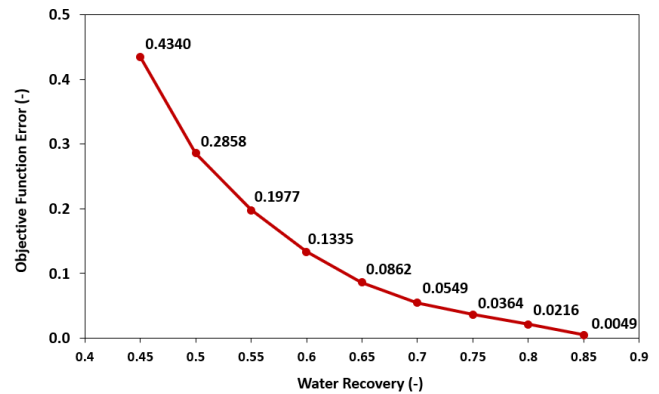
Here, the specific energy consumption ( $R^2=0.99$ ) and the energy consumption ( $R^2=0.99$ ) of the RO plant can both be calculated based on individual shallow ANNs utilizing ReLUs as activation functions. The RO plant's overall permeate flow ( $R^2=0.98$ ) and water recovery ( $R^2=0.98$ ) can be approximated via first-order Taylor expansions. In this instance, the surrogate model explains approximately 94% of the observed variability of the original model and thus we can certify the introduced error due to approximation to be 6%. Accordingly, the original mixed-integer nonlinear programming (MINLP) problem can be successfully approximated by a mixed-integer linear programming problem (MILP) [11].



**Figure 3:** Results of RO modeling and optimization. Lines: Original MINLP model; Points: MILP Surrogate model.

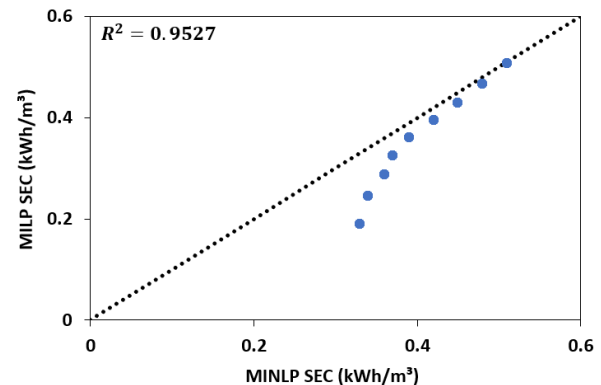
In turn, the error of optimization can be quantified by optimizing both the MINLP and MILP for the same set of constraints. The optimization models have been solved in GAMS with the Gurobi solver for the MILP problem and the BARON solver for the MINLP problem. Results of this analysis in terms of a direct comparison between the optimal energy consumption of the RO plant for changing feed flows and efficiencies of the MINLP and MILP problem are displayed in Figure 3. Here, we can appreciate how the differences in minimized energy consumption

between the MILP and MINLP model shrink with increasing water recovery and decreasing feed flow. The difference between the two optimization approaches results in  $R^2=0.98$ , meaning that solely based on the optimization approach an additional error of 2% is introduced in the objective function values. This change in objective function value can also be quantified by calculating  $Err_{obj}$  (Eq. (3)) for every pair of feedflow and water recovery. Interestingly, the error of objective function value does not change significantly with changes in feed flow, which is exemplified by a maximum standard deviation of 0.0026 across all values.



**Figure 4:** Results of RO modeling and optimization. Introduced error due to the selected optimization method.

Accordingly, the average  $Err_{obj}$  for each water recovery has been calculated which is depicted in Figure 4. On first glance we can postulate that the MILP problem provides in all cases a lower bound to the true optimal solution ( $Err_{obj} > 0$ ). It is interesting to see how the error of optimization decreases with increasing water recovery, i.e. system efficiency. For water recoveries higher than 0.6, the introduced error in the objective function value is less than 0.133. Besides the change in objective function value, the utilization of surrogate models can also influence the obtained values of optimization variables.



**Figure 5:** Results of RO optimization. Comparison of the SEC of system solutions for the MINLP and MILP case. For RO systems, key performance indicators besides the



plant's energy consumption are the system's overall water recovery, overall permeate output and specific energy consumption. Both, the MINLP and MILP optimization problem, result in the same system water recovery and permeate output, meaning that the RO systems operate at the same efficiency and output level.

Interestingly, the SEC consumption of both models differ due to the selection of different operational pressures and water recoveries per stage across the plant (see Eq. (4) and (7)). This is possible since the same water recovery of the system can be achieved with different water recoveries per stage and since the same water recovery per stage can be achieved with different pressures per stage and parallel flow across the plant. To measure this difference in operational decisions the SEC for the MINLP and MILP case are visualized in Figure 5. It is important to note that it has been shown that the SEC is only dependent on the water recovery [8] explaining the reduction in data points between Figures 3 and 5, similar to the observation regarding Figures 3 and 4. Ideally, the data points would follow a straight line of slope one indicating that the SEC results of the respective models are identical. Since this is not the case, we can appreciate the introduced optimization method error as departures from said line, which can be quantified as  $R^2=0.95$ . In turn, this represents an introduced error of 5% in terms of decision variable values.

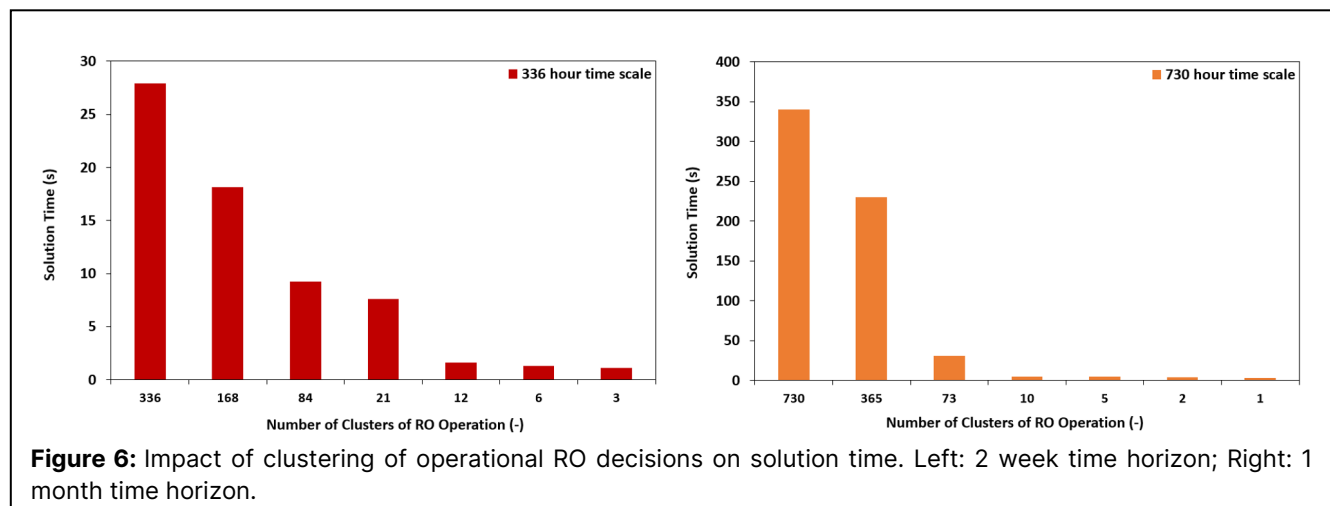
Taking into account this additional error, the optimization model can only capture 88% of the original model's behavior, instead of the initial 94%. Arguably, the error due to the selected approximation approach is approximately the same as the error introduced stemming from the selected optimization approach (both 6%). This observation indeed underlines the importance of quantifying the quality of obtained surrogate model optimization results, although this is a difficult task which might not always be possible because of the potential non-solvability of the underlying physics-based models.

## Computational Complexity

The spatial and temporal multi-scale complexity of the FEWN results in large scale optimization models, which is one of the main reasons that derived optimization frameworks are computationally intractable. It already has been shown that the full energy-water supply system cannot be solved at an hourly time scale over a one-year time horizon. To obtain an upper bound on the true optimal solution, the RO system has been restricted to a single operating point whereas the full dynamic renewable energy system is solved over a one-year time horizon [10]. In this extreme case, the RO system is effectively clustered to a single representative operating point specifying only once its constant  $EC$ ,  $Q_p$ ,  $Q_f$  and  $WR^{sys}$ . This is analogous to assuming that there exists only a single time period of operation of the RO system according to a "short-sighted" rolling horizon approach and can effectively be understood as introducing a different time scale for the water supply system. In turn, this inherently limits the system's flexibility and potentially its ability to adjust to the variability of the renewable energy supply system. Accordingly, altering the frequency of operational clustering corresponds to evaluating the importance of considering varying or equal time scales among the interconnected supply systems.

However, the quality of the obtained upper bound on the true optimal solution has not been quantified yet. In theory, better upper bounds to the true optimal solution can be obtained by adding more operational RO clusters to the system. This clustering of operating points can directly be compared to similar approaches for data aggregation. Alternatively, instead of solving the optimization model over a one-year time horizon, the model can only be solved for one month and then resolved 12 times to approximate the annual solution, similar to a rolling horizon strategy [20].

To evaluate the potential of operational RO clustering for complexity mitigation, this work solves the full energy-water model for two weeks at an hourly time scale



(336 hours), as well as for one month at an hourly time scale (730 hours). In both instances, the RO operation is successively restricted from the full dynamic model to a single operating point. Special attention is placed on how the solution times and objective function values change with increasing operational restriction. It is expected that in all cases  $Err_{obj} \leq 0$  since this clustering approach of the RO system can only provide upper bounds to the true optimal solution. For all cases an optimal solution and mixed-integer programming (MIP) gap of 0% is enforced.

As shown in Figure 6, the solution time decreases significantly with decreasing number of clusters for both cases. However, starting at 12 clusters for the 336 hour time scale and at 10 clusters for the 730 hour time scale no significant speed-ups in solution time can be observed. Moreover, for both cases,  $Err_{obj}$  is negligible over the complete cluster horizon as illustrated by a mean of  $-1.4215 \cdot 10^{-6}$  (standard deviation of  $2.3877 \cdot 10^{-7}$ ) and  $-1.4640 \cdot 10^{-4}$  (standard deviation of  $3.9754 \cdot 10^{-7}$ ), for the 2 week and the 1 month case study, respectively. These results suggest that there is a cluster threshold after which no speed-up in solution time and no significant changes in objective function value are expected, meaning there is an advantageous trade-off between operational flexibility (number of clusters), system solvability (solution time) and accuracy ( $Err_{obj}$ ). However, we can also deduct that collapsing a bi-weekly or monthly operation in a single operating point does not result in a significant loss of information ( $Err_{obj} < 1.4 \cdot 10^{-4}$ ). Therefore, it is suggested to evaluate how the solution of the full-time scale FEWN problem changes with increasing RO operational flexibility (yearly, monthly, bi-weekly, daily, every 12 hours). In this case the model was solved to an optimal solution and MIP gap of 0%.

**Table 1:** Results of solving the full-scale energy-water nexus supply model (8760 hours). Model size is specified according to the dimensions of the constraint matrix.

RO Operation	Clusters	Model Size	$Err_{obj}$
Yearly	1	[175331, 122778]	Basis
Monthly	12	[176310, 123801]	$3.925 \cdot 10^{-6} \%$
Bi-Weekly	24	[177378, 124917]	$7.860 \cdot 10^{-6} \%$
Daily	365	[207727, 156630]	$1.426 \cdot 10^{-5} \%$
12 hours	730	[240212, 190575]	$1.630 \cdot 10^{-5} \%$

Interestingly, the objective function value does not change for all practical purposes by allowing for more flexible operational RO decisions over one year, as shown in Table 1. In addition, the energy and water storage

schedules differ only insignificantly among obtained solutions. For all intents and purposes, the obtained optimal solutions are virtually identical. The solution times on the other hand vary from 250 seconds for the yearly case to 6400 seconds (around 2 hours) and even 53,000 seconds (more than 14 hours) for the daily and every 12 hours RO operational decision strategy, respectively. For the case study presented here, the RO clustering does not seem to have an impact on the overall solution accuracy. This novel analysis overall strengthened our confidence in the harsh sounding assumption of allowing only one RO operating point over one year to obtain a valid upper bound on the true optimal solution. However, this statement can only be generalized if the solution of the full-time scale FEWN model can be obtained, i.e. via decomposition strategies. It is expected that there exists a solution drop-off with an increasing number of clusters towards the full-time scale model. The obtained results can also aid in the convergence time of the full-scale optimization model by providing valid upper bounds to the problem which potentially shrinks the solution gap significantly depending on the quality of the lower bound.

## CONCLUSION

In this work, we revisited the FEWN problem statement by addressing data uncertainty and system complexity through the lens of a representative surrogate energy-water supply system. Sources of introduced data uncertainty and complexity mitigation errors have been discussed in detail. While it is difficult to quantify the introduced modeling and data uncertainty errors it is an essential task to maintain confidence in derived FEWN optimization frameworks and results. Implications of approximation and optimization methods for a RO water supply system have been discussed in detail, which has then been extended to an operational RO clustering approach to mitigate the computational complexity of the interconnected energy-water supply system optimization.

Policy makers and stakeholders are crucial in translating obtained PSE FEWN methodologies and results into real-world implementation. Being able to quantify a level of confidence one can have into proposed modeling strategies aids in this process. The proposed error quantification strategies can be leveraged to certify how much uncertainty in the obtained optimal solution is added with every data processing and modeling step. In turn, this uncertainty can be utilized to provide interpretable ranges of optimal results instead of single values, similar to confidence intervals, for transparent and effective decision-making across all levels.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge support from Texas A&M University, USA, the Texas A&M Energy Institute, USA, and Texas A&M University at Qatar.

## REFERENCES

1. M. Di Martino, P. Linke, E.N. Pistikopoulos. A comprehensive classification of food-energy-water nexus optimization studies: State of the art. *J. Clean. Prod.* 420 (2023).
2. D.J. Garcia, F. You. The water-energy-food nexus and process system engineering: A new focus. *Comput. Chem. Eng.* 91, 49-67 (2016).
3. M. Dalla Fontana, D. Wahl, F. de Araujo Moreira et al. The five Ws of the water-energy-food nexus: a reflective approach to enable the production of actionable knowledge. *Front. Water* 3 (2021).
4. B.A. McCarl, Y. Yang, R. Srinivasan et al. Data for WEF nexus analysis: A Review of Issues. *Curr. Sustain./Renew. Energy Rep.* 4, 137-143 (2017).
5. B.A. McCarl, Y. Yang, K. Schwabe et al. Model Use in WEF Nexus Analysis: a Review of Issues. *Curr. Sustainable Renewable Energy Rep* 4, 144-152 (2017).
6. B. Daher, R.H. Mohtar, S.-H. Lee, A. Assi. Modeling the Water-Energy-Food Nexus - A 7-Question Guideline. *Geophysical Monograph Series* (2017).
7. J. Cook, M. Di Martino, R.C. Allen, E.N. Pistikopoulos, S. Avraamidou. A decision-making framework for the optimal design of renewable energy systems under energy-water-land nexus considerations. *Sci. Total Environ.* 827, 154185 (2022).
8. M. Di Martino, St. Avraamidou, E.N. Pistikopoulos. A Neural Network Based Superstructure Optimization Approach to Reverse Osmosis Desalination Plants. *Membranes* 12(2), 199 (2022).
9. F. Mahmood, R. Govindan, A. Bermak et al. Energy utilization assessment of a semi-closed greenhouse using data-driven model predictive control. *J. Clean. Prod.* 324, 129172 (2021).
10. M. Di Martino, P. Linke, E.N. Pistikopoulos. Towards Optimal Energy-Water Supply System Operation for Agricultural and Metropolitan Ecosystems. *FOCAPO/CPC* 2023.
11. M. Di Martino, P. Linke, E.N. Pistikopoulos. Modeling strategies in multi-scale food-energy-water nexus system optimization. *Comput. Aided Chem. Eng.* 52, 51-56 (2023).
12. L. Yung, E. Louder, L.A. Gallagher, K. Jones, C. Wyborn. How Methods for Navigating Uncertainty Connect Science and Policy at the Water-Energy-Food Nexus. *Front. Environ. Sci.* 7, 37 (2019).
13. G.E. Constante-Flores, M.S. Illindala. Data-Driven Probabilistic Power Flow Analysis for a Distribution System With Renewable Energy Sources Using Monte Carlo Simulation. *IEEE Transactions on Industry Applications* 55 (1), 174-181 (2019).
14. S.G. Baratsas, A.M. Niziolek, O. Onel et al. A framework to predict the price of energy for the end-users with applications to monetary and energy policies. *Nat. Commun.* 12, 18 (2021).
15. S.G. Baratsas, A.M. Niziolek, O. Onel et al. A novel quantitative forecasting framework in energy with applications in designing energy-intelligent tax policies. *Appl. Energy* 305, 117790 (2022).
16. C. Deb, F. Zhang, J. Yang, S.E. Lee, K.W. Shah. A review on time series forecasting techniques for building energy consumption. *Renew. Sustain. Energy Rev.* 74, 902-924 (2017).
17. O.Y. Edelenbosch, K. Kermeli, W. Crijns-Graus et al. Comparing projections of industrial energy demand and greenhouse gas emissions in long-term energy models. *Energy* 122, 701-710 (2017).
18. W.W. Tso, C.D. Demirhan, C.F. Heuberger, J.B. Powell, E.N. Pistikopoulos. A hierarchical clustering decomposition algorithm for optimizing renewable power systems with storage. *Appl. Energy* 270, 115190 (2020).
19. M. Di Martino, R.C. Allen, E.N. Pistikopoulos. The Food-Energy-Water Nexus in Sustainable Energy Systems Solutions. In: Fathi, M., Zio, E., Pardalos, P.M. (eds) *Handbook of Smart Energy Systems*. Springer, Cham (2022). [https://doi.org/10.1007/978-3-030-72322-4\\_168-1](https://doi.org/10.1007/978-3-030-72322-4_168-1)
21. G. Homsj, B. Gendron, S.D. Jena. Rolling horizon strategies for a dynamic and stochastic ridesharing problem with rematches. *Discret. Appl. Math.* 343, 191-207 (2024).

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Optimal Transition of Ammonia Supply Chain Networks via Stochastic Programming

Ilias Mitrai<sup>a</sup>, Matthew J. Palys<sup>a</sup>, and Prodromos Daoutidis<sup>a\*</sup>

<sup>a</sup> University of Minnesota, Department of Chemical Engineering and Materials Science, Minneapolis, 55455 MN, US

\* Corresponding Author: daout001@umn.edu.

---

## ABSTRACT

This paper considers the optimal incorporation of renewable ammonia production facilities into existing supply chain networks which import ammonia from conventional producers while accounting for uncertainty in this conventional ammonia price. We model the supply chain transition problem as a two-stage stochastic optimization problem which is formulated as a Mixed Integer Linear Programming problem. We apply the proposed approach to a case study on Minnesota's ammonia supply chain. We find that accounting for conventional price uncertainty leads to earlier incorporation of in-state renewable production sites in the supply chain network and a reduction in the quantity and cost of conventional ammonia imported over the supply chain transition horizon. These results show that local renewable ammonia production can act as a hedge against the volatility of the conventional ammonia market.

---

**Keywords:** Design and Sustainability, Stochastic Optimization, Capacity Expansion, Supply Chain Optimization, Green Ammonia

## INTRODUCTION

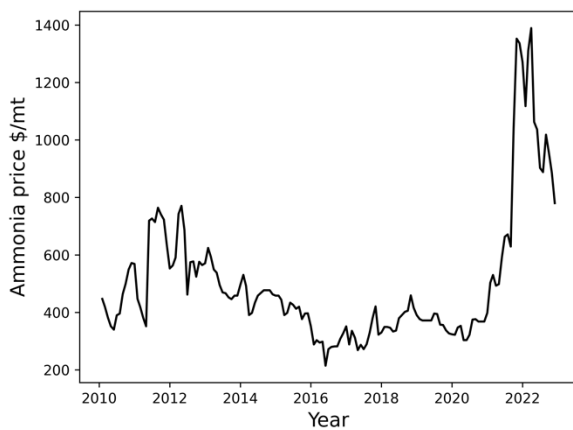
Ammonia is one of the most important industrial chemicals and serves as the backbone of modern agriculture in its use either directly or as a precursor to other nitrogen fertilizers. The standard production of ammonia is based on the Haber-Bosch process, which uses fossil fuels as the feedstock hydrogen source and operates at high pressure and temperature [1]. These facilities generally have capacities greater than 1,000,000 metric tons per year (mt/y) to take advantage of economies of scale [2]. This production paradigm leads to high transportation costs and carbon emissions in the operation of the supply chain because ammonia is transported through national and even global networks of ships, pipelines, rail, and trucks from a few production sites to the final customers [3].

The transition to a more sustainable supply chain network can be achieved by reducing the carbon emissions related to the manufacturing and distribution of ammonia. Renewable or green ammonia production recently has been the subject of extensive research and development as an alternative to the standard ammonia manufacturing paradigm [4]. In this approach, renewable

resources such as wind and solar are used to produce hydrogen via electrolysis and nitrogen via air separation, reducing the carbon intensity of producing ammonia. The Midwest region of the United States uses the most nitrogen fertilizer in the country while also being home to rich wind resources [5]. This gives rise to an opportunity to produce renewable ammonia closer to where it is used, thus reducing the cost and carbon intensity of ammonia distribution [6]. Producing ammonia using renewable energy also offers the potential for ammonia production cost stability. The feedstock renewable energy can have a close-to-constant price in this production setting, whether this energy is sourced through multi-year power purchase agreements (PPA) or the ammonia producer owns and operates the necessary renewable generation. In contrast, ammonia is currently traded on a global market and its price is subject to variability due to a number of factors including natural gas prices, food prices, and global conflict (see Figure 1). Given the transformative potential of renewable ammonia, achieving economical deployment through optimal design of manufacturing facilities and the supply chain network is of critical importance. In this work, we focus on the latter.

The transition of existing ammonia supply chain

networks to incorporate renewable production will likely occur over multiple years and will be affected by multiple sources of uncertainty. Identifying optimal investment decisions over a fixed planning horizon is a widely studied problem in process systems engineering and operations research and is formally known as the capacity expansion problem [7]. However, the application of the capacity expansion formalism to the transition of ammonia supply chain networks is rather limited. Recently, we have proposed a multiperiod deterministic capacity expansion model that considers the optimal transition of ammonia supply chain networks [8]. The model optimizes the investment decisions regarding the installation year and capacity, such that the overall net present cost is minimized while ammonia fertilizer demand is satisfied.



**Figure 1.** U.S. Gulf Coast ammonia price from 2010 to 2022 [9].

In this work, we consider the effect of uncertainty on the optimal transition of existing ammonia supply chain networks. The primary sources of uncertainty in an ammonia supply chain are the ammonia demand and the market price of ammonia. Although the demand for ammonia can be predicted from total fertilizer demand estimates, the ammonia price is more volatile. Accounting for the significant price variability and uncertainty is essential for the optimal expansion of existing supply chain networks.

We propose a two-stage stochastic programming approach where the uncertainty in price is accounted for in the form of scenarios [10,11]. Such a conceptual approach has been previously employed in supply chain optimization models in a number of different industries, for example, waste-to-bioethanol [12], biodiesel production from wastewater treatment byproducts [13], and coal-to-liquids [14]. In our model, the installation decisions (the location, capacity, and construction year for new renewable ammonia manufacturing facilities) are the first stage decisions, and the distribution of ammonia from the installed renewable sites and the conventional producers

to the customers for the different ammonia prices (scenarios) are the second stage decisions. We consider a case study on Minnesota’s ammonia supply chain network. The results show that accounting for uncertainty in the price of ammonia, especially high prices, requires investments earlier in the planning horizon, compared to assuming a nominal price. Furthermore, we simulate the supply chain obtained from the deterministic and stochastic models, and we find that for high ammonia prices, the design obtained via stochastic programming results in lower net present costs. These results highlight the ability of locally-produced renewable ammonia to act as a hedge against high prices on the conventional ammonia market. The rest of the paper is organized as follows: First, we present the two-stage stochastic optimization model, then we present the case study and, finally, the numerical results.

## TWO-STAGE STOCHASTIC OPTIMIZATION MODEL

We consider an existing supply chain network that delivers ammonia to a set of counties  $\mathcal{C} = \{1, \dots, C\}$  via distribution centers  $\mathcal{D} = \{1, \dots, D\}$ . In the original network, the demand  $\delta_c$  at each county is satisfied by purchasing ammonia from conventional producers  $\mathcal{P} = \{1, \dots, P\}$  with price  $\alpha_p$ . Given a set of candidate locations for renewable ammonia production facilities  $\mathcal{R} = \{1, \dots, R\}$ , the goal is to find the optimal investment decisions over a planning horizon  $\mathcal{K}$ , such that the total net present cost of the supply chain is minimized, demands are met for each period of the planning horizon, and at the end of the horizon the entire demand is satisfied using renewable ammonia. We assume that the capacity investment decisions are made annually and the planning horizon  $\mathcal{K}$  is discretized into  $K$  time periods. We define variable  $x_{rk}$  as the capacity installed at candidate renewable site  $r$  at time period  $k$ , and binary variable  $z_{rk}$  which is equal to one if an investment is made at candidate site  $r$  at time period  $k$  and zero otherwise. We assume that the only uncertain parameter is the price of ammonia imported from conventional producers. We model the renewable ammonia production investment decisions, specifically the time period when an investment is made  $z_{rk}$  and the production capacity  $x_{rk}$  at a given candidate location  $r$ , as first-stage decisions. The amount of ammonia sent to each county through a combination of purchases from conventional producers routed through distribution centers and from new renewable production facilities are the second stage decisions. We follow a scenario-based formulation and define the set  $\mathcal{S} = \{1, \dots, S\}$  which represents the scenarios of the price of ammonia, where each scenario has probability  $p_s$ , and the price of ammonia for producer  $p$  and scenario  $s$  is  $\alpha_{ps}$ .

Given this problem setting, first, we define



constraints related to the maximum and minimum capacity that can be installed in each location and time period by the following constraints

$$x_{rk} \leq \bar{x}^U z_{rk} \quad \forall r \in \mathcal{R}, k \in \mathcal{K} \quad (1)$$

$$x_{rk} \geq \bar{x}^L z_{rk} \quad \forall r \in \mathcal{R}, k \in \mathcal{K}, \quad (2)$$

where  $\bar{x}^U, \bar{x}^L$  are the upper and lower bounds on the size of renewable sites. Each renewable candidate site has a certain wind capacity  $\Omega_r$ , electrolysis capacity  $\Omega_{k_i}$ , and a construction period of two years, which constrain the maximum capacity that can be installed and the time that the capacity is available as follows

$$\sum_{k'=1}^k x_{rk'} \omega_{rk'} \leq \Omega_r \quad \forall r \in \mathcal{R}, k \in \mathcal{K} \quad (3)$$

$$\sum_{r \in \mathcal{R}} x_{rk} \xi_{rk} \leq \Xi_k \quad \forall k \in \mathcal{K}. \quad (4)$$

We define variable  $y_{pdks}$  as the amount of ammonia purchased from conventional producer  $p$  and shipped to distribution center  $d$  at time period  $k$  and scenario  $s$ . We also define variable  $y_{dcks}$  as the amount of ammonia shipped from distribution center  $d$  to county  $c$  at time period  $k$  and scenario  $s$ , and the amount of ammonia shipped from the renewable site  $r$  to county  $c$  at time period  $k$  and scenario  $s$  is  $y_{rcks}$ . The demand satisfaction constraints are

$$\sum_{r \in \mathcal{R}} y_{rcks} + \sum_{d \in \mathcal{D}} y_{dcks} \geq \delta_{ck} \quad \forall s \in \mathcal{S}, c \in \mathcal{C}, k \in \mathcal{K} \quad (5)$$

$$\sum_{p \in \mathcal{P}} y_{pdks} \geq \sum_{c \in \mathcal{C}} y_{dcks} \quad \forall s \in \mathcal{S}, d \in \mathcal{D}, k \in \mathcal{K} \quad (6)$$

$$\sum_{d \in \mathcal{D}} y_{pdks} \leq \Lambda_p \quad \forall s \in \mathcal{S}, p \in \mathcal{P}, k \in \mathcal{K} \quad (7)$$

$$\sum_{c \in \mathcal{C}} y_{rcks} \leq \sum_{k'=1}^{k-2} x_{rk'} \quad \forall s \in \mathcal{S}, r \in \mathcal{R}, k \in \mathcal{K} \quad (8)$$

$$\sum_{p \in \mathcal{P}} \sum_{d \in \mathcal{D}} y_{pdks} = 0 \quad \forall s \in \mathcal{S}. \quad (9)$$

The objective function is the net present cost of the supply chain transition over the planning horizon. It can be partitioned into two terms. The first term,  $Z_k$ , is the sum of the capital  $CAP_k$  and operating costs  $OP_k$  which depends only on the first-stage decisions, and are computed as follows

$$CAP_k = \frac{1}{\theta} \sum_{r \in \mathcal{R}} \sum_{k'=1}^k x_{rk'} \sigma_{rk'} + z_{rk'} \gamma_{rk'} \quad (10)$$

$$OP_k = \sum_{r \in \mathcal{R}} \sum_{k'=1}^k x_{rk'} \zeta_{rk'}. \quad (11)$$

The capital cost of a renewable production facility is modeled as a piece-wise affine function of the installed capacity, with slope  $\sigma_{rk}$  and intercept  $\gamma_{rk}$ , to capture the effect of economies of scale. These parameters vary with both renewable site  $r$  and time period  $k$  to capture the effects of varying renewable potential and expected technology cost reductions respectively. The capital cost is annualized using scaled plant lifetime  $\theta$  which is equal to  $10.23 \text{ y}^{-1}$ . The operating cost is assumed to scale linearly with the installed capacity with proportionality constant  $\zeta_{rk}$ . This parameter is assumed to remain constant after installation and also varies with renewable site and

period to capture the effects described above.

The second term in the objective,  $\bar{Z}_{ks}$ , is the sum of the distribution of renewable ammonia  $DR_{ks}$ , transportation of ammonia from the conventional producers to the distribution centers  $TC_{ks}$ , and distribution of conventional ammonia  $DC_{ks}$ . The individual costs are equal to

$$DP_{ks} = \sum_{r \in \mathcal{R}} \sum_{c \in \mathcal{C}} y_{rcks} \tau_{rc} \quad \forall k \in \mathcal{K}, s \in \mathcal{S} \quad (12)$$

$$PC_{ks} = \sum_{p \in \mathcal{P}} \sum_{d \in \mathcal{D}} y_{pdks} \alpha_{ps} \quad \forall k \in \mathcal{K}, s \in \mathcal{S} \quad (13)$$

$$TC_{ks} = \sum_{p \in \mathcal{P}} \sum_{d \in \mathcal{D}} y_{pdks} \tau_{pd} \quad \forall k \in \mathcal{K}, s \in \mathcal{S} \quad (14)$$

$$DC_{ks} = \sum_{d \in \mathcal{D}} \sum_{c \in \mathcal{C}} y_{dcks} \tau_{dc} \quad \forall k \in \mathcal{K}, s \in \mathcal{S} \quad (15)$$

$$Z_k = CAP_k + OP_k \quad \forall k \in \mathcal{K} \quad (16)$$

$$\bar{Z}_{ks} = DC_{ks} + PC_{ks} + TC_{ks} + DP_{ks} \quad \forall k \in \mathcal{K}, s \in \mathcal{S}. \quad (17)$$

The two-stage optimization problem is

$$\begin{aligned} \min & \sum_{k \in \mathcal{K}} \phi_k Z_k + \sum_{s \in \mathcal{S}} \rho_s (\sum_{k \in \mathcal{K}} \phi_k \bar{Z}_{ks}) \\ \text{s. t.} & \text{Eq. 1 - 9} \end{aligned} \quad (18)$$

$$x_{rk} \geq 0, z_{rk} \in \{0,1\}$$

$$y_{rcks} \geq 0, y_{dcks} \geq 0, y_{pdks} \geq 0$$

The parameter  $\phi_k$  is the discounted value of cost contributions in time period  $k$ .

## CASE STUDY

We herein provide a concise description of the case study on Minnesota's ammonia supply chain. For a more detailed case study description, please refer to our previous work [8]. We consider a supply chain planning horizon from 2024 to 2032. We consider an 8.5% discount rate as it pertains to cost flows beyond 2024. Minnesota has 82 counties in which ammonia demand must be met for each period of the supply chain transition optimization. In 2024, the total ammonia demand assumed to be 795,000 mt/y and this is assumed to increase by 0.5% per year. This ammonia can be purchased from 10 conventional producers which are located outside of Minnesota. The average conventional ammonia price from 2010 to 2022 was \$500/mt. Ammonia purchased from these producers is routed through one of three distribution centers in Minnesota. Conventional producers located further from Minnesota have higher associated costs to transport ammonia to a given distribution center. Producer-to-distribution transportation costs range from \$59/mt to \$141/mt, while distribution center-to-county transportation costs range from \$1/mt to \$36/mt.

We consider 26 candidate locations for new in-state renewable ammonia production. Capital costs of a given facility are incurred two years before renewable production begins to represent a two year construction period. The operating cost includes renewable power purchases from PPAs with wind generators assumed to be co-

located with the renewable ammonia production facility. The operating cost also includes revenue from hydrogen production tax credits (PTCs) contained in the U.S. federal government Inflation Reduction Act [15]. These credits provide \$3/kg of hydrogen produced for the first 10 years of facility operation. Renewable ammonia is assumed to be transported from new production facilities directly to counties. These transportation costs range from \$1/mt to \$47/mt. The capacity of each new renewable ammonia facility is constrained to a maximum associated wind generation capacity of 250 MW (Eq. 3). The total installed capacity of renewable ammonia production across all facilities in a given year is constrained by electrolysis availability, which increases from 250 MW in 2024 to 850 MW in 2032 (Eq. 4). The optimization model is implemented in Pyomo [16] and is solved using Gurobi 10.0.2.0 [17] on an Apple MacBook Pro M1 with 8 physical cores and 16 GB of RAM.

## NUMERICAL RESULTS

### Deterministic case

First, we solve a deterministic model using a \$500/mt historical average conventional ammonia price, i.e., using the two-stage model with one scenario with probability one and price \$500/mt. The optimization problem has 22,194 (234 binary and 21960 continuous) variables, 1,803 constraints and is solved in 6.7 seconds. The total net present cost is \$3,002 million (MM) and renewable ammonia production is installed at eight new sites, as presented in Table 1. Three renewable ammonia facilities are installed in 2027, meaning that in-state production does not begin until 2029 and all ammonia is purchased from the conventional market for the first five years of the planning horizon (see Figure 3).

**Table 1:** Installation year, location, and capacity for new renewable ammonia production in the deterministic case. Annual average wind capacity factors for each selected candidate location are provided in parentheses to describe wind potential.

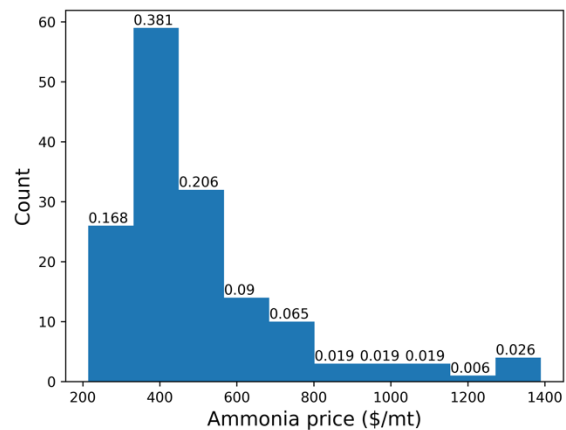
Year	Location	Capacity (1,000 mt/y)
2027	Lake Wilson (52%)	121.4
2027	Chandler (52%)	121.4
2027	Worthington (52%)	55.2
2028	Luverne (52%)	121.2
2028	Wilmont (52%)	110.4
2028	Worthington (52%)	66.0
2030	Blue Earth (47%)	116.1
2030	Winnebago (47%)	114.8

Three additional facilities are installed in 2028. These first six facilities are all located in Southwest

Minnesota, which has the highest wind capacity factors at 52%. In both 2027 and 2028, 575 MW of electrolysis is procured, the maximum allowable amount in each year. This is the reason that two smaller facilities are installed in Worthington in consecutive years. The final facilities are installed in 2030 to ensure that all ammonia is obtained via renewable production by 2032. These are both installed in Southeast Minnesota, which also has a high wind capacity factor at 47%. With the exception of the two smaller facilities in Worthington, all others use at least 225 MW of co-located wind generation in an attempt to achieve economies of scale.

### Stochastic case

We use the data presented in Figure 1 and generate the histogram presented in Figure 2 using ten bins, and obtain the price at the edge of each bin and the number of data points in each bin. Given these data, we generate the scenarios where the price of ammonia in scenario  $s$  is set equal to the edge price in the bin  $s$  and the probability  $p_s$  is the number of data points in the bin  $s$  divided by the total number of data points.



**Figure 2.** Histogram of U.S. Gulf Coast ammonia prices from 2010 to 2022. The width of each histogram bin is \$117.5/mt. In the stochastic optimization model, the probability of the conventional ammonia price being within each bin is listed above that bin.

The optimization problem has 217,566 (234 binary and 217332 continuous) variables, 11,927 constraints, and the solution time is 78 seconds. The total net present cost is \$3,230MM and the renewable sites installed are presented in Table 2. As with the optimal solution of the deterministic model, eight new renewable production facilities are installed. However, these installations occur earlier in the optimal stochastic supply chain transition.

One new facility each is installed in 2024 and 2025, which allows some market penetration of renewable ammonia by 2026 (see Figure 3). Both of these facilities are

installed in locations with the highest wind potential and use the maximum allowable 250 MW of co-located wind generation capacity. In 2027, three additional facilities are added and these cumulatively use 575 MW of electrolysis, the upper bound for that year. The facilities in both Luverne and Worthington use 250 MW of wind generation. We point out that these facilities have slightly higher production capacity than those in Lake Wilson (2024) and Chandler (2025) due to the more efficient electrolysis expected to be available in future years. The third 2027 facility is located in Southeast Minnesota despite its lower wind potential, unlike in the deterministic supply chain. This enables another facility which uses 250 MW wind to be installed in Wilmont in 2028; all five facilities with the highest wind potential (52% capacity factor) use this maximum amount of wind capacity to achieve economies of scale. Another facility is installed in Southeast Minnesota in 2028, also at the 2050 MW scale. Finally, a smaller facility is installed in Southeast Minnesota in 2030 to ensure a fully renewable supply chain by 2032.

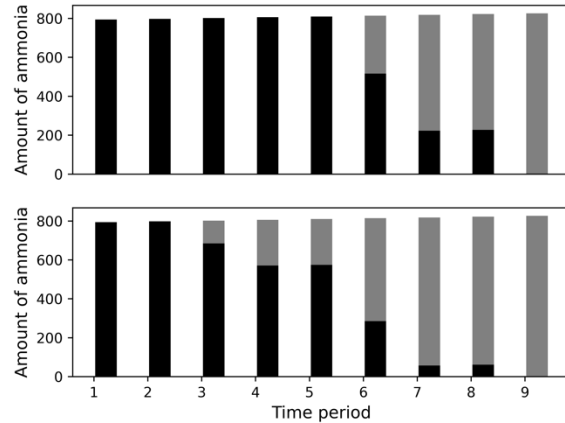
**Table 2:** Installation year, location, and capacity for new renewable ammonia production in the two-stage stochastic case. Annual average wind capacity factors for each selected candidate location are provided in parentheses to describe wind potential.

Year	Location	Capacity (1,000 mt/y)
2024	Lake Wilson (52%)	117.7
2025	Chandler (52%)	117.7
2027	Luverne (52%)	121.14
2027	Worthington (52%)	121.14
2027	Winnebago (47%)	50.71
2028	Wilmont (52%)	121.24
2028	Blue Earth (47%)	111.11
2030	Fairmont (47%)	65.34

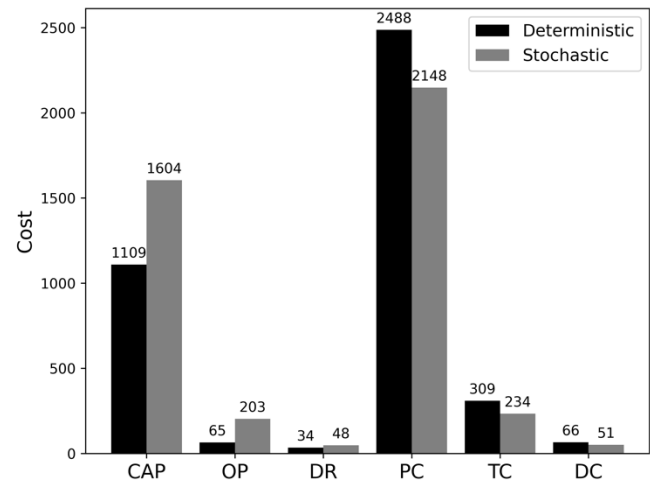
### Comparison between deterministic and stochastic cases

The optimal transition in the stochastic case results in less conventional ammonia purchases and more in-state renewable production over the planning horizon compared to the deterministic transition. This difference leads to a reduction in the total amount of ammonia purchased leading to lower cumulative purchase, transportation, and distribution costs for conventional ammonia over the planning horizon (see Figure 4). Conversely, the capital cost for renewable ammonia production is higher in the stochastic transition even though the same total capacity of renewable production is installed in both cases. This is due to earlier installation using more expensive constituent technologies in the stochastic case.

This also contributes to higher renewable ammonia operating costs over the planning horizon, though these are also higher simply because more renewable ammonia is being produced. Overall, the net present cost of the supply chain transition is 7.6% (228 MM\$) higher in the stochastic case than the deterministic case.



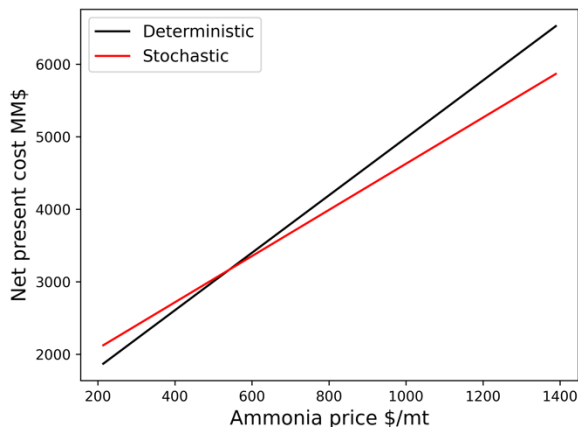
**Figure 3.** Amount of ammonia (mt) in each time period from conventional purchases (black bar) and in-state renewable ammonia production (gray bar) for the deterministic (top figure) and stochastic (bottom figure) cases.



**Figure 4.** Cost contributions to optimal net present cost for the deterministic and stochastic cases. The cost acronyms are defined as follows: CAP - Renewable capital, OP - Renewable operating, DR - Renewable distribution, PC - Conventional purchase, TC - Conventional transportation to Minnesota, DC - Conventional distribution.

We compare the supply chain configurations obtained via the deterministic and stochastic transition optimizations for different conventional ammonia prices to

elucidate the benefit of the stochastic approach. Specifically, we discretize the ammonia price uniformly in 100 points between \$214/mt and \$1389/mt, and for each price, we fix the investment decisions (timing of investments, i.e., binary variables and installed capacity) and compute the net present cost of supply chain transition (see Figure 5). We observe that for high prices of conventional ammonia (above \$700/mt), the net present cost of the supply chain obtained via stochastic programming is meaningfully lower than the cost of the deterministic design, whereas if the price is low (below \$400/mt) the design obtained by the deterministic model has a meaningfully lower net present cost. This difference can be attributed to the different investment strategies for the stochastic and deterministic cases (i.e., more ammonia is manufactured in-state for the stochastic case) as discussed in the previous paragraph.



**Figure 5.** Levelized cost for the deterministic and stochastic design as a function of ammonia price.

## CONCLUSIONS

In this work, we focused on quantifying the effect of uncertainty in conventional ammonia prices on Minnesota's transition from importing fossil-derived ammonia from out-of-state conventional producers to in-state renewable ammonia production. We used a two-stage stochastic programming approach to determine the optimal investment profile for in-state renewable ammonia production over a fixed planning horizon such that the demand of ammonia is satisfied while the net present cost is minimized. The results show that when accounting for uncertainty in the conventional ammonia price, investments are made earlier in the planning horizon compared to a deterministic supply chain transition model. This leads to a reduction in the amount of ammonia purchased from conventional producers. This can bring significant cost savings for higher-than-average ammonia prices. Overall, the stable production cost afforded by in-state

renewable ammonia production can act as a hedge against conventional ammonia price uncertainty and the possibility of very high prices on the conventional ammonia market.

This work used a two-stage approach to account for conventional price uncertainty, but in practice these prices could evolve different multi-year trajectories. Future work will therefore develop multi-stage stochastic programming models for the supply chain transition problem. Furthermore, technology cost reductions are also subject to uncertainty, especially further into a given planning horizon. For example, electrolysis costs are expected to decrease, but the magnitude of this reduction is not well-established at present. Thus, these types of uncertainties will also be incorporated into future supply chain transition models.

## ACKNOWLEDGMENTS

The work was funded in part by NSF CBET (award number 2313289) and the Advanced Research Projects Agency-Energy (ARPA-E), U.S. Department of Energy, under Award Number DE-AR0001479. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## REFERENCES

1. Smith, C., Hill, A.K., Torrente-Murciano, L. Current and future role of Haber-Bosch ammonia in a carbon-free energy landscape. *Energy Environ. Sci.* 13: 331-344 (2020).
2. Rouwenhurst, K.H.R., Travis, A.S., & Lefferts, L. 1921-2021: A century of renewable ammonia synthesis. *Sustain. Chem.* 3(2): 149-171 (2022).
3. The Royal Society. *Ammonia: Zero-carbon fertiliser, fuel, and energy store* (2020). [royalsociety.org/-/media/policy/projects/green-ammonia/green-ammonia-policy-briefing.pdf](https://royalsociety.org/-/media/policy/projects/green-ammonia/green-ammonia-policy-briefing.pdf)
4. Palys, M.J. & Daoutidis, P. Power-to-X: A review and perspective. *Comput. Chem. Eng.* 165: 107948 (2022).
5. U.S. Environmental Protection Agency. *Commercial Fertilizer Purchased* (2023). [epa.gov/nutrient-policy-data/commercial-fertilizer-purchased](https://epa.gov/nutrient-policy-data/commercial-fertilizer-purchased)
6. Allman, A., Daoutidis, P., Tiffany, D., & Kelley, S. A framework for ammonia supply chain optimization incorporating conventional and renewable generation. *AIChE J* 63(10): 4390-4402 (2017).
7. Ahmed, S. Strategic planning under uncertainty: Stochastic integer programming approaches. *University of Illinois at Urbana-Champaign* (2000)
8. Palys, M.J. & Daoutidis, P. Optimizing renewable ammonia production for a sustainable fertilizer

- supply chain transition. *ChemSusChem* 16(22): e202300563 (2023).
9. The World Bank. *Commodity Markets* (2023) [worldbank.org/en/research/commodity-markets](https://worldbank.org/en/research/commodity-markets)
  10. Grossmann, I.E., Apap, R.M., Calfa, B.A., García-Herreros, P. & Zhang, Q. Recent advances in mathematical programming techniques for the optimization of process systems under uncertainty. *Comput. Chem. Eng.* 91: 3-14 (2016).
  11. Govindan, K., Fattahi, M. and Keyvanshokoh, E. Supply chain network design under uncertainty: A comprehensive review and future research directions. *Eur. J. Oper. Res.* 263(1): 108-141 (2017)
  12. Chen, C.W., & Fan, Y. Bioethanol supply chain system planning under supply and demand uncertainties. *Transport. Res. E-LOG* 48(1), 150-164 (2012).
  13. Marufuzzaman, M., Eksioğlu, S.D., & Huang, Y.E. Two-stage stochastic programming supply chain model for biodiesel production via wastewater treatment. *Comput. Oper. Res.* 49: 1-17 (2014).
  14. Zhou, X., Zhang, H., Qiu, R., Lv, M., Xiang, C., Long, Y., & Liang, Y. A two-stage stochastic programming model for the optimal planning of a coal-to-liquids supply chain under demand uncertainty. *J. Clean. Prod.* 228: 10-28 (2019).
  15. 17th U. S. Congress, *Inflation Reduction Act of 2022, H.R.5376* (2022) [congress.gov/bill/117thcongress/house-bill/5376](https://congress.gov/bill/117thcongress/house-bill/5376)
  16. Hart, W.E., Watson, J.P. & Woodruff, D.L. Pyomo: modeling and solving mathematical programs in Python. *Math. Program. Comput.* 3: 219-260 (2011).
  17. Gurobi Optimization LLC, *Gurobi Optimizer Reference Manual* (2023) [gurobi.com](https://gurobi.com)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# A Fast Computational Framework for the Design of Solvent-Based Plastic Recycling Processes

Aurora del C. Munguía-López, Panzheng Zhou, Ugochukwu M. Ikegwu, Reid C. Van Lehn, and Victor M. Zavala\*

University of Wisconsin-Madison, Department of Chemical and Biological Engineering, Madison, WI, United States of America

\* Corresponding Author: [victor.zavala@wisc.edu](mailto:victor.zavala@wisc.edu)

## ABSTRACT

Multilayer plastic films are widely used in packaging applications because of their unique properties. These materials combine several layers of different polymers to protect food and pharmaceuticals from external factors such as oxygen, water, temperature, and light. Unfortunately, this design complexity also hinders the use of traditional recycling methods, such as mechanical recycling. Solvent-based separation processes are a promising alternative to recover high-quality pure polymers from multilayer film waste. One such process is the Solvent-Targeted Recovery and Precipitation (STRAP™) process, which uses sequential solvent washes to selectively dissolve and separate the constituent components of multilayer films. The STRAP™ process design (separation sequence, solvents, operating conditions) changes significantly depending on the design of the multilayer film (the number of layers and types of polymers). Quantifying the economic and environmental benefits of alternative process designs is essential to provide insights into sustainable recycling and film (product) design. In this work, we present a fast computational framework that integrates molecular-scale models, process modeling, techno-economic and life-cycle analysis to evaluate STRAP™ designs. The computational framework is general and can be used for complex multilayer films or multicomponent plastic waste streams. We apply the proposed framework to a multilayer film commonly used in industrial food packaging. We identify process design configurations with the lowest economic and environmental impact. Our analysis reveals trends that can help guide process and product design.

**Keywords:** Polymers, Process Design, Modelling and Simulations, Technoeconomic Analysis, Life Cycle Analysis

## INTRODUCTION

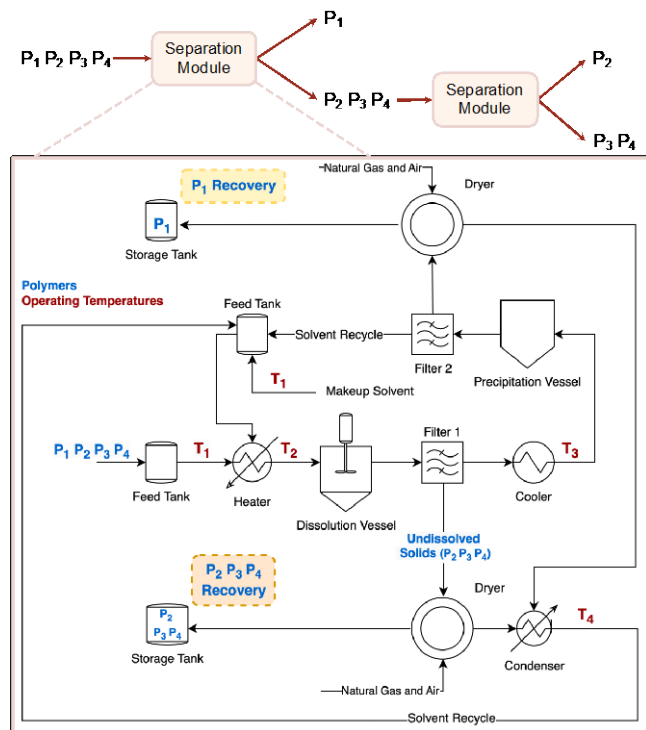
A wide range of packaging applications use multilayer plastic films to protect products from external factors (e.g., oxygen, water, temperature, and light). Multilayer films are complex and diverse because they combine several layers of distinct polymers to leverage their unique properties. The multilayer film design can vary for different applications; for instance, industrial films can include more than ten polymer layers. This complexity hinders direct mechanical recycling. Solvent-based approaches have emerged as a promising alternative to recover and recycle their constituent polymers. One such technology is the Solvent-Targeted Recovery and Precipitation (STRAP™) process, which uses sequential

solvent washes to selectively dissolve and separate the constituent components of multilayer plastic films [1]. The solvents used need to be properly selected to dissolve only a target component in each step. Thus, each selected solvent must have a high solubility for the target component and a low solubility for the other components of the multilayer film. This target component can refer to one or multiple polymers [2].

Molecular-scale models have been employed to predict polymer solubilities for solvent selections in such recycling processes. We recently introduced a joint computational and experimental workflow that can perform large-scale temperature-dependent polymer solubility predictions. Based on this approach, we generated a solubility database for common polymers and a large

number of solvents [3]. The database has enabled the creation of a computational tool that automates solvent selections for solvent-based processes. This tool can screen and rank solvent candidates based on the solubility requirements of separation sequences. Once the solvent for the target component is selected, the next steps of the STRAP™ process can be taken as described below.

**Figure 1** provides an example of the modules and process flow diagram that we use to represent the separation of different components using the STRAP™ process. For instance, for a multilayer film of 4 components, the first component can be separated as shown in the process flow diagram. In this example, each component comprises one polymer. The process steps are as follows: first, the multilayer film is mixed with a previously chosen solvent (that selectively dissolves only the target polymer  $P_1$ ) and heated; next, the solution is filtered to separate the undissolved solids (polymers  $P_2, P_3, P_4$ ); finally, the target polymer  $P_1$  is precipitated via temperature reduction, filtered, and recovered. Most of the solvent is also recovered and recycled. Then, the process is repeated to separate the remaining components (polymers  $P_2, P_3, P_4$ ). Hence, the separation of an  $n$ -component multilayer film requires  $n-1$  separation stage(s).



**Figure 1.** Simplified flow diagram of the STRAP™ process for one separation module.

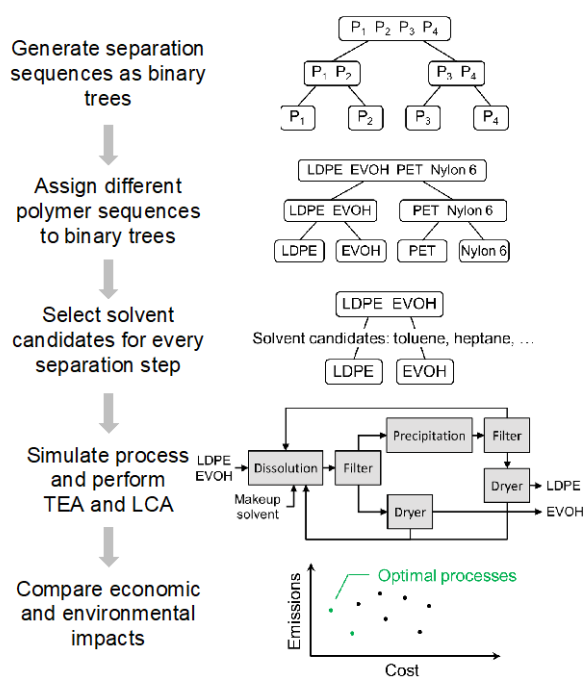
A couple of precipitation techniques have been previously reported [2]. One is temperature-driven

precipitation, which is economically and environmentally feasible and can be used for several polymer-solvent combinations. If the precipitation cannot be done via temperature, an antisolvent can be added to enable precipitation. However, adding an antisolvent leads to higher energy requirements because a distillation unit is required to separate the resulting mixture of solvents. Previous research has shown that the economic and environmental performance of the process is better when the precipitation is temperature-driven [4]. Therefore, in this work, we will select solvents that can enable temperature-driven precipitation only. It is worth highlighting that the solvent and operating temperatures can change if the separation sequence changes (for example, with  $P_2$  as the target polymer). These changes can impact the economic and environmental benefits of the process. Additionally, identifying the proper conditions to achieve a high solvent recovery rate is key for the process benefits. Previous techno-economic analysis (TEA) and life cycle assessment (LCA) of the different STRAP™ process variations have also shown significant differences in costs and  $\text{CO}_2$  emissions due to solvent selection and separation sequence [5,6]. While these models have been useful to determine the economic and environmental feasibility as well as process design bottlenecks, they have also shown the need for a general framework to guide the design of solvent-based processes.

In this work, we present a fast computational framework that integrates molecular-scale models, process modeling, TEA and LCA to provide insights into sustainable solvent-based process design. We also aim to provide guidelines for multilayer film designs that are easier to recycle or have a lower recycling impact. This framework can determine the economic and environmental impacts of several process design scenarios, including different separation sequences, solvents that enable temperature-driven precipitation, and process operating conditions (for dissolution, precipitation, and solvent recovery). The computational framework is general and can be used for complex multilayer films or multicomponent plastic waste streams.

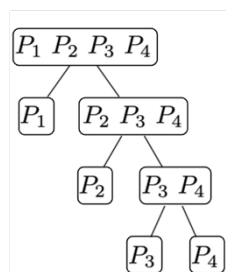
## COMPUTATIONAL FRAMEWORK

Our computational framework follows a series of steps to determine the economic and environmental impacts of each process design feasible scenario. The computational steps are summarized in **Figure 2**.



**Figure 2.** Schematic representation of the proposed framework.

First, for any given multilayer film, the framework generates all possible separation sequences. **Figure 3** provides an example of the binary trees that we use to represent the separation sequences for a multilayer film composed of 4 polymers. Here, we use the left branch to denote the dissolved polymer(s) and the right branch to refer to the undissolved polymer(s). For this example, we can see that every separation step only dissolves one polymer. Therefore, this example has a  $D^n = 1$ , which represents the number of polymers being dissolved in each step. Also, this binary tree only shows such separation sequence. Changing the separation order of the sequence provides all possibilities for the case of  $D^n = 1$ . Therefore, the permutation of the set of polymers  $\{P_1, P_2, P_3, P_4\}$  results in 24 different sequences.



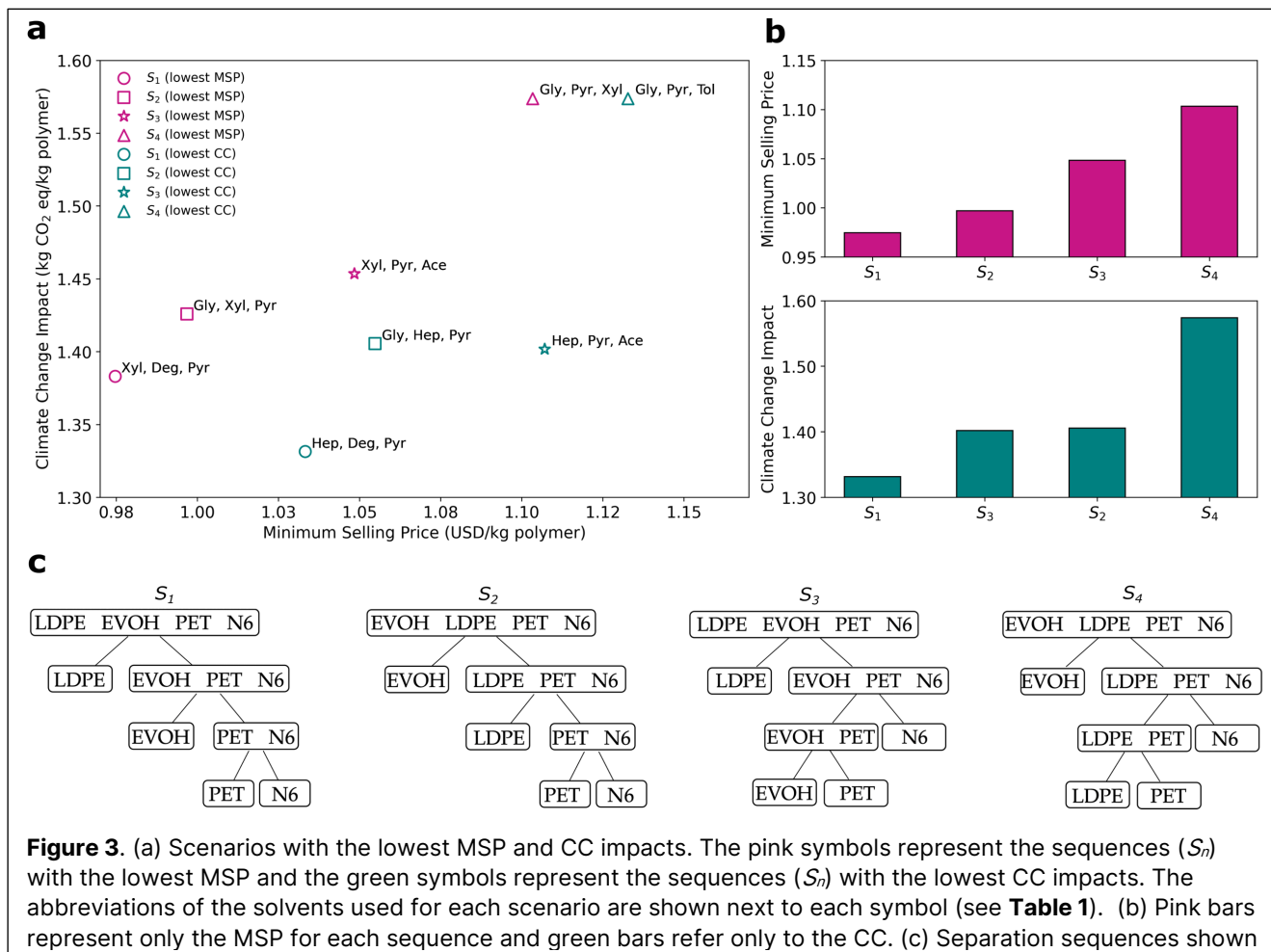
**Figure 3.** Binary tree illustrating a possible separation sequence for a multilayer film composed of 4 polymers.

To determine all the possible sequences, we also consider the binary trees where 2 polymers are dissolved in a certain step ( $D^n = 2$ ), resulting in 2 different binary trees and 48 sequences. Similarly, when there are 3 polymers dissolved together, 2 different binary trees and 48 sequences are possible. Therefore, the separation of a 4-polymer mixture has in total 5 possible binary tree structures, which correspond to 120 sequences. **Table 1** presents a summary of the number of sequences and binary trees generated based on the value of  $D^n$  for the 4-polymer mixture. It is worth highlighting that this combinatorial complexity increases rapidly with the number of polymers. For example, a 5-polymer mixture will have 14 different binary tree structures, which leads to 1680 possible sequences.

**Table 1:** Possible binary trees and sequences for a 4-polymer mixture considering the different possible numbers of dissolved polymers in a certain step ( $D^n$ ).

$D^n$	Binary Trees	Sequences
1	1	24
2	2	48
3	2	48

As shown in **Figure 2**, the next step is to assign different polymer sequences to the binary trees and eliminate the unfeasible sequences based on empirical rules from previous research (e.g., composition, polymer solubility, and maximum number of polymers dissolved in each step) [3-6]. After this, we use the computational tool to identify all the solvent candidates for the target polymer(s) in each separation step. This tool selects solvents (only based on solubility) from the previously reported database that includes predicted temperature-dependent polymer solubilities from molecular-scale models [3]. The criteria to select the feasible solvents is that the solubility of the target polymer(s) should be greater than 15 wt% and the solubility of the non-target polymer(s) should be lower than 3 wt%. Once the solvents have been identified, we estimate the required temperature of the condenser to achieve a high solvent recovery rate (~99.90%) given the inlet temperature from the dryer. Previous studies indicate that achieving a solvent recovery of ~99.99% enhances the economic viability of the separation process. However, solvents with low boiling points, such as toluene, pyridine, and acetic acid, face challenges in attaining this recovery rate. In contrast, solvents like glycol, dodecane, and diethylene glycol can achieve this target. It is equally essential to avoid solvents with high boiling points, as they can escalate the energy consumption in the distillation process during solvent mixture recovery (which occurs in precipitation via antisolvent addition). Therefore, the solvent recovery is



**Figure 3.** (a) Scenarios with the lowest MSP and CC impacts. The pink symbols represent the sequences ( $S_n$ ) with the lowest MSP and the green symbols represent the sequences ( $S_n$ ) with the lowest CC impacts. The abbreviations of the solvents used for each scenario are shown next to each symbol (see **Table 1**). (b) Pink bars represent only the MSP for each sequence and green bars refer only to the CC. (c) Separation sequences shown

set at ~99.90%, and future considerations will include evaluating the trade-offs associated with each solvent. The condenser temperature is estimated by simulating a heat exchanger and iterating over different temperatures to achieve the recovery rate specification as a pre-process simulation. The process model simulation for the heat exchanger is performed in the open-source platform BioSTEAM [7]. This Python process simulator has been validated against proprietary software (SuperPro Designer and Aspen Plus).

Next, we use the collected inputs (mass, polymers, composition, separation sequence, solvent, and operating conditions) to simulate the STRAP™ process and perform the TEA in BioSTEAM. We use the minimum selling price (MSP) as the economic metric to compare different scenarios. After this, we evaluate the environmental impact of each scenario using an LCA methodology. The LCA was performed using the open-source software openLCA [8], the Environmental Footprint and AGRIBALYSE databases and the Environmental footprint impact assessment method [9,10]. We consider all the inputs to the process (electricity, steam, water) and we estimate the climate change (CC) impact for each scenario. Since there are limited data for solvents in the LCA

databases and we are considering high recycling rates, we do not include the impact of solvents for this analysis. Finally, we store the economic (MSP, expressed in USD per kg of polymer sold) and environmental (CC impact, expressed in kg CO<sub>2</sub> per kg of polymer) outputs of all feasible process design scenarios (including different sequences and solvents).

## RESULTS AND DISCUSSION

We applied the proposed computational framework to a multilayer film composed of LDPE, EVOH, PET, and Nylon 6 (N6). This 4-polymer multilayer film is commonly used for food packaging applications [11]. As mentioned above, there are 120 possible sequences for a 4-polymer mixture.

To identify the sequences and solvents that lead to the best economic and environmental performance, we compare the MSP and CC impact of all the generated scenarios. We consider all the separation modules required to recover the 4 constituent polymers of the multilayer film using the STRAP™ process. As a benchmark, we also estimate the climate change impact of producing the 4 polymers from fossil sources and their average

market prices.

First, we identify the solvents that lead to the lowest MSP in each separation step of the different generated sequences and report the corresponding CC impact of these sequence-solvent combinations. Similarly, we identify the solvents that lead to the lowest CC impact and their corresponding MSP. **Figure 3a** presents the scenarios with the lowest MSP (pink symbols) and the lowest CC impact (green symbols). It should be noticed that each symbol refers to a different sequence (represented by  $S_i$ ) and set of solvents used for all the required separation modules. Solvents are represented by an abbreviation which is related to the solvent common name in **Table 2**. From **Figure 3a**, we can see that there are trade-offs between the economic and environmental metrics. For instance, the sequence  $S_7$  with the set of solvents p-Xylene, Diethylene glycol and Pyridine (represented by the pink circle) leads to the lowest MSP (0.97 USD/kg), but it does not lead to the lowest CC impact. On the other hand, the sequence  $S_7$  with the set of solvents n-heptane, Diethylene glycol and Pyridine (represented by the green circle) leads to the lowest CC impact (1.33 kg CO<sub>2</sub>/kg) but leads to a higher MSP.

We compare these results with previous STRAP<sup>TM</sup> studies to show the viability of the proposed framework. A TEA [2] and an LCA [4] for a similar multilayer film composed of LDPE, EVOH, PET, and EVA are used as references. Here, the sequence used was similar to sequence  $S_7$  and was demonstrated experimentally. We find that the reported MSP is 1.1 USD/kg and the CC is 1.18 kg CO<sub>2</sub>/kg. Although the polymers, composition, and solvents used are not equal to the multilayer film addressed in this study, we can see that the economic and environmental impacts are similar which highlights the applicability of the proposed framework. Future work will include evaluating the previously reported multilayer film (LDPE, EVOH, PET, and EVA) with our framework to identify other feasible sequences and solvents that could be tested experimentally. **Figure 3b** provides a summary of the MSP (pink bars) and CC impact (green bars) for each sequence. To compare these values to the price and climate change impact of virgin polymers we consider the following. As a reference, we find that the average market prices of the recovered polymers are 1.2–2.6 USD/kg. Furthermore, the multilayer film addressed in this study has a market price of 2.9 USD/kg [11, 12]. From previous research, we know the range of the required scale to achieve an MSP comparable to these market prices [5]. Therefore, the results presented in this work were obtained for a fixed plant capacity of 6,400 tons per year. As we can see in **Figure 3b**, the MSP of the recovered polymers is comparable to the average market prices. However, this trend is sensitive to the scale of the process and can change with other capacity factors.

**Table 2:** Common name of selected solvents for the scenarios with the lowest impacts.

Solvent Abbreviation	Solvent Common Name
Ace	Acetic acid
Deg	Diethylene glycol
Gly	Glycol
Hep	n-heptane
Xyl	p-Xylene
Pyr	Pyridine
Tol	Toluene

Regarding the environmental impact of virgin polymers, we estimate the CC impact of the production from fossil sources of the constituent polymers of the multilayer film considered in this work (LDPE, EVOH, PET, and N6). Considering that every polymer represents 25 wt% of the film, we find that 4.81 kg CO<sub>2</sub> per kg of polymer are generated. From **Figure 3b**, we can see that all the different STRAP<sup>TM</sup> sequences have a lower CC impact. Specifically, sequence  $S_7$  results in around 60% fewer emissions than the production of the multilayer film from fossil sources.

**Figure 3c** presents the different sequences as binary trees showing the dissolved polymers in each separation step (left branch). From the MSP and CC plots, we observe that sequence  $S_7$  results in the lowest economic and environmental impacts. This sequence refers to dissolving one polymer in each separation step in the following order: LDPE, EVOH, PET, and N6. However, the set of solvents that lead to the lowest MSP and CC are different for each sequence. There are different factors that determine if one sequence and solvent is better for the economic or environmental metrics, such as the amount of required solvent (which is determined by the polymer solubility) and the thermodynamic properties of the solvents. This first analysis provides initial suggestions of the key variables that impact the economic and environmental performance. For instance, the dissolution temperature is dependent on the boiling point of solvents. Therefore, solvents with low boiling points result in lower energy requirements and environmental impacts. On the other hand, we found that the solvents that require a lower polymer-solvent ratio (due to higher solubility) were selected for the scenarios with the best economic performance. Such solvents had a higher boiling point.

## CONCLUSIONS AND FUTURE WORK

This work presented a fast computational framework to provide insights into the STRAP<sup>TM</sup> process design. The proposed framework helps identify the separation sequence, solvents, and process operating



conditions with the lowest economic and environmental impact. We use the framework to identify process designs with low impacts for a multilayer film composed of 4 polymers. From these results, we conclude that the environmental performance is mainly driven by the boiling point of solvents while the economic performance is guided by the polymer solubility. Our framework is general and can help guide the design of other solvent-based processes to efficiently treat complex multilayer waste streams. Therefore, future work will include analyzing case studies with different polymers and number of layers. Furthermore, we will consider varying recycling rates for the solvents. We will also evaluate other processing plant scales and the recovery of only selected polymers (for streams with large number of polymers).

## ACKNOWLEDGEMENTS

This material is based upon work supported by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Bioenergy Technologies Office under Award Number DEEE0009285.

## REFERENCES

- Walker TW, Frelka N, Shen Z, Chew AK, Banick J, Grey S, Kim MS, Dumesic JA, Van Lehn RC, Huber GW. "Recycling of multilayer plastic packaging materials by solvent-targeted recovery and precipitation." *Sci. Adv.* 6, no. 47: eaba7599 (2020)
- Sánchez-Rivera KL, Zhou P, Kim MS, González Chávez LD, Grey S, Nelson K, Wang SC et al. "Reducing antisolvent use in the STRAP process by enabling a temperature-controlled polymer dissolution and precipitation for the recycling of multilayer plastic films." *ChemSusChem* 14, no. 19 : 4317-4329 (2021)
- Zhou P, Yu J, Sánchez-Rivera KL, Huber GW, Van Lehn RC. "Large-scale computational polymer solubility predictions and applications to dissolution-based plastic recycling." *Green Chem.* 25, no. 11: 4402-4414 (2023)
- Munguía-Lopez AC, Göreke D, Sánchez-Rivera KL, Aguirre-Villegas HA, Avraamidou S, Huber GW, Zavala VM. "Quantifying the environmental benefits of a solvent-based separation process for multilayer plastic films." *Green Chem.* 25(4): 1611-1625. (2023)
- Sánchez-Rivera KL, Munguía-López AC, Zhou P, Cecon VS, Yu J, Nelson K, Miller D et al. "Recycling of a post-industrial printed multilayer plastic film containing polyurethane inks by solvent-targeted recovery and precipitation." *Resour. Conserv. Recycl.* 197: 107086 (2023)
- Jiuling Y, Munguía-López AC, Cecon VS, Sánchez-Rivera KL, Nelson K, Wu J, Kolapkar S et al. "High-purity polypropylene from disposable face masks via solvent-targeted recovery and precipitation." *Green Chem.* 25, no. 12: 4723-4734 (2023)
- Cortes-Pena Y, Kumar D, Singh V, Guest, JS "BioSTEAM: a fast and flexible platform for the design, simulation, and techno-economic analysis of biorefineries under uncertainty." *ACS Sustain. Chem. Eng.* 8, no. 8: 3302-3310 (2020)
- Ciroth A, Di Noi C, Lohse T, Srocka M. "openLCA 1.10 Comprehensive user manual." GreenDelta, Berlin, Germany (2020)
- Fazio S, Castellani V, Sala S, Schau E, Secchi M, Zampori L, Diaconu E. "Supporting information to the characterisation factors of recommended EF Life Cycle Impact Assessment methods: new methods and differences with ILCD." (2018).
- Colomb V, Ait-Amar S, Basset-Mens C, Gac A, Gaillard G, Koch P, Mousset J, Salou T, Tailleur A, Van Der Werf HM. "AGRIBALYSE®, the French LCI Database for agricultural products: high quality data for producers and environmental labelling." (2015).
- Shandong Top Leader Plastic Packing CO. <https://www.sdzplastic.com/pa-pe/62713046.html>
- POLYMERSCAN, S. & P Global. <https://www.spglobal.com/commodityinsights/en/products-services/petrochemicals/polymerSCAN>

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Optimal Design of Food Packaging Considering Waste Management Technologies to Achieve Circular Economy

Paola A. Munoz-Briones<sup>a</sup>, Aurora del C. Munguía-López<sup>a</sup>, Kevin L. Sánchez-Rivera<sup>a</sup>, Victor M. Zavala<sup>a</sup>, George W. Huber<sup>a</sup> and Styliani Avraamidou<sup>a\*</sup>

<sup>a</sup> University of Wisconsin-Madison, Department of Chemical and Biological Engineering, Madison, Wisconsin, USA

\* Corresponding Author: avraamidou@wisc.edu.

## ABSTRACT

Plastic packaging plays a fundamental role in the food industry, avoiding food waste and facilitating food access. The increasing plastic production and the lack of appropriate plastic waste management technologies represent a threat to the environmental and human welfare. Therefore, there is an urgent need to identify sustainable packaging solutions. Circular economy (CE) promotes reducing waste and increasing recycling practices to achieve sustainability. In this work, we propose a CE framework based on multi-objective optimization, considering both economic and environmental impacts, to identify optimal packaging designs and waste management technologies. Using mixed-integer linear programming (MILP), techno-economic analysis (TEA), and life cycle assessment (LCA), this work aims to build the first steps in packaging design, informing about the best packaging alternatives and the optimal technology or technologies to process packaging waste. For the economic analysis, we consider the minimum increase in price (MIP) when adding recycling to the cost of each packaging solution, while for the environmental analysis, the greenhouse gas emissions impact was considered. A case study on ground coffee packaging is used to illustrate the proposed framework. The results demonstrate that the multilayer bag option is the most convenient when considering both the chosen economic and environmental impacts.

**Keywords:** Optimization, Life Cycle Analysis, Technoeconomic Analysis, Supply Chain, Modelling

## INTRODUCTION

Approximately 35-40% of plastics produced are discarded as waste after their first use, this percentage is expected to rise in the coming years based on current projections [1]. Containers and packaging represent around 28% of the municipal solid waste. Plastic containers account for around the 18%, and more than half of them are disposed in landfills [2]. The growing generation of plastic waste and the lack of initiatives to upcycle it, along with the irresponsible disposal of waste, threatens both marine life and humans [3].

The food industry is one of the largest manufacturing sectors contributing to the global economy. Packaging plays a crucial role in the food industry considering factors such as food quality and preservation, marketing appeal, and proper product identification. The final disposal or management of the packaging at the end of its

life is a critical factor that should be incorporated during the packaging design stage, given the environmental concerns around plastic pollution. Over the past years, food packaging has evolved to increase efficiency (extend shelf-life while maintaining the freshness of the product) increasing food accessibility and reducing food waste. However, plastic food packaging is considered to follow a linear economic model, where it is designed and produced just in accordance with its intended use [4]. The need for sustainable food packaging solutions has become increasingly significant as the proportion of food packaging waste in municipal waste is on the rise.

Moreover, the increasing production of plastic implies a growing stress on resources. Materials extraction along with the processing of materials, fuels, and food contribute to water stress, biodiversity loss, and accounts for around 50% of the greenhouse gas (GHG) emissions [5]. The growing production of plastics not only

implicates using more virgin materials, but also the use of more resources, including energy and water, for the manufacturing of products. To tackle this, circular economy emerges as a system to promote the transition towards achieving environmental, economic, and social sustainability [6]. CE endorses, among other objectives, the closed-loop handling of natural resources to achieve minimal waste and to maintain the value of materials and products to their highest quality for as long as possible. Zhu et al notes that the design of packaging has been acknowledged as the essential initial step in moving toward CE [4]. Transitioning to more sustainable packaging solutions requires analyzing the impacts of alternative materials and their corresponding waste management technologies. Reducing the waste generated and increasing the compatibility of materials with recycling processes are fundamental to achieving recyclability targets [7].

Often, recycling complex plastic containing products can be more resource intensive than using the virgin plastic resins. For instance, PET possess a great resistance to high temperature which involves a considerable amount of energy to recycle it. Another example is multi-layer plastic films that require fewer resources given the efficient combination of materials. However, this configuration affects its recyclability. Therefore, the waste management processes for packaging should be taken into consideration during the design of the packaging.

To this end, we present a circular economy system engineering framework and decision-making tool that examines alternative food packaging design options and waste management possibilities. Using a superstructure approach, the present work employs a mixed-integer linear programming model for determining the optimal combination of packaging and recycling technologies based on economic and environmental criteria. As a case study, we have chosen coffee, which is the second largest traded commodity.

## LITERATURE REVIEW

The state-of-the-art of sustainable packaging design consists of articles focusing on material selection, conceptual design, design development, and design validation. The material selection work focuses on reusability, biodegradable materials, and processability. The conceptual design phase involves topics such as the integration of reusing, and the abstention of a variety of impactful materials. The design development can incorporate the definition of appropriate features like size and weight, as well as the use of labels. Regarding the design validation phase, LCA tools and CE indicators are commonly used [4]. Although, the design guidelines should consider the facilitation of packaging waste management [8].

Karayilan et al [9] proposes a single-objective optimization model to evaluate environmental benefits (climate change and marine ecotoxicity) and material circularity within the packaging value chain. However, other circularity metrics and the economic aspect are not evaluated. Other studies, [10], focus on assessing the performance of different waste management systems for plastic waste, without considering the design of packaging.

Life Cycle Assessment (LCA) studies have been conducted to analyze the environmental performance of food packaging. Toniolo et al [11] performed a comparative LCA to estimate the appropriate recycling percentage for food packaging. However, most studies typically focus on evaluating metrics for the production and use of different packaging materials or designs. Siracusa et al [12] performed an environmental assessment for a single type of multilayer food packaging with a cradle to factory-gate scope. On the other hand, other LCA and techno-economic assessments evaluate various packaging options, but they concentrate on a specific recycling technology. Xie et al [13] reported an LCA on the recycling of Al-PE composite packaging. Some focus on comparing one or few alternative packaging solutions to the conventional, for example, reusable containers with single-use containers [14]. Jeswani et al [3] presented an LCA for mixed plastic waste considering pyrolysis, mechanical recycling, and energy recovery. Similarly, Munguia-Lopez et al [15] studied the environmental benefits of using solvent-based processes to recycle multilayer plastic films.

Optimization studies can enable the identification of optimal solutions for scenarios considering CE principles. Systems engineering frameworks involving multi-objective modeling can be used for the optimization of supply chains or technology allocation. Komly et al [16] presents a multi-objective optimization strategy to find the optimal allocation of PET bottles and recycling technologies to minimize environmental impacts.

However, a comprehensive economic and environmental assessment and/or optimization study considering the production, use, and waste management for different packaging designs is still lacking.

## METHODOLOGY

The methodology followed will be described based on the case study for coffee packaging. Ground coffee currently comes in several packaging types, including disposable and non-disposable alternatives (refilling and reusing options). These options encompass monolayer film bags, containers, multi-layer film bags, and metal cans, that utilize materials ranging from polyethylene and/or aluminum to polyethylene terephthalate and polylactic acid (PLA). Waste management processes also offer various options, spanning traditional landfilling,

incineration, and mechanical recycling, to more recent technologies like pyrolysis, and solvent-targeted recovery and precipitation (STRAP). Previous research has used LCA and multi-objective optimization to achieve a sustainable supply chain for coffee analyzing environmental and economic impacts but considering coffee production and distribution, but not its packaging [17-18]. The proposed methodology focusing on the optimal design and waste management of ground coffee packaging is divided into the four steps described below.

### Packaging option identification

The first step in this analysis, is the identification of coffee packaging options currently available in the market. Between the packaging alternatives, various materials are utilized to consolidate different presentations of packaging. Among the bags considered, there are two plastic film options: monolayer and multilayer films. A multilayer film is a flexible plastic packaging that is composed of different materials that are combined to achieve specific properties that cannot be achieved with monolayer films alone. For example, in the case of coffee, a common multilayer film is composed of PE, PET, and aluminum. By combining the specific attributes of these materials, a packaging with high barrier properties and protection from light and oxygen is achieved, therefore, high-quality coffee is guaranteed. However, multilayer films are difficult to recycle using certain technologies such as mechanical recycling because of the incompatibility of the materials. Before multilayer films arose, monolayer films dominated the market, and some options are still present today.

The use of renewable resources to produce plastics arises as an alternative to avoid the dependence on fossil fuels (non-renewable energy sources) and address environmental concerns. Common plant-based feedstocks like sugarcane and corn can be processed into polymers that would have the exact same structure as the fossil fuels-derived ones. The production of these biobased materials is expected to result in a reduced carbon footprint owing to the carbon sequestration occurred during the cultivation of the plants. However, due to the various estimation methods and the conflicting results based on considerations like land use, the present study considers the carbon uptake outside of the scope.

Therefore, four options for ground coffee packaging were considered in this study: i) multi-layer film bags made with PE, aluminium, and PET, ii) monolayer film bags made with PE, iii) monolayer film bags made with biobased PE, and iv) rigid HDPE containers.

### Waste management options identification

The most common plastic waste management technologies consist of landfilling, incineration, and mechanical recycling. Given the limited landfill disposal and

growing environmental concerns, more efforts have been put into the development of new recycling technologies. Therefore, processes like energy recovery, mechanical recycling, and chemical recycling have been proposed. Chemical recycling through pyrolysis transforms mixed plastic waste (MPW) into a chemical feedstock that can be used instead of virgin materials to remanufacture polymers. Alternatively, STRAP emerges as an alternative to recycle multilayer films by using solvents to separate the film into its different layers [19-20]. The process selectively dissolves target polymers of multicomponent waste with specific solvents.

Packaging waste management begins with the collection and sorting of municipal waste, and the subsequent production of bales from plastic waste all of which takes place at a materials recovery facility (MRF). The second stage that transforms these plastic bales into a suitable feedstock for the different upcycling technologies occurs in a plastics reprocessing facility (PRF). In PRFs, plastic bales are reduced into plastic flakes and impurities are removed from the stream.

Therefore, in this study five options for plastic packaging waste management are considered: i) mechanical recycling, ii) STRAP, iii) pyrolysis, iv) incineration, and v) landfilling. All these waste management options have as a common denominator MRFs. Pyrolysis and mechanical recycling technologies also require PRFs.

### Economic and Environmental Impact Assessment

This case study assumes the coffee companies will be responsible for recycling the post-consumer waste packaging following the Extended Producer Responsibility trend. The scope considered for the TEA starts with the purchasing of the plastic packaging by the coffee company. Moreover, the fixed and variable operational costs for the recycling technologies are considered, as well as the sales of the resulting products. The packaging costs are estimated according to the price market. The recycling technologies costs are calculated based on literature data [21-23]. The products costs are defined based on the quality and the corresponding price in the market. The difference between the costs of packaging and recycling, and the revenue from sales provides the minimum increase in price, a metric that is used to quantify the economic behavior.

For the GHG emissions assessment, the production of the packaging, along with the recycling process are considered. The climate change factors contemplated to perform the LCA are obtained from the Environmental Footprint database [24-25]. The factors of climate change used for the packaging take into consideration the steps that the production of each packaging requires, such as pellets production, film extrusion, or lamination. The same applies to the climate change impact for the

recycling technologies. The factors from the stages of the recycling processes, MRF and PRF, are included in the applicable scenarios. The different inputs required for the recycling technologies are collected from studies in the literature [21-23].

Previous studies have reported the economic and environmental feasibility of using STRAP to recover the constituent polymers of diverse multilayer films [19-20]. However, the multilayer film commonly used for coffee packaging (composed of PE, PET, and Al) has not been studied. Following the previous approaches, in this work, we propose a STRAP variation to treat such multilayer films. We performed STRAP experiments that showed that the recovery of all components and ink removal was feasible. The experimental yields were higher than 95 wt%. and the composition of the packaging was determined to be: 49 wt% of PE, 32 wt% of PET, 17 wt% of aluminium and 2 wt% of ink. The recovered polymers have comparable properties to the corresponding virgin resins. The collected experimental data was used to simulate the STRAP process and perform the TEA and LCA analysis [26].

### Superstructure Representation and multi-objective MILP model

For each packaging, all the possible options among waste management technologies and the subsequent products are contemplated. Figure 1 illustrates the superstructure considered for this study.

Following the alternatives previously described, seven scenarios are considered for the four most common coffee packaging options. Pyrolysis, landfill, and incineration are contemplated for all the packaging, as shown in the superstructure in Figure 1. Pyrolysis is divided into two scenarios: one that extends to the production of the pyrolysis oil and another that additionally involves the polymerization of new polymers from the pyrolysis oil. Mechanical recycling can process only monolayer packaging (films or rigid containers) and STRAP, only multilayer films. Two scenarios were defined for STRAP: one with ink removal that separates all the layers of the packaging using: one solvent to remove the PE layer, a second solvent to separate the aluminum and PET leftover along with the ink, and an eddy current separator to separate the metal from the polymer. The second scenario with no ink removal just removes the initial layer of PE (which constitutes the largest portion following the common composition of coffee packaging [27]) and directs the remaining residue to landfill. The use of more solvents conveys a larger use of resources that generate more emissions, then a scenario considering STRAP and landfill is analyzed as an additional alternative. The basis for the study is defined as 1000 containers.

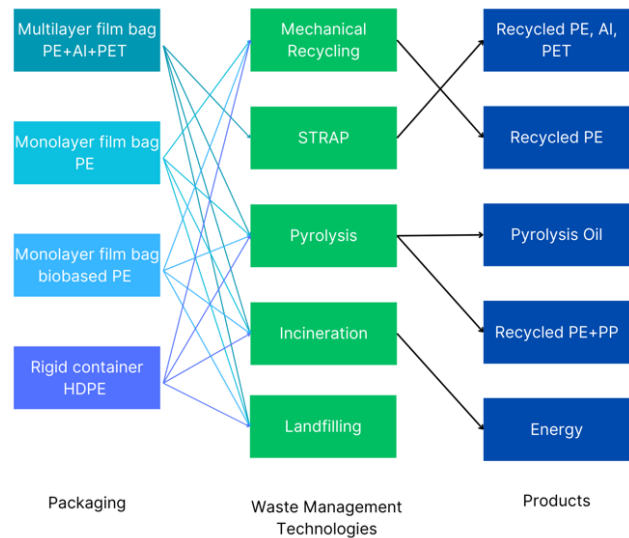


Figure 1. Superstructure considered in the study.

Under the economic criteria, the goal is to find the minimum increase in price that the company or the customer will have to cover. On the other hand, for the environmental criteria, the objective is to select the process that would generate the least GHG emissions. Given these two competing objectives, this work proposes the use of a multi-objective optimization to evaluate the trade-offs between the different packaging designs and waste management technologies. Hence, the resulting pareto front of this circular economy framework serves as a decision-making tool.

The basis for the study was defined based on a typical 500-gram ground coffee container, and the model was defined using the following nomenclature.

### Nomenclature

#### Sets and Index

$P, p$	set of packaging, packaging index
$R, r$	set of recycling technologies, recycling technology index
$m$	Individual materials index
$mt$	multilayer packaging
$mn$	monolayer packaging
$bm$	biobased monolayer packaging
$hp$	HDPE rigid container
$po$	pyrolysis with oil production
$pp$	pyrolysis with polymerization
$sr$	STRAP with ink removal
$sn$	STRAP with no ink removal and landfill
$mr$	mechanical recycling
$lf$	landfill



<i>we</i>	waste to energy
<b>Decision Variables</b>	
$I_p$	total input of packaging
$B_p$	number of bags for packaging p
$F_{p,r}$	flow from packaging p to technology r
$C_p$	total cost for purchasing packaging p
$CR_{p,r}$	total cost for recycling packaging p through technology r
$S_{p,r}$	sales of products obtained from processing packaging p through technology r
$D_{p,r}$	difference between cost of recycling and sales for packaging p and recycling technology r
$EP_p$	CO <sub>2</sub> eq. emissions from the production of packaging p
$ER_{p,r}$	CO <sub>2</sub> eq. emissions from the recycling of packaging p with technology r
<b>Constants</b>	
$M_p$	material per bag for packaging p
$W_p$	weight of packaging p
$W_m$	weight of individual material
$X_p$	cost of packaging p per ton
$X_m$	cost of individual material
$Cost_r$	cost of waste management technology r per ton
$Cost_{MRC}$	cost for mechanical recycling for an hdpe rigid container
$n_{p,r}$	yield of processing packaging p with technology r
$nl_{p,pp}$	yield of the polymerization of LDPE
$np_{p,pp}$	yield of the polymerization of PP
$X_{p,r}$	cost of product from packaging p processed through technology r
$X_{pp}$	cost of PP
$X_{bp}$	cost of polymerization by-products
$G_p$	Climate change factor for packaging p or material
$G_m$	Climate change factor for individual material
$G_{p,r}$	Climate change factor for packaging p processed through technology r

The packaging alternatives have a set of possible recycling technologies  $r$  available. The set of packaging is:  $P = \{mt, mn, bm, hp\}$ . The set of technologies is:  $R = \{po, pp, sr, sn, mr, lf, we\}$ . Additionally, a set of materials is defined to refer to the components present in the multi-layer film:  $M = \{PET, Al, PE\}$ .

The selection or not of the packaging and the technology is given by the following two binary variables:

$$y_p \forall p \in P \quad (1)$$

$$z_r \forall r \in RT \quad (2)$$

The decision variables depend on the following continuous variables:

$$B_p \leq 1000 y_p \quad \forall p \in P \quad (3)$$

$$\sum_p B_p = 1000 \quad \forall p \in P \quad (4)$$

$$I_p = B_p M_p \quad \forall p \in P \quad (5)$$

The amount that should come from packaging p to recycling technology  $r$  is given by the continuous variable:

$$\sum_r F_{p,r} = I_p \quad \forall r \in R, \forall p \in P \quad (6)$$

The following equations are general constraints considered for the model:

$$F_{p,sr} = 0, \forall p \in \{mn, bm, hp\} \forall r \in \{sr, sn\} \quad (7)$$

$$F_{mt,mr} = 0 \quad (8)$$

$$\sum_p y_p \geq 1 \quad \forall p \in P \quad (9)$$

$$\sum_r z_r \geq 1 \quad \forall r \in R \quad (10)$$

$$F_{p,r} \leq 100000 z_r \quad \forall r \in R, \forall p \in P \quad (11)$$

$$F_{p,r} \leq 100000 y_p \quad \forall p \in P, \forall r \in R \quad (12)$$

The packaging cost, the recycling cost and the products sales considered for the economic assessment are given by equations 13-22:

$$C_p = W_p X_p B_p \quad \forall p \neq mt \quad (13)$$

$$C_{mt} = (W_{Al} X_{Al} + W_{PET} X_{PET} + W_{PE} X_{PE}) B_{mt} \quad (14)$$

$$CR_{p,r} = F_{p,r} Cost_r \quad \forall p \in P, \forall r \in R - \{i = hp, j = mr\} \quad (15)$$

$$CR_{hp,mr} = F_{hp,mr} Cost_{MRC} \quad (16)$$

$$S_{p,r} = F_{p,r} n_{p,r} X_{p,r} \quad \forall p \in P, \quad \forall r \in RT : r \neq pp, lf, sr, sn \quad (17)$$

$$S_{p,pp} = F_{p,pp} (nl_{p,pp} X_{PE} + np_{p,pp} X_{PP} + n_{p,po} + X_{bp}) \quad (18)$$

$$S_{p,sr} = (F_{p,sr} n_{p,sr}) (0.49 X_{PE} + 0.32 X_{Al} + 0.17 X_{PET}) \quad (19)$$

$$S_{p,sn} = F_{p,sn} n_{p,sn} 0.49 X_{PE} \quad (20)$$

$$S_{p,lf} = 0 \quad (21)$$

$$D_{p,r} = S_{p,r} - CR_{p,r} \quad \forall p \in P, \forall r \in R \quad (22)$$

$$Profit = \sum_p \sum_r D_{p,r} - \sum_p C_p \quad (23)$$

The computation of the emissions used for the analysis of the environmental impact follow equations 24-27:

$$EP_p = W_p G_p B_p \quad \forall p \in P: p \neq mt \quad (24)$$

$$EP_{mt} = (W_{Al}G_{Al} + W_{PET}G_{PET} + W_{PE}G_{PE}) + M_{mt}G_{mt}B_{mt} \quad (25)$$

$$ER_{p,r} = G_{p,r}F_{p,r} \quad \forall p \in P, \forall r \in R : p \neq mt \quad (26)$$

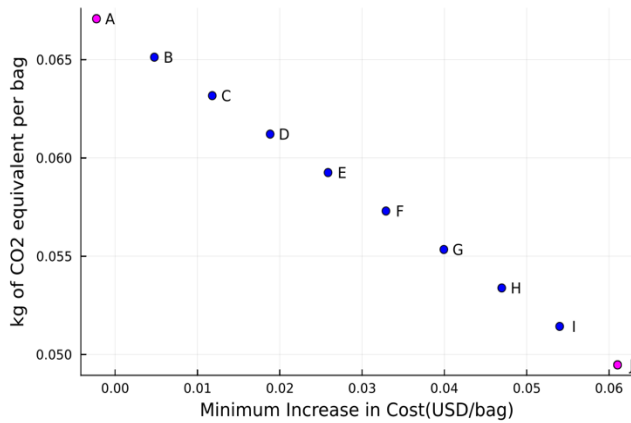
$$ER_{p,sn} = G_{p,sn}F_{p,sn} + (F_{p,sn}0.22)(0.65G_{lf} + 0.35G_{ldA}) \quad (27)$$

$$Emissions = \sum_p EP_p + \sum_p \sum_r ER_{p,r} \quad (28)$$

The model objective functions are given by the maximization of profit and the minimization of emissions, as shown in equations 23 and 28.

## DISCUSSION AND RESULTS

The mixed-integer optimization model was implemented in Julia 1.8.5 and solved with Gurobi v1.0.1. For the multi-objective problem, the  $\epsilon$ -constraint method was utilized [28]. The resulting pareto front is illustrated in Figure 2, with the details of each pareto solution presented in Table 1.



**Figure 2.** Superstructure considered in the study.

**Table 1:** Percentage of multilayer film bags processed by the selected technologies for the points of the obtained pareto front.

Points	Technology	Bags Allocation
● A	STRAP with ink removal	100%
● B	STRAP with ink removal	89%
	Landfilling	11%
● C	STRAP with ink removal	78%
	Landfilling	22%
● D	STRAP with ink removal	67%
	Landfilling	33%
● E	STRAP with ink removal	56%
	Landfilling	44%
● F	STRAP with ink removal	44%
	Landfilling	56%
● G	STRAP with ink removal	33%
	Landfilling	67%
● H	STRAP with ink removal	22%

	Landfilling	78%
● I	STRAP with ink removal	11%
	Landfilling	89%
● J	Landfilling	100%

The results indicate that the most profitable scenario is STRAP with ink removal for a multilayer film with an increase in price of approximately zero per bag due to the high selling price of the recycled polymers. However, its emissions reach 0.067 kg of CO<sub>2</sub>eq. per bag. The minimum increase in price shows a negative value in the pareto front due to a minimum profit of 0.0018 \$ per bag that for simplicity is approximated to zero. On the other hand, the least GHG-emitting alternative is the landfilling of multilayer packaging with the lowest emissions of 0.05 kg CO<sub>2</sub>eq. per bag. Nevertheless, the extra cost reaches \$0.06 per bag.

The packaging type selected in all pareto points is the multilayer PE-Al-PET bag. The set of solutions show possibilities for: a combination of two technologies (blue points) or just one technology (pink points). The technologies contemplated are STRAP with ink removal and landfilling. Given the set of solutions found, it is shown that the reduction of materials together with the proper feasible recycling technology are key to sustainable packaging solutions.

The costs and emissions for processing different packaging types even when using the same technology differ due to the different amounts of material present in each type of container. A trade-off between the amount of material needed and the desired performance of the packaging is highlighted. Multilayer films require less amount of material than monolayer packaging to achieve a high performance. However, the multilayer films characterize by greater complexity that results in a challenge for traditional recycling technologies. Therefore, the recently proposed process STRAP represents a breakthrough technology. It would solve the recyclability limitation generally attributed to multilayer films, while promoting a packaging solution that is less resource intensive. Furthermore, multilayer films represent a substantial advantage as the light weight leads to decreased energy consumption during transportation and consequently lower emissions.

The integration of environmental and economic considerations is suitable to determine the most convenient sustainable solutions. The LCA takes into consideration the emissions generated from the resources used in each of the technologies, like solvents, water withdrawal and energy requirements. Then, the CO<sub>2</sub> equivalent emissions obtained from the LCA, enables an appropriate comparison of the environmental impact of the different processes. However, there are other environmental impacts that should be considered. While landfilling has been selected as the most environmentally convenient alternative when considering the GHG emissions, there are other

potential impacts of landfills to the environment including: soil contamination, water and air pollution, and loss of resources. Landfilling is recognized to have long-lasting environmental impacts. Even after closure treatment and some years, landfills can still release gases and leach contaminants within the surrounding area [29].

The model enables the quantification of alternatives that, while initially deemed beneficial, were ultimately excluded from the set of chosen solutions. A packaging choice anticipated as a green alternative, yet not selected was the biobased polyethylene bag. Despite biobased polyethylene being an environmentally friendly alternative due to its origin, it involves higher prices than fossil-fuel derived polymers and accounts for the same disposal issue (recycling requirement) as fossil-based plastics. The emissions for the biobased packaging (based on the system boundaries) with the landfill alternative increase to 0.29 kg CO<sub>2</sub>eq. per bag. Another packaging option that was not selected as solution is the HDPE rigid container. Even though the packaging facilitates the recycling of the plastic waste, the amount of material required makes it less environmentally friendly. For this alternative, the emissions increase to 0.23 kg CO<sub>2</sub>eq per bag considering landfill.

Besides recycling, the scope of the study takes into consideration the resulting products of the process. The outputs for the different scenarios are analyzed given that the products (and even the quality of the pellets) might vary. Depending on the products, the sales can be altered. Some LCA studies that compare mechanical with chemical recycling, don't take into consideration important factors such as the quality of the outputs of the process, affecting the alternatives that provide a higher quality recycled product. The ideal case would be to obtain the most valuable product, like in the case of pyrolysis with polymerization, however, that would imply higher costs and emissions for the further processing of the material. The output of STRAP for the coffee packaging is of high quality and is obtained at a low cost, providing an economically feasible scenario.

As previously noted, certain recycling technologies may degrade the plastic's quality because they affect its mechanical properties [31]. Others instead produce a virgin-like quality plastic without deterioration and application restrictions [3]. Moreover, some processes might produce more waste or consume more resources than others. Consequently, the next logical step involves implementing a comprehensive circularity assessment framework to holistically explore environmental impacts [32].

## CONCLUSIONS

Plastic packaging plays a significant role in modern society given its multiple benefits; however, a

responsible management is required. Global efforts are put into the development of sustainable alternatives to reduce environmental impacts, such initiatives highlight the need to evaluate the required recycling infrastructure. The CE framework proposed in this work contributes to literature by analyzing the optimal packaging configuration based on the recycling technologies available to process it and materials used for packaging. Packaging solutions that are environmentally and economically viable are required. Both objectives are evaluated individually and as competing objectives. The use of multi-objective optimization generates a set of solutions valuable for decision-making. The case study of coffee provides an insight into the minimum increase in price if coffee companies were responsible for the packaging recycling. The prices are obtained for a combination of technologies with the best packaging resulting in different levels of emissions.

Despite their initial classification as environmentally unfriendly, multilayer films prove to be the most convenient option when equipped with a recycling technology to process it like STRAP. The technologies that are part of the solution encompass STRAP with ink removal and landfilling. The most economically viable technology is STRAP with a zero-price increase, but emissions of 0.07 kgCO<sub>2</sub> eq. per bag. Conversely, the most environmentally friendly alternative is landfill with 0.05 kgCO<sub>2</sub>eq. per bag, but an extra cost of 0.06 USD per bag. However, given the multiple environmental impacts caused by landfilling, further analysis is required. The toxicity of the leachate is one of the principal sources of contamination. Therefore, an analysis of their composition and their related impacts can provide insights in the environmental and human health effects [29,33]. This can be quantified under a metric that considers toxicity or the impact on water, air, land quality and human health. An alternative to assign a monetary value to environmental goods is to use an economic valuation technique [34]. Additionally, for future work a further analysis for biobased packaging can be explored taking into account carbon uptake with different system boundaries and considerations.

Regardless, the environmental and economic assessment is a good approximation and first approach, a holistic evaluation of circularity needs to be developed. For that reason, the future work will focus on applying a circularity calculator to the different management technologies to extend the analysis and visualize the trade-off between circular and economically feasible scenarios. The ultimate goal would be to provide insights to design more sustainable packaging and supply chains. Furthermore, there are a few biodegradable packaging options emerging today. However, it will be considered in the future study together with reusing practices and alternative disposal technologies such as composting and home-composting.

## ACKNOWLEDGEMENTS

This material is based upon work supported by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Bioenergy Technologies Office under Award Number DEEE0009285, and the University of Wisconsin-Madison.

## REFERENCES

1. Stegmann P, Daioglou V, Londo M, van Vuuren DP, Junginger M. Plastic futures and their CO<sub>2</sub> emissions. *Nature* 612, 7939:272-276 (2022)
2. EPA. <https://www.epa.gov/facts-and-figures-about-materials-waste-and-recycling/containers-and-packaging-product-specific#:~:text=Containers%20and%20packaging%20make%20up,beverages%2C%20medications%20and%20cosmetic%20products>.
3. Jeswani H, Krüger C, Russ M, Horlacher M, Antony F, Hann S, Azapagic A. Life cycle environmental impacts of chemical recycling via pyrolysis of mixed plastic waste in comparison with mechanical recycling and energy recovery. *Sci. Total Environ.* 769, 144483 (2021)
4. Zhu Z, Liu W, Ye S, Batista L. Packaging design for the circular economy: A systematic review. *Sustain. Prod. Consum.* 32, 817-832 (2022).
5. UNEP. <https://www.unep.org/news-and-stories/story/were-gobbling-earths-resources-unsustainable-rate>
6. Avraamidou S, Baratsas SG, Tian Y, Pistikopoulos EN. Circular Economy-A challenge and an opportunity for Process Systems Engineering. *Comput. Chem. Eng.* 133: 106629 (2020)
7. Czarnecka-Komorowska D, Wiszumirska K. Sustainability design of plastic packaging for the Circular Economy/Zrownowazone projektowanie opakowan z tworzyw sztucznych w gospodarce cyrkularnej. *Polimery* 65(1), 8-18 (2020).
8. Grégoire N, Chauvelot I. Accelerating transition to a circular economy in plastics. *Field Actions Sci. Rep.* (Special Issue 19):44-53 (2019)
9. Karayılan S, Yılmaz Ö, Uysal Ç, Naneci S. Prospective evaluation of circular economy practices within plastic packaging value chain through optimization of life cycle impacts and circularity. *Resour Conserv Recycl* 173:105691 (2021).
10. Mastellone ML. Technical description and performance evaluation of different packaging plastic waste management's systems in a circular economy perspective. *Sci. Total Environ.* 718:137233 (2020)
11. Toniolo S, Mazzi A, Niero M, Zuliani F, Scipioni A. Comparative LCA to evaluate how much recycling is environmentally favourable for food packaging. *Resour Conserv Recycl.* 77:61-8 (2013)
12. Siracusa V, Ingrao C, Giudice AL, Mbohwa C, Dalla Rosa M. Environmental assessment of a multilayer polymer bag for food packaging and preservation: An LCA approach. *Food Res. Int.* 62:151-61 (2014)
13. Xie M, Bai W, Bai L, Sun X, Lu Q, Yan D, Qiao Q. Life cycle assessment of the recycling of Al-PE (a laminated foil made from polyethylene and aluminum foil) composite packaging waste. *J. Clean. Prod.* 112:4430-4 (2016)
14. Accorsi R, Cascini A, Cholette S, Manzini R, Mora C. Economic and environmental assessment of reusable plastic containers: A food catering supply chain case study. *Int. J. Prod. Econ.* 152:88-101 (2014)
15. Munguia-Lopez AC, Göreke D, Sánchez-Rivera KL, Aguirre-Villegas HA, Avraamidou S, Huber GW, Zavala VM. Quantifying the environmental benefits of a solvent-based separation process for multilayer plastic films. *Green Chem.* 25(4): 1611-1625. (2023)
16. Komly CE, Azzaro-Pantel C, Hubert A, Pibouleau L, Archambault V. Multiobjective waste management optimization strategy coupling life cycle assessment and genetic algorithms: Application to PET bottles. *Resour Conserv Recycl.* 69:66-81 (2012)
17. Sakamoto H, Bruschi LT, Kulay L, Yamakami A. Using the Life Cycle Approach for Multiobjective Optimization in the Context of the Green Supply Chain: A Case Study of Brazilian Coffee. *Sustainability.* 15(18):13987 (2013)
18. Baratsas SG, Pistikopoulos EN, Avraamidou S. A systems engineering framework for the optimization of food supply chains under circular economy considerations. *Sci. Total Environ.* 794:148726 (2021)
19. Sánchez-Rivera KL, Munguía-López AC, Zhou P, Cecon VS, Yu J, Nelson K, Miller D et al. Recycling of a post-industrial printed multilayer plastic film containing polyurethane inks by solvent-targeted recovery and precipitation. *Resour. Conserv. Recycl.* 197: 107086 (2023)
20. Sánchez-Rivera KL, Zhou P, Kim MS, González Chávez LD, Grey S, Nelson K, Wang SC et al. Reducing antisolvent use in the strap process by enabling a temperature-controlled polymer dissolution and precipitation for the recycling of multilayer plastic films. *ChemSusChem* 14, no. 19 : 4317-4329 (2021)
21. Ma J, Tominac PA, Aguirre-Villegas, H. A., Olafasakin OO, Wright MM, Benson CH, Huber GW, Zavala VM Economic evaluation of infrastructures

- for thermochemical upcycling of post-consumer plastic waste. *Green Chem.* 25, 1032-1044 (2023)
22. Larrain M, Van Passel S, Thomassen G, Van Gorp B, Nhu TT, Huysveld S, Van Geem KM, De Meester S, Billen P. Techno-economic assessment of mechanical recycling of challenging post-consumer plastic packaging waste. *Resour Conserv Recycl* 170: 105607(2021)
  23. Rudolph N, Kiesel R, Aumnate C. Understanding Plastics Recycling: Economic, Ecological, and Technical Aspects of Plastic Waste Handling. Carl Hanser Verlag GmbH Co KG (2020)
  24. Ciroth A, Di Noi C, Lohse T, Srocka M. OpenLCA 1.10 Comprehensive User Manual. GreenDelta, Berlin, Germany (2020)
  25. Fazio S, Castellani V, Sala S, Schau E, Secchi M, Zampori L, Diaconu E. Supporting Information to the Characterisation Factors of Recommended EF Life Cycle Impact Assessment Methods: New Methods and Differences with ILCD. (2018).
  26. Cortes-Pena Y, Kumar D, Singh V, Guest, JS. BioSTEAM: a fast and flexible platform for the design, simulation, and techno-economic analysis of biorefineries under uncertainty. *ACS Sustain. Chem. Eng.* 8, no. 8: 3302-3310 (2020)
  27. Büsser S, Jungbluth N. The role of flexible packaging in the life cycle of coffee and butter. *Int J LCA.* 14:80-91. (2009)
  28. Mavrotas G. Effective implementation of the  $\epsilon$ -constraint method in multi-objective mathematical programming problems. *Appl. Math. Comput.* 213(2):455-65 (2009).
  29. Vaverková MD. Landfill impacts on the environment. *Geosci.* 9(10):431(2019)
  30. Benavides PT, Lee U, Zarè-Mehrjerdi O. Life cycle greenhouse gas emissions and energy use of polylactic acid, bio-derived polyethylene, and fossil-derived polyethylene. *J. Clean. Prod.* 277:124010 (2020)
  31. Yin S, Tuladhar R, Shi F, Shanks RA, Combe M, Collister T. Mechanical reprocessing of polyolefin waste: A review. *Polym Eng Sci.* 55(12):2899-909 (2015)
  32. Baratsas SG, Pistikopoulos EN, Avraamidou S. A quantitative and holistic circular economy assessment framework at the micro level. *Comput Chem Eng.* 160:107697 (2022)
  33. Khoiron K, Probandari AN, Setyaningsih W, Kasjono HS, Setyobudi RH, & Anne O. A review of environmental health impact from municipal solid waste (MSW) landfill. *Annals of tropical medicine and public health.*, 23(3), 60-67. (2020)
  34. European Commission. A Study on the Economic Valuation of Environmental Externalities from Landfill Disposal and Incineration of Waste. (2000)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Integrating the Design of Desalination Technologies into Produced Water Network Optimization

Sakshi Naik<sup>a</sup>, Miguel Zamarripa<sup>b,c</sup>, Markus Drouven<sup>b</sup>, Lorenz T. Biegler<sup>a\*</sup>

<sup>a</sup> Carnegie Mellon University, Pittsburgh, PA 15232, USA

<sup>b</sup> National Energy Technology Laboratory (NETL), Pittsburgh, PA 15236, USA

<sup>c</sup> NETL Support Contractor, Pittsburgh, PA 15236, USA

\* Corresponding Author: [lb01@andrew.cmu.edu](mailto:lb01@andrew.cmu.edu)

---

## ABSTRACT

The oil and gas energy sector uses billions of gallons of water for hydraulic fracturing each year to extract oil and gas. The water injected into the ground for fracturing along with naturally occurring formation water from the oil wells surfaces back in the form of produced water. Produced water can contain high concentrations of total dissolved solids and is unfit for reuse outside the oil and gas industry without desalination. In semi-arid shale plays, produced water desalination for beneficial reuse could play a crucial role in alleviating water shortages and addressing extreme drought conditions. In this paper we co-optimize the design and operation of desalination technologies along with operational decisions across produced water networks. A multi-period produced water network model with simplified split-fraction-based desalination nodes is developed. Rigorous steady-state desalination mathematical models based on mechanical vapor recompression are developed and embedded at the desalination sites in the network model. An optimal common design is ensured across all periods using global capacity constraints. The solution approach is demonstrated for multi-period planning problems on networks from the PARETO open-source library. Model formulation and challenges associated with scalability are discussed.

---

**Keywords:** Water Networks, Process Design, Modelling, Pyomo, Optimization

## 1 INTRODUCTION

Hydraulic fracturing accounts for about half of current US crude oil production [1]. It involves the injection of large volumes of water underground to build pressure and force open rocks to release underground oil and gas reserves. The injected water along with natural formation water is returned to the surface as shown in Figure 1 and contains dissolved solids, oil, grease, and chemicals used during the well stimulation. The volume and quality of the water varies over time and is different from well to well. The oil and gas industry produces billions of gallons of high total dissolved solids (TDS) produced water every year due to oil and gas operations. There is a need to manage that water in cost-efficient and environmentally sustainable ways.

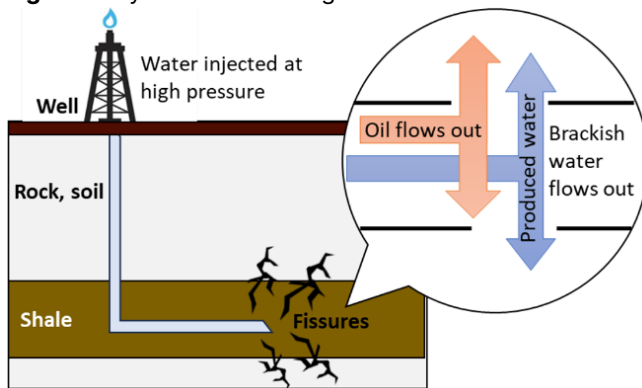
Currently, produced water is either disposed or recycled for well development purposes with pre-treatment by adding friction reducers or other chemicals. Strategic planning and optimal operational policies in

produced water management have shown to increase water reuse within the oil and gas industry, making the process more sustainable [2-3]. Several efforts have focused on building infrastructure for produced water management for economic and environmental benefits. For example, optimization models that minimize the cost of produced water management have been developed that consider reuse, storage and simplified treatment [4]. Multi-period planning models involving network infrastructure expansion and strategic decisions have been shown to significantly improve profitability for upstream operators [5]. Due to large quantities of produced water and fracking scheduling, it may not be feasible to recycle a 100% of produced water within the oil and gas industry [6].

To reuse produced water outside the oil and gas industry, it requires extensive treatment (i.e., desalination) to bring its quality to acceptable levels by lowering TDS concentrations. Desalination technologies for produced water are mainly classified into thermal-based [7-9] and

membrane-based [10–11] technologies. Several studies have focused on developing detailed mathematical models for the desalination units to develop techno-economic analyses and determine the effectiveness of different desalination technologies [12–14]. Multi-effect evaporation with mechanical vapor recompression (MEE-MVR) is a potential desalination technology for produced water at an industrial scale [7]. A thermo-economic model for an MEE-MVR process has been developed in [15]. Also, detailed nonlinear programming (NLP) models for the MEE-MVR process were developed in [16] with heat integration for the desalination of produced water from shale gas plays. These models were further used to minimize the total annualized cost (TAC) of the desalination unit and determine the optimal design and operation given the feed flow and salinity of produced water.

**Figure 1:** Hydraulic fracturing



Although produced water network optimization and optimization of rigorous desalination units have been individually explored, there is a gap with respect to integrating these two research areas to create an overall optimal decision-making system. In [12] a network optimization model was developed for the Marcellus shale play by using a treatment cost model, obtained using techno-economic assessment of membrane distillation. Simplified costing models for treatment have been used in several works which consider produced water management optimization [2, 4]. However, these formulations don't provide the optimal treatment system design and operational decisions of the treatment units.

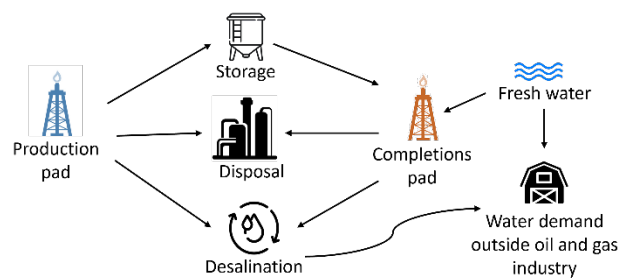
This work supports the development of PARETO – a DOE-sponsored, free, and open-source optimization framework for onshore produced water management ([www.project-pareto.org](http://www.project-pareto.org)), by integrating detailed desalination models into the existing produced water management framework. Specifically, in this paper, we introduce a formulation for integrating steady-state desalination units into multi-period produced water network models for the co-optimization of design and operational decisions. A multi-period produced water network model consisting of production pads, completion pads, storage, disposal, and desalination units has been developed.

Since the resulting integrated formulation is nonlinear, it is crucial to ensure that there are no singularities arising due to zero flows in the network. Therefore, separate flow variables for solid and liquid flows are used for the mass balances to avoid bilinear terms in each network arc. We use the MEE-MVR model [16] with heat integration for the rigorous desalination unit with a separate MEE-MVR model for each period. A common design is ensured across all periods using global capacity constraints on the design variables. In Section 2 we provide the mathematical formulation for the produced water network model, desalination units, and the integrated problem. Section 3 consists of case studies on a multi-period network with integrated desalination models. Section 4 concludes the paper with a discussion on scaling up the problem to larger networks.

## 2 MODEL FORMULATION

### Network Model

A typical produced water network consists of production pads (well sites) comprising mature oil or gas wells, which produce hydrocarbons and water, and completion pads, which have a water demand for drilling and hydraulic fracturing. It also consists of disposal wells, storage units, and – in the future – may involve desalination units where excess produced water can be sent. Completions pads' water demand (for hydraulic fracturing) can be met using existing water from the production pads or by procuring fresh water; this depends upon the cost of transportation vs. the cost of sourcing. Figure 2 shows a schematic of a produced water network along with all possible network arcs.



**Figure 2:** Produced water network.

In this section we describe our model formulation with mass balances for the produced water network model. Table 1 describes all the sets, variables, and parameters used in the formulation.

**Table 1:** Description of sets, variables, and parameters for the network model.

#### Sets

$\mathcal{N}^P, \mathcal{N}^C, \mathcal{N}^{MS}$	Production pads, completion pads, mixers, or splitters respectively
$\mathcal{N}^{T,IN}, \mathcal{N}^{T,TW}, \mathcal{N}^{T,CW}$	Desalination inlet, treated water, and concentrated water nodes
$\mathcal{T}$	Total time periods
$\mathcal{T}_n^P, \mathcal{T}_n^C, \mathcal{T}_n^0$	Periods when node $n$ is producing, consuming, or neither, respectively
$\mathcal{Q}$	Components in the stream
$\mathcal{A}$	Arcs in the network
$\mathcal{A}_n^{In}$	Arcs to the inlet of node $n$
$\mathcal{A}_n^{Out}$	Arcs from the outlet of node $n$
$\mathcal{N}^{Up}(n)$	Upstream node of node $n$
$a^{In}(n), a^{Out}(n)$	Arcs in or out of node $n$ , respectively

### Variables

$F_{at}$	Flowrate of water in arc $a$ at time $t$
$S_{aqt}$	Flowrate of solids $q$ in arc $a$ at time $t$
$C_{nqt}$	Concentration of $q$ at node $n$ at time $t$
$C_{nt}^{total}$	Total concentration at node $n$ at time $t$
$I_{nt}$	Water in inventory unit $n$ at time $t$
$I_{nqt}^s$	Solids $q$ in inventory unit $n$ at time $t$
$\alpha_{nt}$	Water recovery fraction at desalination unit $n$ at time $t$

### Parameter

$\hat{F}_{nt}^{Prod}, \hat{F}_{nt}^{Cons}$	Flowrate of water produced and consumed at node $n$ at time $t$
$\hat{S}_{nqt}^{Prod}, \hat{S}_{nqt}^{Cons}$	Flowrate of solids $q$ produced and consumed at node $n$ at time $t$
$\hat{C}_{nqt}^{Prod}$	Concentration of solids $q$ produced at node $n$ at time $t$
$\hat{I}_n^0$	Initial water inventory at the inventory node
$\hat{I}_n^{s0}$	Initial TDS level at the inventory node
$\hat{C}_{min}^{total}, \hat{C}_{max}^{total}$	Concentration bounds for the desalination inlet stream
$\hat{F}_n^{min}, \hat{F}_n^{max}$	Flowrate bounds for desalination inlet node $n$
$\Delta t$	Time discretization

A general flow balance for water and solids can be written for each network connecting the mixers, splitters, and the production pads:

$$\sum_{a \in \mathcal{A}_n^{In}} F_{at} + \hat{F}_{nt}^{Prod} = \sum_{a \in \mathcal{A}_n^{Out}} F_{at} + \hat{F}_{nt}^{Cons} \quad \forall n \in \mathcal{N}^P \cup \mathcal{N}^{MS}, t \in \mathcal{T}$$

$$\sum_{a \in \mathcal{A}_n^{In}} S_{aqt} + \hat{S}_{nqt}^{Prod} = \sum_{a \in \mathcal{A}_n^{Out}} S_{aqt} + \hat{S}_{nqt}^{Cons} \quad \forall n \in \mathcal{N}^P \cup \mathcal{N}^{MS}, q \in \mathcal{Q}, t \in \mathcal{T}$$

The outlet lines of the splitters contain equal concentration of solids. This constraint can be expressed as:

If  $|\mathcal{A}_n^{Out}| > 1$ ,

$$\hat{S}_{iqt} \sum_{a \in \mathcal{A}_n^{In}} F_{at} = F_{it} \sum_{a \in \mathcal{A}_n^{Out}} S_{aqt} \quad \forall i \in \mathcal{A}_n^{Out}, n \in \mathcal{N}^{MS}, q \in \mathcal{Q}, t \in \mathcal{T}$$

For the inventory nodes, the mass balance is given by:

$$I_{nt} = I_{nt-1} + \sum_{a \in \mathcal{A}_n^{In}} F_{at} - \sum_{a \in \mathcal{A}_n^{Out}} F_{at} \quad \forall n \in \mathcal{N}^S, t \in \mathcal{T}$$

If there is a pretreatment unit installed before the inventory node, the concentration of solids in the inventory is zero. Otherwise, a solids mass balance is required to track the concentration of solids in the inventory:

$$I_{nqt}^s = I_{nqt-1}^s + \sum_{a \in \mathcal{A}_n^{In}} S_{aqt} - \sum_{a \in \mathcal{A}_n^{Out}} S_{aqt} \quad \forall n \in \mathcal{N}^S, q \in \mathcal{Q}, t \in \mathcal{T}$$

The solids flow out of the inventory, depend on the concentration of the stored water and are given by:

$$S_{aqt} = F_{at} C_{nqt}, \quad C_{nqt} I_{nt} = I_{nqt}^s \quad \forall n \in \mathcal{N}^S, q \in \mathcal{Q}, t \in \mathcal{T}$$

Completions pads can consume water, produce water, do both simultaneously, or neither produce nor have demand. Depending upon the mode of operation, the completions pad mass balances need to be written differently. If the completions pad is consuming water:

$$\hat{F}_{nt}^{Cons} = \sum_{a \in \mathcal{A}_n^{In}} F_{at} \quad \forall n \in \mathcal{N}^C, t \in \mathcal{T}_n^C$$

$$\sum_{a \in \mathcal{A}_n^{Out}} F_{at} = 0, \quad \sum_{a \in \mathcal{A}_n^{Out}} S_{aqt} = 0 \quad \forall n \in \mathcal{N}^C, q \in \mathcal{Q}, t \in \mathcal{T}_n^C$$

If the completions pad is producing water:

$$\hat{F}_{nt}^{Prod} = \sum_{a \in \mathcal{A}_n^{Out}} F_{at} \quad \forall n \in \mathcal{N}^C, t \in \mathcal{T}_n^P$$

$$\hat{S}_{nqt}^{Prod} = \sum_{a \in \mathcal{A}_n^{Out}} S_{aqt} \quad \forall n \in \mathcal{N}^C, t \in \mathcal{T}_n^P, q \in \mathcal{Q}$$

$$\sum_{a \in \mathcal{A}_n^{In}} F_{at} = 0, \quad \sum_{a \in \mathcal{A}_n^{In}} S_{aqt} = 0 \quad \forall n \in \mathcal{N}^C, q \in \mathcal{Q}, t \in \mathcal{T}_n^P$$

If the completions pad is simultaneously producing and consuming:

$$\hat{F}_{nt}^{Prod} = \sum_{a \in \mathcal{A}_n^{Out}} F_{at}, \quad \hat{F}_{nt}^{Cons} = \sum_{a \in \mathcal{A}_n^{In}} F_{at} \quad \forall n \in \mathcal{N}^C, t \in \mathcal{T}_n^{PC}$$

$$\hat{S}_{nqt}^{Prod} = \sum_{a \in \mathcal{A}_n^{Out}} S_{aqt} \quad \forall n \in \mathcal{N}^C, t \in \mathcal{T}_n^{PC}, q \in \mathcal{Q}$$

If the completions pad is neither consuming nor producing water:

$$\sum_{a \in \mathcal{A}_n^{In}} F_{at} = 0, \quad \sum_{a \in \mathcal{A}_n^{Out}} F_{at} = 0 \quad \forall n \in \mathcal{N}^C, t \in \mathcal{T}_n^0$$

$$\sum_{a \in \mathcal{A}_n^{In}} S_{aqt} = 0, \quad \sum_{a \in \mathcal{A}_n^{In}} S_{aqt} = 0 \quad \forall n \in \mathcal{N}^C, q \in \mathcal{Q}, t \in \mathcal{T}_n^0$$

The desalination unit node in the network is broken down into three parts: the inlet node, the concentrated water node, and the treated water node. For the desalination inlet node, we ensure that there is only one inlet connection. If there are multiple inlet lines going into a desalination unit, we add a mixer node with one outlet that goes to the desalination unit. The desalination inlet node equations are given by:

$$C_{nt}^{total} \sum_{a \in \mathcal{A}_n^{In}} F_{at} = \sum_{q \in \mathcal{Q}} \sum_{a \in \mathcal{A}_n^{In}} S_{aqt} \quad \forall n \in \mathcal{N}^{T,In}, t \in \mathcal{T}$$

The treated water stream from the desalination unit is considered fresh with zero salinity. This assumption is true in the case of MEE-MVR, since the condensate from the evaporators contains negligible amounts of solids. The concentrated water node consists of the brine from the desalination unit. The flow in and flow out are related by the water recovery fraction:

$$F_{a^{out}(n)t} = \alpha_{n't} F_{a^{in}(n')t} \quad \forall n \in \mathcal{N}^{T,TW}, n' \in \mathcal{N}^{Up}(n), t \in \mathcal{T}$$

$$F_{a^{out}(n)t} = (1 - \alpha_{n't}) F_{a^{in}(n')t} \quad \forall n \in \mathcal{N}^{T,CW}, n' \in \mathcal{N}^{Up}(n), t \in \mathcal{T}$$

$$C_{nqt} F_{a^{out}(n)t} = C_{n'qt} F_{a^{in}(n')t}, S_{a^{out}(n)qt} = C_{nqt} F_{a^{out}(n)t} \quad \forall n \in \mathcal{N}^{T,CW}, n' \in \mathcal{N}^{Up}(n), t \in \mathcal{T}, q \in \mathcal{Q}$$

We assume that the desalination unit has a minimum flow and concentration restriction. This assumption is reasonable as the startup procedure for the desalination units can be time consuming and expensive compared to running them at some minimum capacity for the entire planning horizon. Therefore, the bounds on desalination operation can be written as:

$$\hat{C}_{min}^{total} \leq C_{nt}^{total} \leq \hat{C}_{max}^{total} \quad \forall n \in \mathcal{N}^{T,In}, t \in \mathcal{T}$$

$$\hat{F}_n^{min} \leq \sum_{a \in \mathcal{A}_n^{In}} F_{at} \leq \hat{F}_n^{max} \quad \forall n \in \mathcal{N}^{T,In}, t \in \mathcal{T}$$

The disposal sites also have disposal limits on the volume of water that can be disposed in a particular period:

$$\sum_{a \in \mathcal{A}_n^{Out}} F_{at} \leq \hat{F}_n^{max} \quad \forall n \in \mathcal{N}^D, t \in \mathcal{T}$$

We assume that the salinity of the fresh water sourced is near-zero and that there are capacity constraints on the volume of water that can be sourced:

$$\sum_{a \in \mathcal{A}_n^{Out}} F_{at} \leq \hat{F}_n^{max}, \sum_{a \in \mathcal{A}_n^{Out}} S_{aqt} = 0 \quad \forall n \in \mathcal{N}^D, q \in \mathcal{Q}, t \in \mathcal{T}$$

The main advantage of including the solids flow stream into the network model is that we can bound the solids flow when the flow of water in any arc is zero:

$$S_{aqt} \leq \hat{C}^{max} F_{at} \quad \forall a \in \mathcal{A}, q \in \mathcal{Q}, t \in \mathcal{T}$$

This constraint avoids singularities in the problem formulation and improves the robustness of the model while optimizing.

The operating cost of the network includes costs for transportation, storing water in the inventory units, sourcing fresh water, and produced water disposal or injection. The desalination of water and the removal of water from the inventory unit has a reward associated with it to encourage desalination and discourage storage of produced water. The desalination cost comes from the detailed MEE-MVR process model costs and hence is not included in the network cost here:

$$C_{net} = \sum_{t \in \mathcal{T}} \left[ \sum_{a \in \mathcal{A}} \beta_{ar} F_{at} + \sum_{n \in \mathcal{N}^S} \sum_{a \in \mathcal{A}_n^{In}} \beta_s F_{at} + \sum_{n \in \mathcal{N}^W} \sum_{a \in \mathcal{A}_n^{Out}} \beta_s F_{at} \right. \\ \left. + \sum_{n \in \mathcal{N}^D} \sum_{a \in \mathcal{A}_n^{In}} \beta_d F_{at} - \sum_{n \in \mathcal{N}^S} \sum_{a \in \mathcal{A}_n^{Out}} \gamma_s F_{at} \right. \\ \left. - \sum_{n \in \mathcal{N}^{T,TW}} \sum_{a \in \mathcal{A}_n^{Out}} \gamma_T F_{at} \right] \Delta t$$

## Desalination Model

An MEE-MVR process unit consists of single or multiple evaporator stages where the feed water enters on the shell side. As seen in Figure 3, hot steam flows through the evaporator tubes and the resulting heat transfer evaporates the feed water, which is split into liquid brine and vapor streams. In a multistage evaporation setting, the vapor from the  $i^{th}$  evaporator is fed into the tubes of the  $(i+1)^{th}$  evaporator and the brine from the  $i^{th}$  evaporator is fed into the  $(i-1)^{th}$  evaporator. A compressor is used for vapor recompression from the last evaporator stage to feed it into the tubes of the first evaporator in the sequence. A preheater is used to exchange heat between the condensate and the feed to pre-heat the feed for better energy integration.

A rigorous MEE-MVR process model is developed based on [16], the mathematical model consists of mass and energy balances and non-linear pressure and temperature relationships along with bounds for feasible operation. The detailed modeling equations are provided in the digital supplementary material. The main costs associated with the desalination unit are the CAPEX from the evaporators, preheater, and the compressor and the OPEX from using electricity to run the compressor. The CAPEX was calculated using correlations from [17] and was annualized using an amortization period of 10 years and an interest rate of 10%.

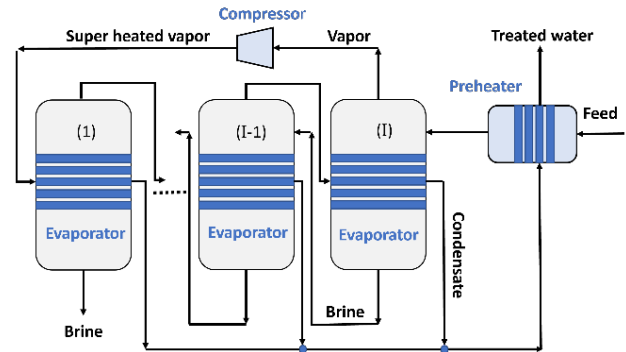


Figure 3. MEE-MVR process with heat integration.

## Integrated Model

To formulate the multi-period integrated problem, a steady-state MEE-MVR unit model is written in each period of the network. The variables from the network are linked with the input variables (represented by a bar on top of the variables) to the desalination unit using linking constraints. To ensure a common desalination design across all periods and feasible operation in every period, we use global capacity constraints on the capacity variables of the desalination model. The annualized CAPEX and OPEX are converted to daily CAPEX and OPEX given by  $CAPEX_{\Delta t}$  and  $OPEX_{\Delta t}$ . The integrated model formulation is given by:

$$\min C_{net} + \sum_{t \in \mathcal{T}} (CAPEX_{\Delta t} + OPEX_{\Delta t}^t) \Delta t$$

$$\text{s.t } \sum_{a \in \mathcal{A}_n^{\text{in}}} F_{at} = \bar{F}_{nt}^{\text{in}} \quad \forall n \in \mathcal{N}^{T, \text{In}}, t \in \mathcal{T}$$

$$C_{nt}^{\text{total}} = \bar{C}_{nt}^{\text{in}} \quad \forall n \in \mathcal{N}^{T, \text{In}}, t \in \mathcal{T}$$

$$\alpha_{nt} = \bar{\alpha}_{nt} \quad \forall n \in \mathcal{N}^{T, \text{In}}, t \in \mathcal{T}$$

$$CAPEX_{\text{evap}}^{(i) \text{global}} \geq \bar{CAPEX}_{\text{evap}, t}^{(i)} \quad \forall i \in \{1, \dots, I\}, t \in \mathcal{T}$$

$$CAPEX_{\text{ph}}^{\text{global}} \geq \bar{CAPEX}_{\text{ph}, t} \quad \forall t \in \mathcal{T}$$

$$CAPEX_{\text{comp}}^{\text{global}} \geq \bar{CAPEX}_{\text{comp}, t} \quad \forall t \in \mathcal{T}$$

All network equations and desalination model equations in each period

The  $CAPEX_{\Delta t}$  is calculated using the global CAPEX of individual equipment. It is assumed that there are no uncertainties in the model. The integrated problem is a non-linear model, developed in Pyomo [18] and solved using the open source NLP solver IPOPT [19].

### 3 RESULTS

In this section we demonstrate the results of our integrated optimization approach on a small network case study provided by “Project PARETO”. The network consists of four production pads, one completions pad, two disposal sites, two fresh water sources, one storage site, and one prospective desalination site as shown in Figure 4. A planning horizon of 52 weeks is considered with a weekly discretization. The completions pad has a water demand between weeks 1 to 12 and weeks 45 and 52. In the remaining time there is completions flowback, which means there is no demand for water within the network during that time. We perform three case studies:

1. Embedding a single-stage MVR at R01
2. Embedding at two-stage MVR at R01
3. Embedding a three-stage MVR at R01

#### Case Studies

A single-stage desalination unit is embedded at desalination site R01. The problem consists of 6,972 variables, 6,015 equality constraints, and 2,860 inequality constraints. The optimal network operation profiles are shown in Figure 5. When there is no water demand in the network, the desalination unit is run at its maximum capacity as shown in Figure 5b. When there is completions demand, the optimal solution balances recycling and desalination to reduce costs incurred from freshwater sourcing. From Figure 5c, we observe that initially, when there is completions demand, the inventory is depleted to meet the demand, and then there is inventory build-up to meet the completions demand in the later weeks of the year. Finally, at the end of the year the inventory is completely depleted as there is no incentive to store excess water. The optimal solution indicates that 53% produced water is sent to desalination, 14% is reused to meet completions demand, and 33% is sent to disposal. 4.2 million

barrels of fresh water are sourced to meet the completions demand. Table 2 shows the optimal design variables for the desalination plant.

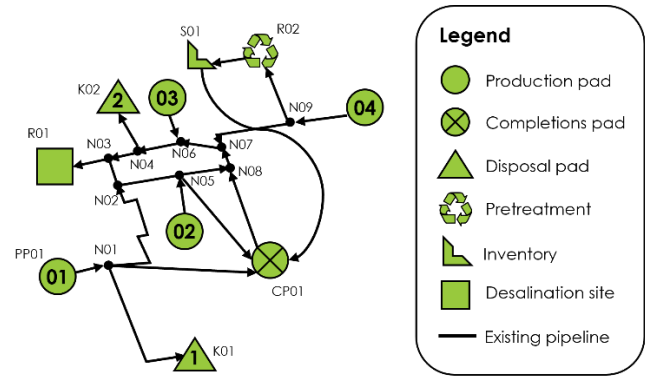


Figure 4. Produced water network case study.

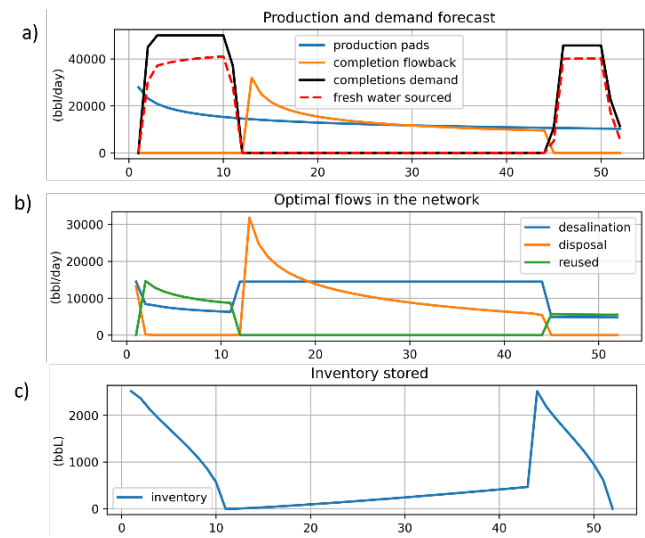


Figure 5. Optimal flow profiles for a single-stage desalination unit at R01. A) Production and demand forecasts against time and fresh water sourced in each period. b) Optimal water disposal, desalination, and reuse in each time period. c) Optimal volume of water stored in S01 in each period.

When a two-stage desalination unit is embedded at R01, the problem has 8,064 variables, 7,055 equality constraints, and 3,328 inequality constraints. The optimal flow profiles obtained are nearly identical to the single-stage unit. The optimal design variables for the two-stage desalination plant are shown in Table 2. The two-stage desalination plant needs a smaller compressor due to the separation process being split over two evaporator stages. Hence, the CAPEX and the OPEX are both smaller for a two-stage unit compared to a single-stage unit. Overall, the network with a two-stage desalination unit is 31% cheaper than with a single-stage unit in this example.

When a three-stage unit is installed, the desalination



unit doesn't run at its full capacity because the compressor reaches its maximum compression ratio at a lower inlet flowrate to the desalination unit. Thus, the water sent to desalination is decreased, with only 48% water sent to desalination compared to the two-stage and single-stage case. Also, a smaller compressor is installed but the cost of installing three evaporator stages offsets the cost reduction obtained by installing the smaller compressor. Table 3 shows the cost breakdown of the network costs and desalination costs for each case study. Overall, the two-stage desalination unit has the smallest objective function and is therefore optimal in this case. All problems solve with IPOPT within an order of 100 s.

**Table 2:** Optimal design of the desalination units obtained by solving the integrated problem.

Unit type → Equipment ↓	Single-stage	Two-stage	Three-stage
Evaporator area 1	372 m <sup>2</sup>	372 m <sup>2</sup>	372 m <sup>2</sup>
Evaporator area 2	–	372 m <sup>2</sup>	372 m <sup>2</sup>
Evaporator area 3	–	–	372 m <sup>2</sup>
Preheater area	73.87 m <sup>2</sup>	118 m <sup>2</sup>	78 m <sup>2</sup>
Compressor capacity	2750 Hp	1790 Hp	1449 Hp

**Table 3:** Optimal costs obtained by solving the integrated problem.

Unit type → Costs (kUSD) ↓	Single-stage	Two-stage	Three-stage
Objective function	9882.9	9094	9575
Transportation	370	370	366
Storage	5.8	5.03	6.12
Disposal	917.5	918	1060
Fresh water	12682	12684	12680
Desalination CAPEX	406	374.8	388
Desalination OPEX	2143.2	1384.2	1075
Storage reward	0.58	0.50	0.61
Desalination reward	6640.6	6641.7	6002

## 4 CONCLUSIONS

In this paper, we develop an integrated optimization framework for the co-optimization of desalination units with operational decision-making in produced water networks. The framework is demonstrated on a multi-period network problem with three case studies embedding MVR desalination units with different number of stages into the network. The results show that in this case, a two-stage MVR leads to the lowest objective function. The framework gives the optimal operation variables in the network along with the optimal design and operation variables in the rigorous desalination units.

Future work involves developing decomposition techniques for scaling the framework for application on larger network instances. We also plan on incorporating

bi-directional flow and strategic decision making into the network model.

## DIGITAL SUPPLEMENTARY MATERIAL

The detailed desalination unit model for MEE-MVR process is provided:

[https://pareto.readthedocs.io/en/latest/model\\_library/water\\_treatment/index.html](https://pareto.readthedocs.io/en/latest/model_library/water_treatment/index.html)

## ACKNOWLEDGEMENTS

We gratefully acknowledge support from the U.S. Department of Energy, Office of Fossil Energy and Carbon Management, through the Environmentally Prudent Stewardship Program.

### Disclaimer:

This project was funded by the Department of Energy, National Energy Technology Laboratory an agency of the United States Government, through a support contract. Neither the United States Government nor any agency thereof, nor any of their employees, nor the support contractor, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## REFERENCES

1. Troy Cook, Jack Perrin. <https://www.eia.gov/todayinenergy/detail.php?id=25372>
2. Markus Drouven, Andrés Caldéron, Miguel Zamarripa, and Keith Beattie. PARETO: An open-source produced water optimization framework. *Optimization and Engineering*, vol. 24. pp. 2229–2249 (2022)
3. Tania Itzel, Serrano-Areválo, Luis Fernando, Lira-Barragán, Mahmoud El-Halwagi, and José María Ponce-Ortega. Strategic Planning for Optimal Management of Different Types of Shale Gas Wastewater. *ACS Sustainable Chem. Eng.*, vol. 10, no. 4, pp. 1451–1470 (2022)
4. Linlin Yang, Ignacio Grossmann, and Jeremy Manno. Optimization models for shale gas water management. *AIChE Journal*, vol. 60, no. 10, pp. 3490–3501 (2014)
5. Markus Drouven and Ignacio Grossmann. Multi-period planning, design, and strategic models for

- long-term, quality-sensitive shale gas development. *AIChE Journal*, vol. 62, no. 7, pp. 2296–2323 (2016)
6. Bohyun Hwang, Joonghyeok Heo, Chungwan Lim, Joonkyu Park. Environmental Implications of Shale Gas Hydraulic Fracturing: A Comprehensive Review on Water Contamination and Seismic Activity in the United States. *Water*, vol 15 (2023)
  7. Devin Shaffer, Laura Arias Chavez, Moshe Ben-Sasson, Santiago Romero-Vargas Castrillón, Ngai Yin Yip, and Menachem Elimelech. Desalination and Reuse of High-Salinity Shale Gas Produced Water: Drivers, Technologies, and Future Directions. *Environ. Sci. Technol.*, vol. 47, no. 17, pp. 9569–9583 (2013)
  8. Alba Carrero-Parreño, Viviani Onishi, Rubén Ruiz-Femenia, Raquel Salcedo-Díaz, José Caballero, and Juan Reyes-Labarta. Optimization of multistage membrane distillation system for treating shale gas produced water. *Desalination*, vol. 460, pp. 15–27, (2019)
  9. Viviani Onishi, Eric Fraga, Juan Reyes-Labarta, and José Caballero. 12 - Desalination of shale gas wastewater: Thermal and membrane applications for zero-liquid discharge In: *Emerging Technologies for Sustainable Desalination Handbook* pp. 399–431. Ed: Veera Gnanaswar Gude. Elsevier (2018)
  10. Nur'Aishah Shahrim, Nada Abounahia, Aya El-Sayed, Haleema Saleem, and Syed Zaidi. An overview on the progress in produced water desalination by membrane-based technology. *Journal of Water Process Engineering*, vol. 51, p. 103479 (2023)
  11. Elmira Shamlou, Radisav Vidic, and Vikas Khanna. Optimization-based technoeconomic comparison of multi-stage membrane distillation configurations for hypersaline produced water desalination. *Desalination*, vol. 543, p. 116098 (2022)
  12. Sakineh Tavakkoli, Omkar Lokare, Radisav Vidic, and Vikas Khanna. A techno-economic assessment of membrane distillation for treatment of Marcellus shale produced water. *Desalination*, vol. 416, pp. 24–34 (2017)
  13. Raquel Salcedo-Díaz, Rubén Ruiz-Femenia, Alba Carrero-Parreño, Viviani Onishi, Juan Reyes-Labarta, and José Caballero. Combining Forward and Reverse Osmosis for Shale Gas Wastewater Treatment to Minimize Cost and Freshwater Consumption In: *Computer Aided Chemical Engineering*, vol. 40, no. 40. Ed: Antonio Espuña, Moisés Graells, Luis Puigjaner. Elsevier, pp. 2725–2730. (2017)
  14. Viviani Onishi, Rubén Ruiz-Femenia, Raquel Salcedo-Díaz, Alba Carrero-Parreño, Juan Reyes-Labarta, Eric Fraga, José Caballero. Process optimization for zero-liquid discharge desalination of shale gas flowback water under uncertainty. *Journal of Cleaner Production*, vol. 164, pp. 1219–1238, (2017)
  15. A. S. Nafey, H. E. S. Fath, and A. A. Mabrouk. Thermoeconomic design of a multi-effect evaporation mechanical vapor compression (MEE–MVC) desalination process. *Desalination*, vol. 230, no. 1, pp. 1–15 (2008)
  16. Viviani Onishi, Alba Carrero-Parreño, Juan Reyes-Labarta, Rubén Ruiz-Femenia, Raquel Salcedo-Díaz, Eric Fraga. Shale gas flowback water desalination: Single vs multiple-effect evaporation with vapor recompression cycle and thermal integration. *Desalination*, vol. 404, pp. 230–248 (2017)
  17. James Couper, W. Roy Penney, James Fair, and Stanley Walas. *Chemical Process Equipment: Selection and Design*. Elsevier (2005)
  18. William Hart, Jean-Paul Watson, and David Woodruff. Pyomo: modeling and solving mathematical programs in Python. *Mathematical Programming Computation*, vol. 3, no. 3, pp. 219–260 (2011)
  19. Andreas. Wächter, Lorenz. T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming, *Math. Programming*, vol 106, pp. 25-57 (2004)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Nature-inspired Bio-Mineral Refinery for Simultaneous Biofuel Feedstock production and CO<sub>2</sub> mineralization

Pavan Kumar Naraharisetti <sup>a, b, \*</sup>

<sup>a</sup> Newcastle University in Singapore, Faculty of Science, Agriculture & Engineering, Newcastle University, Newcastle Upon Tyne NE1 7RU, UK

<sup>b</sup> Institute of Chemical and Engineering Sciences, A\*STAR (Agency for Science, Technology and Research), 1 Pesek Road, Jurong Island, Singapore 627 833

\* Corresponding Author: [pavan-kumar.naraharisetti@ncl.ac.uk](mailto:pavan-kumar.naraharisetti@ncl.ac.uk)

## ABSTRACT

Inspired by Nature, we propose that synergies between biorefinery and mineral refinery can be exploited so that at least a part of the carbon is captured before being released to the atmosphere. In doing so, carbon is captured not only from CO<sub>2</sub> but also from biomass and developing more such processes may be the cornerstone for controlling CO<sub>2</sub> emissions. A comparison of circular economy in traditional biorefineries and bio-mineral refineries is done by using general chemical formulas and it is shown that the bio-mineral refinery captures carbon. In this work, we have shown that Serpentine may be used to partially neutralise biomass pyrolysis oil. The extracted oil may be used as feedstock to produce downstream chemicals and further studies are required to produce the same.

**Keywords:** Biofuels, Biomass, Carbon Capture, Carbon Dioxide Sequestration, Environment

## INTRODUCTION

Circular economy, a concept that focuses on recycle, is viewed as a potential means of reducing new resource usage, thus limiting CO<sub>2</sub> emissions. Applying the concept of circular economy to the chemical process industries; biomass, fossil resources, all the chemical products and CO<sub>2</sub> can be considered as the intermediates, while energy to the system comes from the Sun. Currently, the rate of CO<sub>2</sub> emissions is greater than the rate of biomass regeneration and this is leading to CO<sub>2</sub> accumulation in the atmosphere, and this is viewed as being linear. Furthermore, CO<sub>2</sub> utilization/mineralization to reduce atmospheric CO<sub>2</sub> accumulation may be viewed as the process of closing the loop in a circular economy; however, doing so is energy intensive and leads to further CO<sub>2</sub> emissions. Biorefineries have been proposed as a panacea since they can potentially substitute fossil sources, and the energy for CO<sub>2</sub> capture and production of biomass is taken directly from the sun. However, biorefineries have had limited success, and most are single product chemical plants.

The critical difference between fossil sources such as crude oil, natural gas and biomass is that biomass has about 40% (w/w) oxygen and the same is almost non-

existent in the fossil sources. In view of this, one can expect that biomass to chemicals/fuels is going to be expensive especially if oxygen elimination is necessary. Another critical difference is that fossil source is produced by natural interaction of dead biomass with the environment surrounding them and it is unknown how many tonnes of biomass produced one tonne of fossil source. On the contrary, biomass to chemicals process must have high productivity and the mass balance is stacked against bio-refineries. While this is challenging, our efforts must continue to focus on the use of biomass since it replaces fossil sources. Where possible, this skewed mass balance must be turned to our advantage by furthering our research. Further, the effect of the natural environment on biomass is not completely simulated in a biorefinery. In view of this, we try to simulate the interaction of biomass with the environment by combining biorefinery and mineral refinery. This inspiration comes from the fact that animals, which are complex chemical systems, take biomass as feed, produce complex (bio) chemicals within their body and leave mineral carbonates as bones at the end of their life cycle.

The objective of this work is thus to demonstrate that it is possible to combine biorefineries and mineral refineries such that synergies can be explored for the

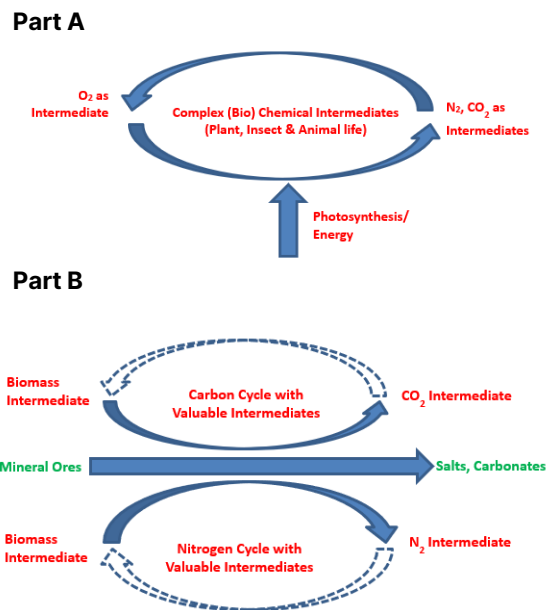
benefit of CO<sub>2</sub> mineralization. In view of this, we present details of two chemical processes where the synergies between biomass refining and mineral refining can be explored, and carbon is captured. They are the production of acetone and acetamide/acetonitrile. While the two processes are based on existing know-how, we also apply the concept to demonstrate that the field of “Bio-mineral refinery” can be developed to exploit the synergies while using raw biomass source. In view of this, we use biomass pyrolysis oil with Serpentine and the concept of Biomineral refinery is introduced by conducting preliminary experiments. In doing so, a multiproduct bio-mineral refinery is designed; carbon is captured both from biomass and CO<sub>2</sub> and good productivity achieved.

## CIRCULAR CHEMICAL ECONOMY WITH BIO-MINERAL REFINERIES

In the natural world, dead biomass degrades by bacterial action while producing organic acids that dissolve the surrounding mineral ores. The process produces mineral salts that serve as nutrients to plant life and plants support the insect and animal life. At the end of their life, the biomass from the insects, plants and animals degrades and the cycle continues. This entire process can be viewed as a circular (bio) chemical system with several recycle loops where the biomass, insects, animals, O<sub>2</sub> and CO<sub>2</sub> are considered as intermediates and energy is input to the system in the form of sunlight (Fig. 1A). Similarly, the processes in the chemical and allied industries can also be visualized where all the materials move in a loop with several recycles and energy is input to the system in the form of heat and electricity.

On further examination, we can see that the animal life leaves mineral carbonates at the end of their life cycle. Hence, the entire process can be viewed as a circular (bio) chemical system that is supported by a linear process where mineral ores are converted to mineral carbonates. We can see that there are synergies between the circular (bio) chemical system and linear mineral processes, and it is this synergy that we must exploit to develop new chemical processes. The overall process can be viewed as the carbon and nitrogen cycles interacting with the linear mineral processes (Fig. 1B). Here we present two processes in support of the Nature inspired ‘linear supported circular (bio) chemical system’. They are theoretically feasible and further research must be carried out to evaluate the productivity, energetics, and the economic feasibility. The first is a bio-mineral refinery based on the carbon cycle and the second is based on carbon and nitrogen cycles. Briefly, mineral ores such as Wollastonite and Serpentine are used as raw materials to produce mineral carbonates and other inorganic products. One of the common products of biomass degradation is acetic acid and we present on the use of palm

empty fruit bunch (EFB, C6H10.8O3.72) [1] derived acetic acid as one of the raw materials to produce organic intermediates such as acetonitrile and acetone. The materials thus produced can be treated as intermediates and these chemicals can be used as platform materials to produce downstream chemicals. In the proposed Circular Chemical Economy, we consider carbon capture both from biomass and from CO<sub>2</sub> thus limiting CO<sub>2</sub> emissions to the atmosphere.



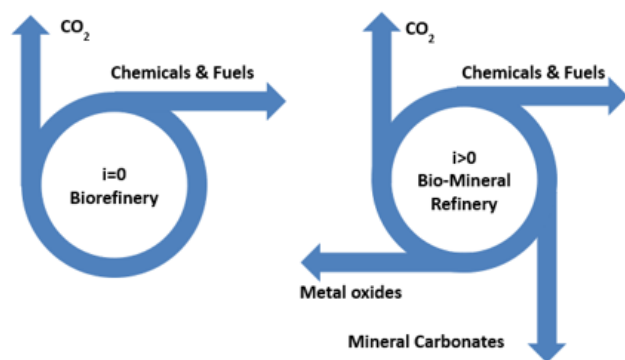
**Figure 1.** Part A: A traditional circular (bio) chemical system. Intermediate materials move in cycles with energy from the sun. Part B: A circular (bio) chemical system that is dependent on a linear mineral process. Dotted arrows represents natural processes.

## CO<sub>2</sub> MINERALIZATION POTENTIAL

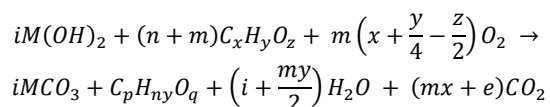
The reactions that may occur in a circular chemical economy to produce fuels and/or chemicals may be written by assuming a general formula for biomass. Given, ‘n’ is the number of moles of biomass used to produce the products and capture CO<sub>2</sub>, ‘i’ is the number of moles of CO<sub>2</sub> mineralized, ‘m’ is the number of moles of biomass combusted in order to provide heat to the process and ‘e’ is the CO<sub>2</sub> consumed/emitted by the process; where ‘i’ is a positive integer, ‘e’, ‘m’ and ‘n’ are real numbers; when i=0, the proposed bio-mineral refinery becomes a traditional biorefinery i.e., no carbon is captured. The same is depicted in Fig. 2 with reference to the circular economy. Biorefineries emit CO<sub>2</sub> either by way of energy uses in the process or as a byproduct and a sequestration process is necessary to minimize the CO<sub>2</sub> emissions. On the contrary, a part of the carbon is mineralized in Bio-mineral refineries and over a period of several cycles, the amount of carbon mineralized would be far greater than the

carbon emitted as CO<sub>2</sub> and thus, the process turns from being a CO<sub>2</sub> emitting process to carbon sequestration process.

We can see from Fig. 2. that carbon is mineralized in each cycle as mineral carbonates; hence the Bio-Mineral refineries become net CO<sub>2</sub>/carbon negative processes after a few cycles. The total CO<sub>2</sub> emissions (tonnes/tonne of CO<sub>2</sub> mineralized) can be calculated to be equal to  $(mx+e)/i$  and the number of cycles required for the process to become net carbon negative from being a CO<sub>2</sub> emitting process is also  $(mx+e)/i$ . We can conclude that 'e' and 'm' must be as small as possible and 'i' must be as large as possible so that maximum amount of carbon is mineralized while the input energy to the process is low. However, as 'm' and 'n' increases, the amount of valuable products produced per ton of CO<sub>2</sub> increases and this needs to be high to recover the cost of production. In view of this, detailed analysis on individual processes is necessary to understand the trade-off between the opposing objectives of maximizing the amount of CO<sub>2</sub> mineralized (with low chemical production) and maximizing the mass of valuable products produced (with low CO<sub>2</sub> mineralization).



**Figure 2.** A comparison of circular economy in traditional biorefineries ( $i=0$ ) and bio-mineral refineries ( $i>0$ ) is depicted here.



$$\text{where, } p = nx - i - e; q = nz - 2i - 2e$$

### Bio-Mineral refinery – Carbon Cycle

Acetic acid can be produced from biomass by (bio) chemical action. The chemical process has faster kinetics when compared to the bioprocess and may be suitable for large scale production. First, methanol is produced from biomass derived Syngas, next acetic acid is produced by the carbonylation of methanol (Monsanto and Cativa process) where the carbon monoxide may be produced from biomass derived carbon. Acetic acid can be

used for leaching Wollastonite and SiO<sub>2</sub> is precipitated [2] in the process. The salt that is produced, calcium acetate, can be separated and used to produce acetone while calcium carbonate is precipitated [3]. Acetone serves as a platform chemical [4] to produce downstream chemicals such as isopropyl alcohol, propane, solvents (Methyl isobutyl alcohol, methyl isobutyl ketone), Bisphenol-A (Polycarbonates, Polyurethanes, epoxy resins), Methyl methacrylate, among others. Acetone thus produced can potentially reduce the dependence on acetone from the Cumene process and hence some of the crude oil can be used in the production of other petro-chemicals and the overall use of crude oil reduced. Benzene may also be produced in a bio-mineral refinery to offset loss on production via Cumene process. Hydrogen from biomass derived Syngas may be used to meet downstream demand. The schematic of the processes is shown in Fig.3 and the detailed chemical reactions for Fig. 3. are given in Table 1. Briefly, Carbon either from biomass or from CO<sub>2</sub> is converted into mineral form thus removing some of it from the system. While such processes occur in geological time scales in the earth, chemical reactions in chemical plants are targeted to have faster chemical kinetics and it is hoped that the process of mineralization can be fastened.

The carbon and nitrogen cycles-based bio-mineral refinery produces acetamide/acetonitrile by using serpentine. These details, corresponding to Fig. 4., are only shown briefly due to limitation of space.

EFB is used to produce both Syngas which is later used to produce acetic acid. We can see from Table 1 that 1.32 moles of EFB successfully mineralizes 1.9 moles of CO<sub>2</sub> as CaCO<sub>3</sub> when Wollastonite is used. This is equivalent to sequestering 0.44 tons of CO<sub>2</sub>/tonne of EFB used while producing 0.59 tonnes of acetone and 0.61 tonnes of Silica that can generate income. Like the Wollastonite process, we observed that 1.32 moles of EFB successfully mineralizes 1.9 moles of CO<sub>2</sub> as MgCO<sub>3</sub> when Serpentine is used. The same amount of CO<sub>2</sub> is mineralized while producing 1.19 tons of acetamide. The amount of acetamide produced per tonne of EFB is higher compared to the amount of acetone produced. This is obvious because CO<sub>2</sub> is mineralized from biomass when acetone is produced, and flue gas is used as the CO<sub>2</sub> source when acetamide is produced. Here, we have assumed 100% conversion, and the actual values will be much different. In view of this, we take the basis as the amount of valuable product produced as it is straight forward and these values won't be dependent on the conversion, but only on reactions R8 (Fig. 3) or R12 (Fig. 4). When Wollastonite and Serpentine are used, 0.76 tonnes of CO<sub>2</sub>/tonne of acetone and 0.37 tonnes of CO<sub>2</sub>/tonne of acetamide (or 0.54 tonnes of CO<sub>2</sub>/tonne of acetonitrile) is mineralized respectively. We can see that the amount of CO<sub>2</sub> mineralized per tonne of valuable product is large and these

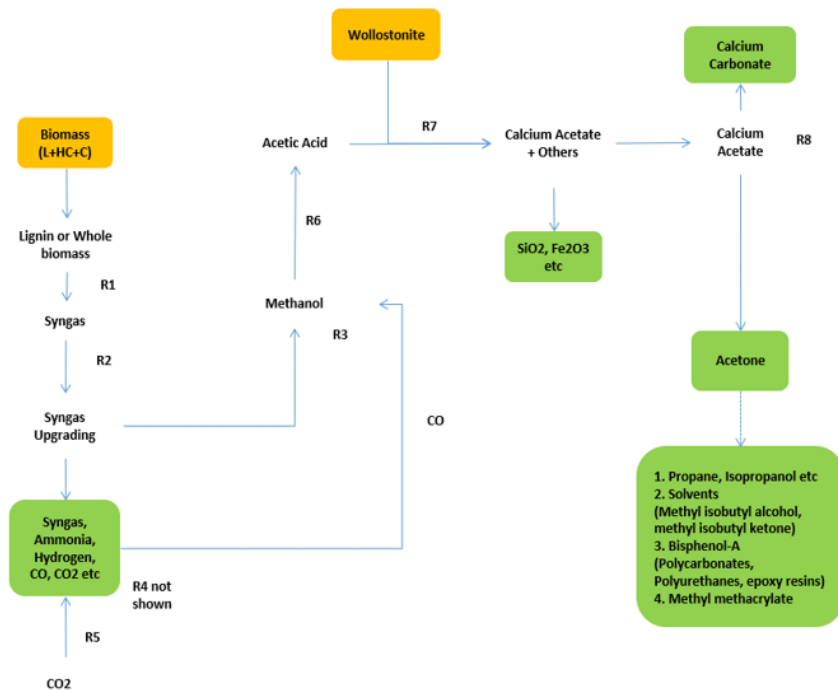


processes potentially can be used to produce valuable chemicals while simultaneously capturing CO<sub>2</sub>. The CO<sub>2</sub> emissions from the use of energy and the penalty thereof must be estimated separately and this is estimated from the values presented in Table 1. To calculate the energy requirements of the processes, we only consider the endothermic energy of the processes; any gains by heat integration associated with exothermic processes is neglected as these must be subjected to detailed heat integration calculations.

Based on Fig. 4. each cycle 6.95 tonnes or 8.78 tonnes of CO<sub>2</sub> is released into the atmosphere for every tonne of CO<sub>2</sub> mineralized (detailed reactions not shown). Here, the amount of CO<sub>2</sub> released to the atmosphere is large compared to the CO<sub>2</sub> mineralized, and hence the proposed concept of circular chemical economy through bio-mineral refineries holds. In view of this, over a period of 7 and 9 cycles the net CO<sub>2</sub> emitting processes become CO<sub>2</sub> neutral processes since the atmospheric CO<sub>2</sub> is effectively an intermediate that will be taken up in biomass

**Table 1:** Reactions in a Bio-Mineral refinery that uses palm EFB and Wollastonite as feed.  $\Delta H$  at a given temperature is calculated using HSC 6.1 software [5]. Calculations for Serpentine (Carbon-Nitrogen cycle) are not shown.

Use of Wollastonite as Feed & Production of Acetone		Rxn. T (°C)	$\Delta H$ (kJ)
Syngas Production (R1)	$C_6H_{10.8}O_{3.72} + 1.14O_2 \rightarrow 6CO + 5.4H_2$	700	2802
Syngas Upgrading (R2)	$6CO + 5.4H_2 + 2.2H_2O \rightarrow 3.8CO + 7.6H_2 + 2.2CO_2$	300	86
Methanol Production (R3)	$3.8CO + 7.6H_2 \rightarrow 3.8CH_3OH$	250	373
Carbon production (R4)	$0.32C_6H_{10.8}O_{3.72} + 0.532O_2 \rightarrow 1.9C + 1.71H_2O$	1300	563
CO production (R5)	$1.9C + 1.9CO_2 \rightarrow 3.8CO$	800	-323
Acetic Acid Production (R6)	$3.8CH_3OH + 3.8CO \rightarrow 3.8CH_3COOH$	200	620
Mineral Digestion (R7)	$1.9CaSiO_3 + 3.8CH_3COOH \rightarrow 1.9(CH_3COO)_2Ca + 1.9SiO_2 + 1.9H_2O$	50	209
Mineralisation (R8)	$1.9(CH_3COO)_2Ca + Heat \rightarrow 1.9CH_3COCH_3 + 1.9CaCO_3$	160	-183



**Figure 3:** Schematic of the process of producing Acetone from Biomass and Wollastonite. Acetone can be used a precursor to many other chemicals.

regeneration. After the 7 or 9 cycles the processes become purely CO<sub>2</sub> mineralization processes. A cycle is complete when the CO<sub>2</sub> emitted by the process is taken up in biomass regeneration. Hence, biomass doubling time can be considered as time for one cycle i.e., the cycle time is dependent on the biomass growth rate which in turn is dependent on the land area and the type of crop that is cultivated. Clearly, over the long term these processes have the potential to sequester large amounts of CO<sub>2</sub> and effectively de-carbonize the atmosphere. In contrast to the traditional biorefinery where fuels and/or chemicals are produced and no CO<sub>2</sub> is captured, a bio-mineral refinery captures at least some amount of carbon as mineral carbonates and partially prevents atmospheric CO<sub>2</sub> accumulation. This contrast is the key to the successful development of circular chemical economy because the focus is on carbonate accumulation and not on CO<sub>2</sub> accumulation in the atmosphere. EFB is produced at the rate of 18 to 30 tonnes/hectare/year [6]. Considering the lower value of 18 tonnes/hectare/year, each tonne of CO<sub>2</sub> mineralized requires about 0.21 hectares or 0.26 hectares of land for CO<sub>2</sub> recycle and biomass regeneration. However, the land required to produce biomass that is used as feedstock is larger at 0.33 hectares or 0.38 hectares per tonne of CO<sub>2</sub> mineralized since all the carbon in the biomass is not released to the atmosphere, but carbonates and valuable products are produced. Thus, we have demonstrated through two examples that there is scope for developing circular chemical economy with carbon mitigation and the development of other processes to produce valuable products must be explored.

While the proposed routes are interesting, the apparent question is how fast each loop in the circular economy is. To answer such a question, detailed kinetics studies are necessary. While it suffices to say here that since CO<sub>2</sub> of about 10 % (based on less than 10 cycles) is mineralized and that reactions in chemical industries are typically faster in the matter of minutes to hours, we also need to answer this in terms of the amount of mineralization with reference to the mass of products produced. The same was presented earlier as 0.44 tonnes of CO<sub>2</sub>/tonne of EFB is mineralized for each of the two processes. While this is theoretical conversion, the efficiency related to energy penalty in such processes are expected to be 50% [7, 8] and may be appropriate to assume that 0.22 tonnes of CO<sub>2</sub> are mineralized per tonne of EFB after scale-up.

Although more such examples exist in literature where biomass derived chemicals can be used for CO<sub>2</sub> mineralization and the production of valuable chemicals in a bio-mineral refinery, the key for successful implementation of the processes in a commercial scale is to use whole biomass so that the number of processing steps and the associated costs are reduced. The possibility that whole biomass, which has a large amount of

oxygen, be de-oxygenated by Ketonic decarboxylation must be explored so that valuable hydrocarbons and inorganic products are produced, and carbon is mineralized. Biomass has phenols and other aromatic compounds, and it may (not) be possible to produce higher hydrocarbons which have a high C/O ratio using biomass as feedstock. Considering the variety of compounds in biomass, it would be rather difficult to predict the type of products that can be produced in hydrothermal/thermochemical treatment of biomass & mineral ore mixtures. Further, addition of acid for acid hydrolysis of biomass and/or leaching of the mineral ore and production of salts thereof would lead to increased complexity. In view of this, experiments to study the products and the mechanism of product formation in biomass-mineral interactions are warranted for the development of bio-mineral refinery based circular chemical economy that can at least partially replace fossil fuel based chemical economy. Hence, we have conducted experiments using biomass pyrolysis oil and serpentine mineral and present the same here.

**Table 2:** CO<sub>2</sub> capture potential of the proposed processes. W-wollastonite; S-serpentine.

	<b>W</b>	<b>S</b>
Endothermic energy required (MJ)	4.65	5.91
Biomass for energy supply (80% energy efficiency, kg)	0.31	0.39
CO <sub>2</sub> emissions from Biomass combustion (kg)	0.57	0.72
Total Biomass required (tons/ton CO <sub>2</sub> mineralized)	5.92	6.91
Total CO <sub>2</sub> emissions (kg)	0.58	0.73
CO <sub>2</sub> emissions (tonnes/tonne CO <sub>2</sub> mineralized)	6.95	8.78
Products produced (tonnes/tonne CO <sub>2</sub> mineralized)	1.32 (acetone)	2.68 (acetamide)
Number of Cycles required for the process to become net CO <sub>2</sub> capture process	7	9
Land area for feedstock biomass production (hectares/tonne of CO <sub>2</sub> mineralized)	0.33	0.38

## CAPTURING CARBON FROM BIOMASS

### Preliminary Experiments

In this work, we need to first establish if bio-oil would react with Serpentine. In view of this, 20 grams of serpentine were mixed with 260 ml of water-soluble bio-oil (60 ml water soluble component from 100 ml pure bio-oil and 200 ml water) and stirred overnight. Then the solid samples, serpentine that gets disassociated, were collected, and analyzed. We were able to collect 12.8 grams

of suspended solids (fine solids), 9.5 grams of settled solids (coarse) and 0.75 grams of magnetic material that sticks to the magnetic bar. The samples were left to air dry for three days inside a fume-hood and weighed again. The measured weights are as follows: 10.1 grams of suspended solids (fine solids), 9.3 grams of settled solids (coarse) and 0.69 grams of magnetic material. It was observed that there was a pH change from 1.0 to 4.3 due to the addition of Serpentine.

Separately, 100 ml of biomass pyrolysis oil is mixed with 200 ml of water. The water-soluble mixture is decanted (260 ml) and 10 g of serpentine is added to it and stirred for 4 hrs. A sample was collected for analysis by GC-MC. This mixture is heated in a Parr reactor at 175°C for 4 hrs. at a maximum CO<sub>2</sub> pressure of 20 barg. The Parr reactor is left to cool over-night before opening. The contents are allowed to settle at RTP for 2 hrs. after opening and before filtration to allow for the dissolved CO<sub>2</sub> to escape. 23 g of cake is collected, and this is allowed to air-dry in a fume hood for two days after which the weight was observed to be 16 g. Elemental Analysis and TGA

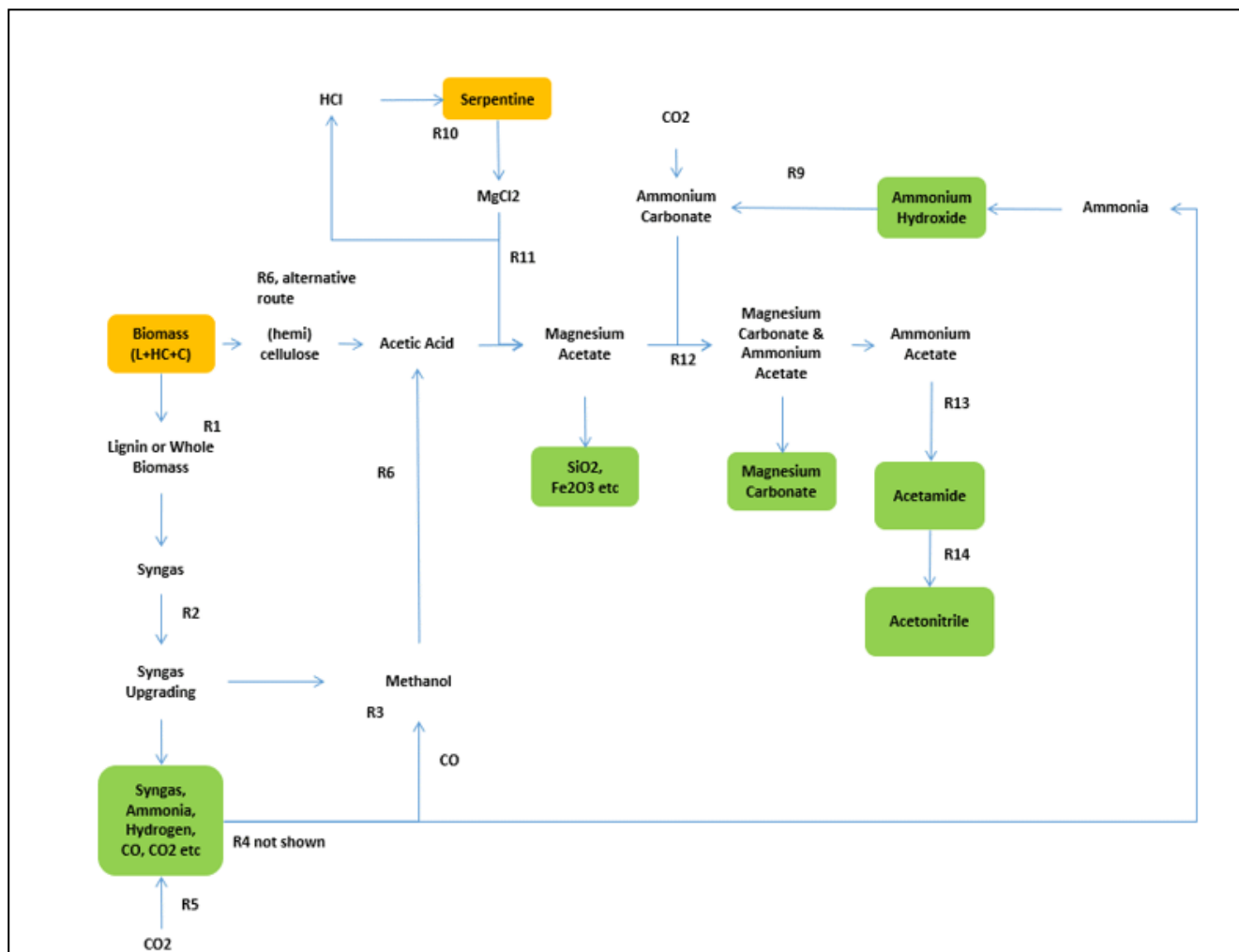
were conducted on this solid residue. A liquid sample is collected for analysis by GC-MS (results not presented).

### Preliminary Results

TGA has shown that there is a 44% weight loss when heated up to 800°C; hence we can conclude that 56% (8.96 g) of the material is inorganic oxides and this is reasonable since serpentine (10 g) is expected to have about 10% water of hydration and this would be lost during calcination at high temperature [9]. The rest of the mass (7.04 g) can be expected to be organic matter. The elemental analysis of the material indicated 41.22% carbon, 0.32% nitrogen, 4.11% hydrogen and 0.21% sulphur, leading to a total weight of 101.8% and this excess can be expected to be due to instrument error.

### Alternative uses

One of the uses for the residual water insoluble solids is the production of construction material. With respect to carbon capture, the precipitate so produced can then be compressed into bricks and stored as a material



**Figure 4:** Schematic of the process of producing Acetamide/Acetonitrile from Biomass and Serpentine.

that effectively captured carbon. The solid material was mixed with water and shaken for 15 min and allowed to settle overnight. It was observed that the solid and water separated indicating that the material is largely insoluble in water. The advantage of this process is that the precipitate holds more carbon than if the mineral serpentine were to be reacted with CO<sub>2</sub> and hence this process would require mining of fewer serpentine. In general, a gram of serpentine can hold up to a third of a gram of CO<sub>2</sub> to produce its mineral carbonate [10] and this is equivalent to about a tenth of a gram of carbon. In contrast, the precipitate we produced holds about 0.66 g of carbon per gram of serpentine used. In view of this, it is worthwhile to explore two options; the first is to hydrogenate these residual water-soluble organics to produce a fuel. Preliminary results of hydrogenation of the mixture without the use of an additional catalyst were performed at 10 barg initial H<sub>2</sub> pressure and a temperature of 175°C in a Parr reactor. The Parr reactor was allowed to cool overnight, and the hydrogenation is repeated at the same conditions and this process is repeated on 4 consecutive days to test the catalytic effect of the mineral; the resultant liquid samples were analyzed by GC-MS. It was observed that the composition of the product is rather different compared the starting material indicating that serpentine has some catalytic effect, but the process did not produce any fuel or water insoluble materials (results not presented) and it may be necessary to use a commercial catalyst at a higher pressure and/or temperature to produce a fuel. The second option is to use this water-soluble bio-oil as a substrate in a biochemical reactor to produce higher value chemicals in view of the current low crude oil price. Preliminary results have shown that the treated bio-oil is not bactericidal unlike the untreated pyrolysis oil. Notwithstanding the economics, we are of the opinion that this concept must be explored further with a view of developing more Bio-Mineral refineries to mitigate climate change and hence we suggest other possible routes that could be explored.

### Other Possible Routes

We present more examples where biomass derived chemicals are used to produce valuable products through decarboxylation. Propionic acid [11], Levulinic acid and Adipic Acid [12], Benzoic Acid [13], Butanol [14] and Butyric Acid [15] can be derived from biomass and used in bio-mineral processes. Briefly, Cyclopentanone and Cyclopentene from Adipic Acid [16], 3-pentanone from Propionic Acid [17], Acetic acid or acetone and a mixture of Cyclopentenones from Levulinic Acid [18], Benzene from Benzoic Acid [19], 5-nonanone from Pentanoic Acid [20, 21] and 4-heptanone from Butanol via Butyric Acid [22] are other examples of bio-mineral processes where different metal oxides are used for decarboxylation. A list of possibilities using two different chemicals in a single

Ketonic decarboxylation reaction can be found elsewhere [23]. The ketones so produced can be hydrogenated and dehydrated to produce olefins [24]. Considering that it is possible to capture carbon as a by-product of a chemical process, it may be possible to reengineer the crude oil, biorefinery and mineral refinery supply chains by the development of Bio-Mineral refineries that capture carbon and control climate change. Hence, it may be worthwhile to pursue research in this direction for the betterment of the planet.

## CONCLUSIONS

We have proposed a circular chemical economy to produce silica, acetone and acetamide/acetonitrile while capturing carbon as mineral carbonates. The proposed concept relies on the notion that valuable chemicals can be produced in a circular economy while synergies between circular bio-economy and linear mineral economy are exploited. Details of two examples of multi-product bio-mineral refineries are given to demonstrate the concept. Such refineries can capture some of the carbon before it is released into the atmosphere, while also capturing CO<sub>2</sub> from flue gas. We have shown that the proposed processes turn from being net CO<sub>2</sub> emitters to being CO<sub>2</sub> neutral processes in 7 or 9 cycles, after which they can be treated as CO<sub>2</sub> mineralization processes with the production of valuable chemicals as by-products. While the proposed processes are technically feasible, it is of importance to optimize them and identify more such processes, keeping in mind the large number of chemical products that potentially need to be replaced with alternative processes to mitigate climate change. Further, it may be worthwhile to study the use of biomass along with mineral ores and reduce the number of processing steps which have a direct impact on the overall cost of production. Hence, we have conducted experiments using biomass pyrolysis oil and serpentine and these show promise for further studies. In view of this, detailed studies are necessary to further develop this research area.

## ACKNOWLEDGEMENTS

Financial support for this work under “ICES/16-230A01: Bio-mineral refineries: Biomass pre-treatment and capturing carbon from biomass/CO<sub>2</sub>” is gratefully acknowledged. The author would like to thank Prof. Paul Sharratt and Dr. Veeresh Juturu for valuable insights.

## REFERENCES

1. Lee, D. H., Yang, H., Yan, R., Liang, D. T., 2007. Prediction of gaseous products from biomass pyrolysis through combined kinetic and thermodynamic simulations. *Fuel*, 86, 410-417.

2. Ghoorah, M., Dlugogorski, B. Z., Balucan, R. D., Kennedy, E. M., 2014. Selection of acid for weak acid processing of Wollastonite for mineralisation of CO<sub>2</sub>. *Fuel*, 122, 277–286.
3. Mora, F. M., 1957. Continuous process for producing acetone using calcium acetate liquor as starting material. US Patent Office, 2794848.
4. Werpy, T., Petersen, G., Eds. Top Value Added Chemicals from Biomass Volume I—Results of Screening for Potential, Candidates from Sugars and Synthesis Gas. U.S. Department of Energy, 2004.
5. HSC. 6.1 ed: Outotec Oy.
6. Wilmar International. <https://www.wilmar-international.com/our-businesses/plantation/oil-palm-plantation-milling> (last accessed 18 Oct 2023)
7. Naraharisetti, P. K., Yeo, T. Y., Jie, B., 2017. Factors influencing CO<sub>2</sub> and energy penalties of CO<sub>2</sub> mineralization processes. *ChemPhysChem*, Special Issue: CO<sub>2</sub> utilization, 18, 22, 3189–3202.
8. Naraharisetti, P. K., Yeo, T. Y., Jie, B., 2019. New classification of CO<sub>2</sub> mineralization processes and economic evaluation. *Renewable and Sustainable Energy Reviews*, 99, 220–233.
9. Balucan, R. D., Dlugogorski, B. Z., 2012 Thermal activation of antigorite for mineralization of CO<sub>2</sub>. *Environmental Science & Technology*, 47, 182–90.
10. Gerdemann, S. J., O'Connor, W. K., Dahlin, D. C., Penner, L. R., Rush, H., 2007. Ex situ aqueous mineral carbonation. *Environmental Science & Technology*, 41, 2587–93.
11. Ekman, A., Börjesson, P., 2011. Environmental assessment of propionic acid produced in an agricultural biomass-based biorefinery system. *Journal of Cleaner Production*, 19 (11), 1257–1265.
12. Nobbs, J. D., Zainal, N. Z. B., Tan, J., Drent, E., Stubbs, L. P., Li, C., Lim, S. C. Y., Kumbang, D. G. A., van Meurs, M., 2016. Bio-based Pentenoic Acids as Intermediates to Higher Value- Added Mono- and Dicarboxylic Acids. *ChemistrySelect*, 1, 539–544.
13. Arceo, E., Ellman, J. A., Bergman, R. G., 2010. A Direct, Biomass-Based Synthesis of Benzoic Acid: Formic Acid-Mediated Deoxygenation of the Glucose-Derived Materials Quinic Acid and Shikimic Acid. *ChemSusChem*, 3, 811 – 813
14. Chukwuemeka Ezejji, T. C., Nasib Qureshi, N., Blaschek, H. P., 2007. Bioproduction of butanol from biomass: from genes to bioreactors. *Current Opinion in Biotechnology*, 18, 3, 220–227.
15. Liu, S., Bischoff, M., Leathers, T. D., Qureshi, N., Rich, J. O., Hughes, S. R., 2013. Butyric acid from anaerobic fermentation of lignocellulosic biomass hydrolysates by *Clostridium tyrobutyricum* strain RPT-4213. *Bioresource Technology*, 143, 322–329.
16. Thorpe, J. F., Kon, G. A. R., 1925. *Cyclopentanone. Org. Synth.*, 5, 37
17. Siegel, H., Eggersdorfer, M., 1990. *Ullmann's Encyclopedia of Industrial Chemistry*. Ed.: Gerhartz, W., Weinheim, VCH, 83–95.
18. Schwartz, T. J., van Heiningen, A. R. P., Wheeler, M. C., 2010. Energy densification of levulinic acid by thermal deoxygenation. *Green Chem.*, 12, 1353–1356
19. Faraday, M., 1825. On new compounds of carbon and hydrogen, and on certain other products obtained during the decomposition of oil by heat. *Philosophical Transactions of the Royal Society* 115, 440–466.
20. Gliński, M., Kijeński, J., 2000. Catalytic Ketonization of Carboxylic Acids Synthesis of Saturated and Unsaturated Ketones. *Reaction Kinetics and Catalysis Letters*, 69, 1, 123–128.
21. Khromova, S. A., Smirnov, A. A., Selishcheva, S. A., Kukushkin, R. G., Dundich, V. O., Trusov, L. I., Yakovlev, V. A., 2013. Magnesium-containing catalysts for the decarboxylation of bio-oil. *Catalysis in Industry*, 5, 3, 260–268.
22. Plint, N., Ghavalas, D., Vally, T., Sokolovski, V.D., Coville, N.J., 1999. Catalysed synthesis of 4-heptanone from 1-butanol. *Catalysis Today*, 49, 1–3, 71–77.
23. Renz, M., 2005. Ketonization of Carboxylic Acids by Decarboxylation: Mechanism and Scope. *Eur. J. Org. Chem.*, 979–988
24. Corma, A., Renz, M., Schaverien, C., 2008. Coupling Fatty Acids by Ketonic Decarboxylation Using Solid Catalysts for the Direct Production of Diesel, Lubricants, and Chemicals. *ChemSusChem*, 1, 739 – 741.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Deciphering the Policy-Technology Nexus: Enabling Effective and Transparent Carbon Capture Utilization and Storage Supply Chains

Manar Y. Oqbi<sup>a,\*</sup>, Dhabia M. Al-Mohannadi<sup>b</sup>

<sup>a</sup> Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX, 77843, USA

<sup>b</sup> Department of Chemical Engineering, Texas A&M University at Qatar, P.O Box 23874, Education City, Doha, Qatar

\* Corresponding Authors: [manar.oqbi@tamu.edu](mailto:manar.oqbi@tamu.edu).

---

## ABSTRACT

In response to the global imperative to address climate change, this research focuses on enhancing the transparency and efficiency of the Carbon Capture Utilization and Storage (CCUS) supply chain under carbon tax. We propose a decision-making framework that integrates the CCUS supply chain's optimization model, emphasizing carbon tax policies, with a blockchain network. Smart contracts play a pivotal role in automating the exchange and utilization of carbon emissions, enhancing the digitalization of the CCUS supply chain from source to sink. This automation facilitates seamless matching of carbon sources with sinks, efficient transfer of emissions and funds besides record-keeping of transactions. Consequently, it improves the monitoring, reporting, and verification processes within the CCUS framework, thereby simplifying compliance with regulatory mandates for net emission reductions and carbon taxation policies. By eliminating reliance on third-party verifiers, our blockchain-based CCUS system reduces verification costs and ensures reliable tracking of emissions, mitigating the risk of carbon leakage. Policymakers and stakeholders gain valuable insights to optimize the CCUS network design, specifically considering the impact of carbon tax. This study represents an advancement in sustainable practices, providing a robust tool for decision-makers engaged in climate change mitigation efforts.

---

**Keywords:** Carbon Dioxide, Carbon Capture, Carbon Dioxide Sequestration, Carbon Capture Utilization and Storage (CCUS), Supply Chain, Optimization, Carbon Reduction Policies, Carbon Tax, Blockchain, digitalization

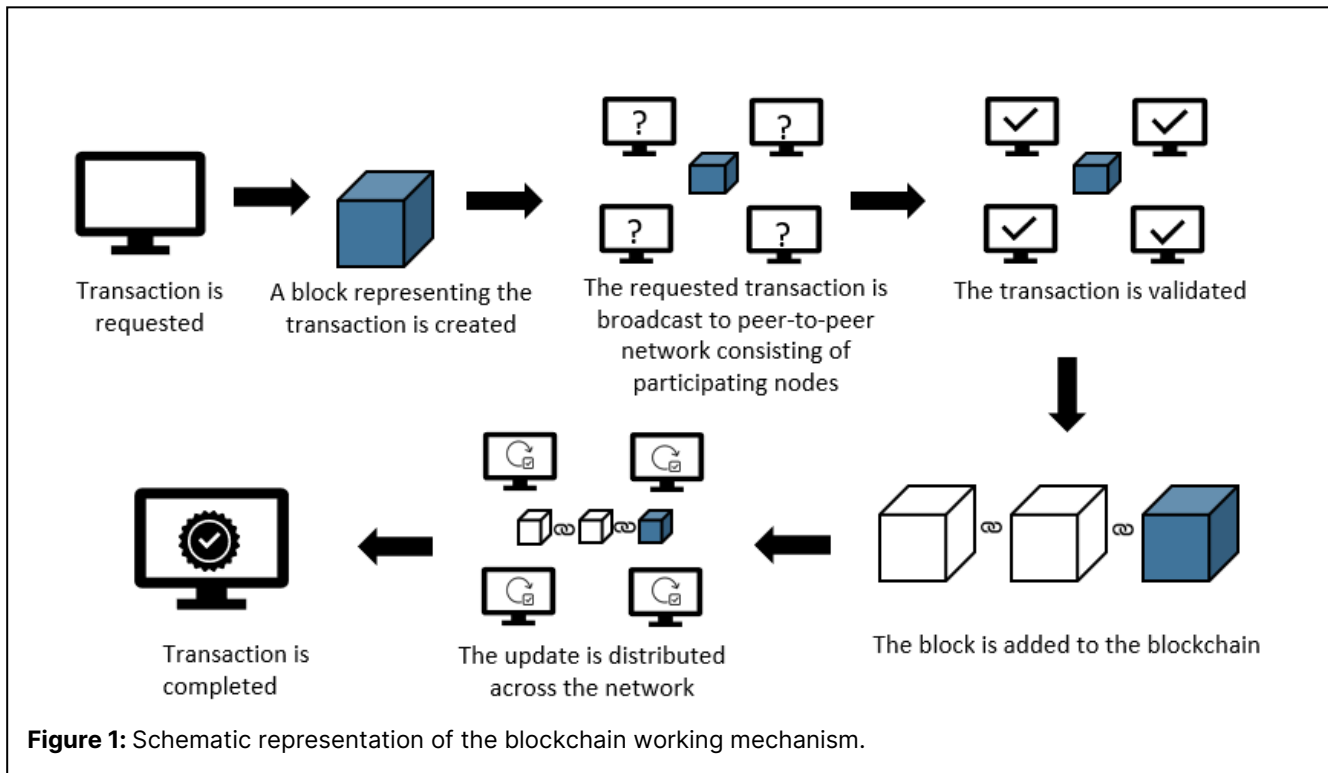
## INTRODUCTION

Climate change arises as an escalating global crisis, primarily fueled by soaring levels of greenhouse gas emissions, notably carbon dioxide. Alarming statistics from 2019 reveal a 54 % surge in net anthropogenic greenhouse gas emissions since 1990, underscoring the urgency of transformative solutions [1]. Carbon Capture Utilization and Storage (CCUS) emerges as a pivotal intervention. Projections from the International Energy Agency [2] highlight the growing significance of CCUS, foreseeing a 12 % cumulative reduction in emissions by 2050, particularly in sectors like cement, steel, and chemicals.

The intersection of CCUS with carbon tax policies introduces a nexus laden with challenges and promises. Precise measurement and verification of captured and

stored carbon dioxide pose intricate hurdles in a carbon tax regime. Challenges like low capture efficiency, operational costs, and the lack of international cooperation inhibit the holistic development of CCUS [3,4]. Carbon taxes, wielding economic influence, are instrumental in steering industries toward sustainable practices. They not only curtail energy consumption but also propel advancements in emission reduction technologies [5].

However, the amalgamation of carbon taxes and CCUS confronts multifaceted obstacles encompassing cost implications, technology maturity, and governance complexities. The efficacy of carbon taxes hinges on meticulous governance mechanisms to ensure precision in measurement, verification, and transparency, necessitating innovative solutions. Blockchain technology emerges as a disruptive force, offering a decentralized, transparent, and tamper-proof solution to address these



challenges and fortify the CCUS-carbon tax framework.

Blockchain, with its immutable ledger system, stands poised to support the monitoring and verification of carbon emissions and cash flows. Its decentralized nature instills trust in reported data, mitigates fraud, and fosters international collaboration in the pursuit of climate change mitigation. This paper proposes a blockchain-based approach for CCUS-carbon tax nexus, propelling us toward a sustainable future. As we explore the challenging landscape of environmental responsibility and technological advancement, this study reveals a promising path where CCUS-carbon tax supply chains, and blockchain technology intersect. The subsequent sections delve into overview of blockchain technology, provide a summary of existing studies in the literature, elucidate the proposed framework, explore current challenges, and conclude with insights into future directions.

## BLOCKCHAIN TECHNOLOGY OVERVIEW

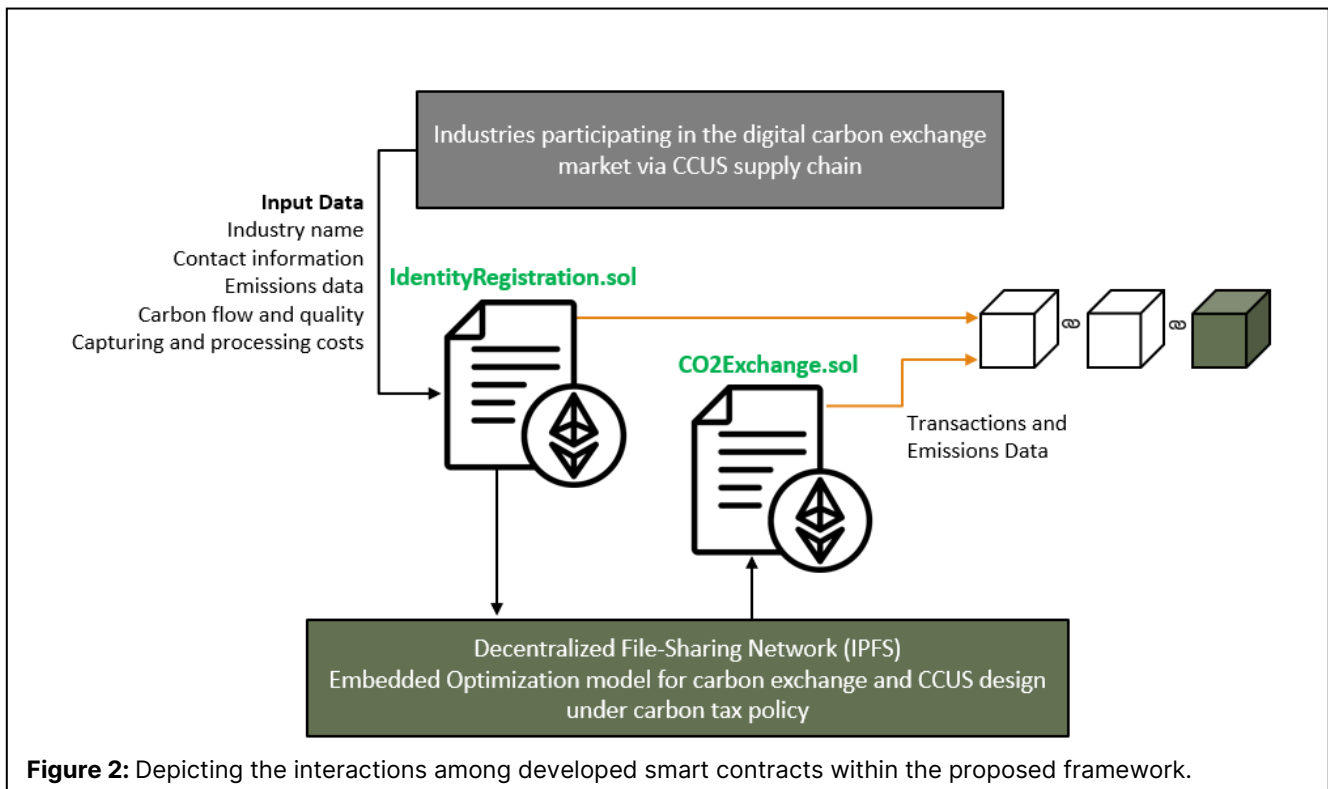
Blockchain is a decentralized technology for recording and managing data and transactions in a transparent, secure and tamper-proof manner [6]. The emerging technology is a distributed ledger that ensures integrity and immutability of data using cryptography and via the replication and allocation of transactions across a network of computers [7]. Blockchain technology has the potential to eliminate security vulnerabilities, eradicate fraudulent activities, and establish an unprecedented level of transparency obviating the necessity for a reliable third-party entity [8]. Blockchain technology is used

to manage tangible and intangible assets using a chain of blocks each containing a set of digital transactions via a decentralized and peer-to-peer network [9].

The decentralization characteristic disperses authority among network nodes, eradicating the need for central oversight. Integral cryptographic techniques bolster data security and integrity [9]. The consensus mechanism arranges validation and transaction sequencing, often via Proof of Work (PoW) or Proof of Stake (PoS). This distributed ledger is bolstered by redundancy across network nodes, thwarting single points of failure. Smart contracts enable self-executing agreements. Blockchains can be public or private, and they imbue data with immutability, making alterations difficult. Furthermore, their permissioned or permissionless nature adds a layer of flexibility to their applications. These multifaceted features collectively empower blockchain's versatility in diverse domains, from cryptocurrencies to supply chain management and beyond [8]. Figure 1 represents blockchain working mechanism, showcasing decentralized, secure, and transparent data sharing among validating nodes, ensuring integrity and traceability in information recording and verification.

## LITERATURE REVIEW

CCUS plays a pivotal role in global climate change mitigation, particularly in challenging industries. Recent advancements emphasize trade-offs between economics and environmental impact, showcasing the benefits of CCUS in achieving net-zero energy systems [10].



**Figure 2:** Depicting the interactions among developed smart contracts within the proposed framework.

Another study focused on optimal cost and environmental benefits of deploying CCUS supply chains at scale in Guangdong Province, China [11]. Some researchers evaluated comprehensive benefits of CCUS networks in coal-fired power plants [12], while others optimized CCUS deployment considering carbon neutrality, cost efficiency, and water stress [13]. These studies highlight evolving CCUS models, addressing economic viability, environmental impact, and sector-specific applications.

Carbon tax is pivotal in incentivizing CCUS adoption. The impact on CCUS source-sink matching in China was analyzed [14], and a stepwise deployment strategies was proposed in Canada based on varying carbon tax levels [15]. CCUS systems in Italy and Germany were explored considering carbon tax in their comprehensive objective function [16]. Other work delved into a risk management framework under carbon tax uncertainties [17], and CO<sub>2</sub> storage potential in shale reservoirs was assessed factoring in economic viability and carbon tax implications [18]. These studies emphasize the role of carbon taxes in shaping CCUS strategies.

Blockchain integration transforms CCUS governance. A blockchain-based carbon trading mechanism was proposed [19], and a permissioned blockchain for emissions trading was modeled [20]. A Sovereign blockchain for carbon trading was explored [21] and a token-based economy for carbon trading was suggested [22]. The intersection of blockchain, carbon trading, and CCUS governance offers a promising avenue for transparency

and efficiency in climate change mitigation.

Existing literature focuses largely on CCUS modeling, optimization, and the impact of carbon taxes, leaving a research gap in exploring blockchain's potential beyond carbon trading in the CCUS landscape. This study addresses this gap by proposing an approach for integrating blockchain into CCUS supply chain optimization models under carbon tax influence. By harnessing blockchain's transparency and tamper-proof nature, we aim to enhance data accuracy, prevent fraud, and foster global collaboration for effective climate change mitigation. This integration offers a comprehensive solution to key challenges in CCUS governance, contributing to a sustainable future.

## PROPOSED APPROACH

Transparency of the CCUS supply chains necessitates the inclusion of emissions and cost data from various stages, encompassing CO<sub>2</sub> capture, compression, utilization, and geological storage. To enhance transparency in designing CCUS supply chains with consideration of carbon tax, this study provides a framework integrating blockchain technology. This integration involves embedding the CCUS-carbon tax optimization model with the blockchain network, aiming for an optimal CCUS network design that fosters improved transparency in emissions reduction, tracking, and cost management. The integration ensures the recording of emissions and cost data, facilitating transparent emission reduction

verification and limiting carbon leakage. The proposed framework empowers CCUS network optimization under carbon tax, fostering a distributed ledger containing transparent CCUS emissions and cost data across various stages.

In this study, we developed two smart contracts utilizing Solidity, the primary programming language for crafting smart contracts on the Ethereum blockchain platform. These contracts, named IdentityRegistration.sol and CO2Exchange.sol, were meticulously designed to support our research framework. To verify their integrity and operational effectiveness, both smart contracts underwent a comprehensive process of compilation, testing, and deployment. This rigorous validation process ensured that they met our stringent requirements for security, efficiency, and reliability, making them integral components of our blockchain-based CCUS system.

The IdentityRegistration.sol smart contract constitutes the foundational layer of our blockchain-based system. This pivotal contract is designed to facilitate a range of critical activities, including the registration of new industries and verification entities. It enables the enrollment of carbon sources, sinks, and verifiers by capturing essential details such as name, industry, address, contact information, email address, industry identification number (ID), verifier ID, carbon flow characteristics, quality parameters, physical properties, and associated costs, including treatment expenses for sources and processing charges for sinks, alongside data on secondary emissions. This smart contract incorporates functions that allow for the retrieval of information about registered entities, enhancing transparency and accessibility within the system. A significant feature of this contract is its built-in mechanism to prevent duplicate registrations, thereby ensuring the integrity of the registration process. It achieves this by conducting preliminary checks against existing records before finalizing any new registrations, thus maintaining an accurate and reliable database of participants in the CCUS ecosystem.

The CO2Exchange.sol smart contract is a critical component of our blockchain-based framework for managing the CCUS supply chain under carbon tax, facilitating the effective exchange and utilization of captured carbon dioxide. This contract is designed to execute several key functions crucial for the transparency and efficiency of the CCUS process. It is responsible for accurately documenting carbon emissions at each stage of the CCUS supply chain, including emission sources, treatment units, transportation mechanisms, and storage or utilization sinks. Additionally, it enables the seamless transfer of funds among carbon sources, sinks, and the government, thereby ensuring compliance with carbon tax regulations and supporting the economic aspects of carbon trading by managing the flow of funds within the

system. This includes the recording of all related transactions, providing a transparent and immutable ledger of carbon emissions and financial exchanges.

**Table 1:** Algorithms of Smart Contracts.

---

**IdentityRegistration.sol: Initializing CO<sub>2</sub> utilization and exchange**

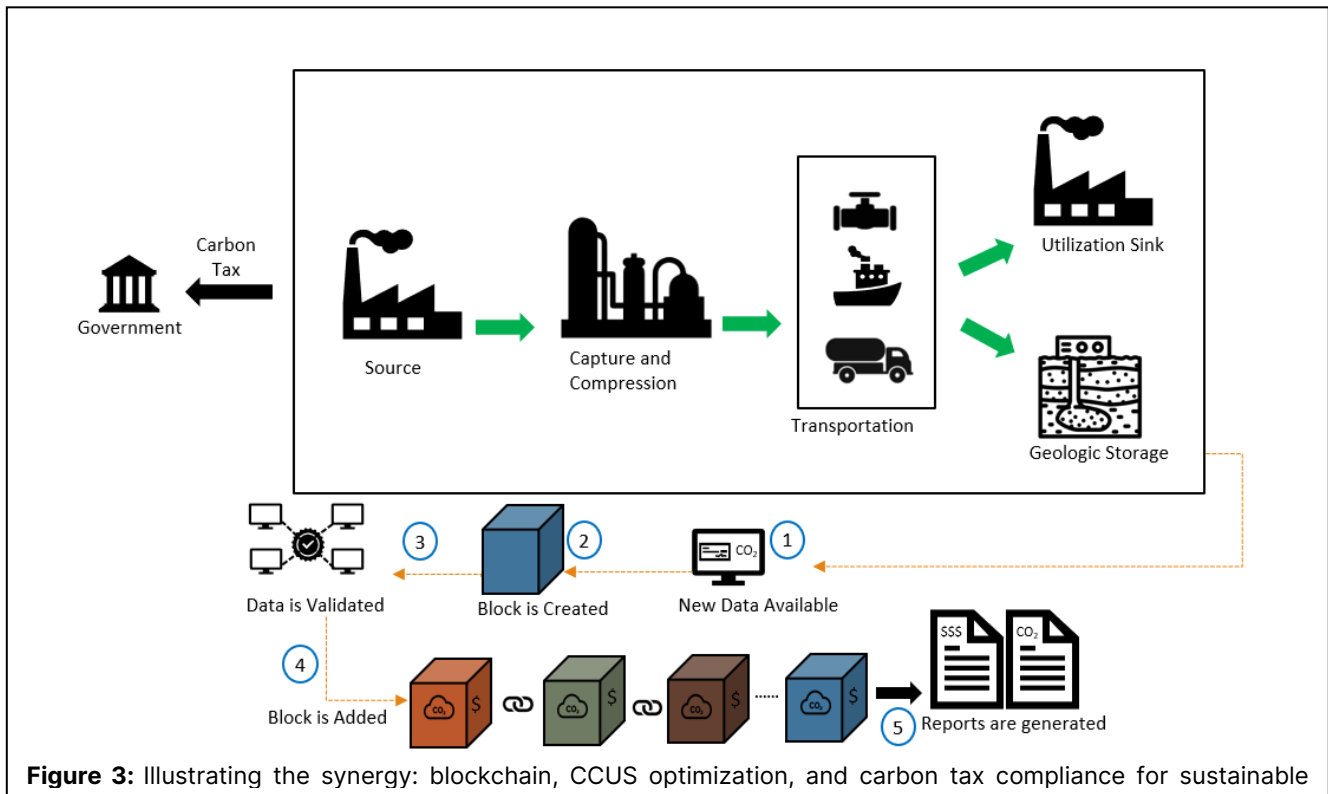
1. Declare a contract called IdentityRegistration.
2. Define structs for Sources, and Sinks considering relevant information such as industry name, type, address, contact information, email, industry ID, carbon flow, and quality.
3. Define struct for Verifiers including relevant information such as verifier name, industry, contact information, email, and verifier ID.
4. Define mapping variables to map industry ID to Source/Sink structs.
5. Define mapping variables to map verifier ID to verifier struct.
6. Define events for NewSource, NewSink, and NewVerifier.
7. Define functions to register new industries and verifiers and retrieve their information.

**CO2Exchange.sol: Transaction and tokenization process**

1. Declare a contract called CO2Exchange.
  2. Define struct for CO<sub>2</sub> transaction.
  3. Define mapping variables to map transaction ID to the transaction struct.
  4. Define function to record emissions from each stage including sources, capturing units, transportation, and sinks.
  5. Define function to transfer funds from sinks to sources and from sources to government.
  6. Define function to record transactions.
- 

The operational logic and step-by-step functionality of these smart contracts are detailed in Table 1, providing a clear schematic of their roles within the CCUS system. Figure 2 visually depicts the dynamic interactions facilitated by the developed smart contracts, illustrating how they interconnect within the blockchain-enabled CCUS supply chain under carbon taxation. This representation underscores the smart contracts' pivotal role in streamlining the CCUS process to form a cohesive system that not only tracks carbon emissions and transactions but also facilitates the regulatory and economic mechanisms essential for effective carbon management and mitigation efforts.

The commencement of the CCUS-blockchain system takes place upon the availability of emissions and cost data, resulting in the creation of a new block. This block then undergoes a secure and decentralized distribution process among validating and authorizing nodes, with the assumption that both government entities and



**Figure 3:** Illustrating the synergy: blockchain, CCUS optimization, and carbon tax compliance for sustainable

participating industries hold authorized access. The confirmed data blocks are smoothly added to the ongoing sequence by network nodes, safeguarded by distinctive hashes and the hash of the preceding block. Emissions and cost data from each stage are documented within authenticated and encrypted blocks, assuring the preservation of data integrity across the entire blockchain network. Figure 3 demonstrates the proposed approach, integrating CCUS-carbon tax optimization models with a blockchain network for enhanced transparency in emission and cash flow tracking.

Incorporating blockchain technology within the CCUS carbon tax nexus presents a promising avenue for enhancing data integrity, transparency, and stakeholder trust. However, it introduces certain risks and drawbacks that necessitate careful consideration for successful implementation. The adoption of blockchain technology significantly increases the operational complexity of the CCUS supply chain. Establishing a blockchain infrastructure requires substantial initial investments not only in technology but also in training personnel and developing new operational protocols. This upfront cost and the effort needed for integration can deter stakeholders, particularly in regions where CCUS technologies are still in nascent stages. The complexity associated with blockchain can also slow down the decision-making processes, potentially delaying critical CCUS deployments needed to meet urgent carbon reduction targets.

While blockchain enhances data sharing and transparency, it raises significant data privacy concerns. The

immutable nature of blockchain means that once data is entered, it cannot be altered or removed, posing potential risks in handling sensitive or proprietary information. Ensuring that the blockchain architecture complies with global data protection regulations requires sophisticated solutions that can segregate and protect sensitive data without undermining the benefits of transparency and traceability that blockchain offers. A notable paradox in utilizing blockchain for environmental initiatives like CCUS is the technology's own environmental footprint, particularly for systems that rely on energy-intensive consensus mechanisms like Proof of Work (PoW). The additional energy consumption required for blockchain operations could contribute to the greenhouse gas emissions that CCUS seeks to mitigate. This aspect is critically important when considering the overall sustainability and environmental impact of integrating blockchain into CCUS frameworks.

As CCUS initiatives expand, the blockchain system must be able to scale accordingly without compromising performance or security. Achieving scalability while maintaining the decentralized nature of blockchain poses technical challenges. Furthermore, the decentralized decision-making process inherent in blockchain networks requires consensus among all participants, which can be difficult to achieve across a diverse stakeholder group with varying interests, technological readiness, and priorities. To overcome these challenges, collective efforts from stakeholders, government bodies, and



technological experts are imperative to ensure the smooth integration of CCUS supply chains with blockchain technology, facilitating transparent and effective strategies for emission reduction and cost management.

Considering these complexities, our research emphasizes the need for a balanced approach that leverages blockchain's strengths while addressing its inherent drawbacks. We propose a blockchain architecture that combines the transparency and immutability of public and private blockchain features through smart contracts. This approach, coupled with the development of smart contracts tailored to automate and streamline CCUS-specific operations, aims to minimize the operational and environmental drawbacks of blockchain, ensuring that its integration into the CCUS carbon tax nexus contributes positively to the global effort against climate change.

The consensus mechanism adopted in this study is the Delegated Proof of Stake (DPoS), which is strategically chosen for its efficiency in reducing the computational power required for transaction validation and consensus achievement. Unlike the conventional Proof of Work (PoW) mechanism, which is known for its high energy consumption and consequent environmental emissions, DPoS significantly minimizes these impacts. This approach not only aligns with our environmental sustainability goals but also enhances the overall efficiency and scalability of the blockchain network by delegating the responsibility of transaction validation and block creation to a select group of trusted elected representatives. This mechanism ensures a more environmentally friendly and cost-effective solution for the blockchain framework employed in our CCUS system under carbon tax policy.

Our approach introduces a blockchain architecture within the CCUS carbon tax nexus, significantly advancing current methodologies by balancing the need for transparency, data integrity, and stakeholder privacy. The proposed approach utilizes public and private features of smart contracts. It ensures the immutability and transparency critical for public trust and regulatory compliance while also providing the privacy and scalability necessary for industrial adoption and efficient operation. This innovative framework addresses key concerns around data privacy, scalability, and operational complexity that have hampered broader blockchain application in environmental and carbon management systems.

A cornerstone of our approach is the strategic use of smart contracts to automate critical processes within the CCUS supply chain, from carbon emission verification to carbon tax computation and compliance reporting. This automation significantly reduces administrative overhead, minimizes human error, and speeds up the decision-making process, allowing for more dynamic and responsive CCUS operations. Moreover, by eliminating the need for third-party verification, we expect to see a reduction in operational costs, further incentivizing the

adoption of CCUS technologies by industries.

The expected benefits of our approach are profound. By enhancing the transparency and integrity of carbon and tax data, we aim to foster greater trust among stakeholders, including regulatory bodies, industries, and the public, thereby facilitating a more collaborative and efficient approach to carbon management. The reduction in operational costs and the improved efficiency of regulatory compliance processes are anticipated to accelerate the deployment of CCUS technologies, contributing significantly to global efforts to mitigate climate change. Furthermore, the environmental impact of blockchain operations is minimized through our energy-efficient blockchain design, aligning the technology's application with the overarching goal of reducing greenhouse gas emissions.

In essence, our approach not only addresses the immediate challenges faced by the CCUS carbon tax nexus but also sets a new standard for the integration of blockchain technology in environmental governance, promising a more sustainable, transparent, and efficient framework for global climate change mitigation efforts.

## ILLUSTRATIVE EXAMPLE

This section presents an illustrative case study for the proposed approach of integrating CCUS optimization models with consideration for carbon tax implications through blockchain technology. The methodology builds upon a previously developed Mixed Integer Non-Linear Optimization Model (MINLP) [23]. The objective function, as expressed in Equation (1), seeks to minimize the Total Annual Cost (TAC), encompassing tax cost  $C_s^{\text{Tax}}$ , treatment  $C_{s,k}^{\text{Treat}}$ , compression  $C_{s,k}^{\text{Comp}}$ , transportation  $C_{s,k}^{\text{Trans}}$ , and utilization  $C_{s,k}^{\text{Sinks}}$  costs.

The objective function subject to constraints ensuring that the carbon flow from each source equals the treated flowrate  $T_{s,k,t}$ , untreated flowrate  $U_{s,k}$ , and CO<sub>2</sub> emissions to the atmosphere  $R_{s,\text{emit}}$  as indicated via Equation (2). The CO<sub>2</sub> utilized in each sink  $F_k$  defined as the sum of treated  $T_{s,k,t}$  and untreated  $U_{s,k}$  CO<sub>2</sub> transported from different sources as shown in Equation (3). Chemical absorption amine units were employed for CO<sub>2</sub> treatment due to its maturity. The costs associated with capturing, compression, transportation, and utilization were computed following the methodology outlined in [23].

$$\text{TAC} = C_s^{\text{Tax}} + C_{s,k}^{\text{Treat}} + C_{s,k}^{\text{Comp}} + C_{s,k}^{\text{Trans}} + C_k^{\text{Sinks}} \quad (1)$$

$$R_s = \sum_{k \in K} \sum_{t \in T} T_{s,k,t} + \sum_{k \in K} U_{s,k} + R_{s,\text{emit}} ; \forall s \in S \quad (2)$$

$$F_k = \sum_{s \in S} \sum_{t \in T} T_{s,k,t} + \sum_{s \in S} U_{s,k} ; \forall k \in K \quad (3)$$

Data detailing CO<sub>2</sub> composition ( $X_{CO_2}$ ), flow rate, treatment cost parameter ( $C_t$ ), sink processing cost parameter ( $C_s$ ), and emission factors ( $C_e$ ) which are expressed in tons of CO<sub>2</sub> emitted per tons of CO<sub>2</sub> processed, for CO<sub>2</sub> sources and sinks were acquired following the registration of emission sources and sinks through the IdentityRegistration.sol smart contract. The data are presented in Tables 2 and 3, respectively. CO<sub>2</sub> compositions and flow rates exhibit variations across industries. This study examined four carbon emission sources: ammonia production facility, steel-iron mill, power generation plant, and refinery. Furthermore, it included five carbon dioxide sinks, specifically: algae-based carbon utilization system, saline aquifer storage, and the production processes for methanol, urea, and acetic acid. The study sets a target net emission reduction limit of 2.48 MM t CO<sub>2</sub>/yr and a tax rate of 50 \$/t CO<sub>2</sub>. Industrial facilities face a decision to either adopt CCUS technologies or remit a carbon tax to meet the government-mandated net emissions reduction target. The choice is based on minimizing the total annual cost. All plants are assumed to operate 8000 h per year. The blockchain-based CCUS-carbon tax optimization model was implemented, and the results are discussed below.

**Table 2:** Carbon dioxide source data.

Source	$X_{CO_2}$ (wt%)	CO <sub>2</sub> Flow (t/d)	$C_t$ (\$/t CO <sub>2</sub> )
Ammonia	100	977	0
Steel-iron	44	3451	29
Power plant	7	9385	43
Refinery	27	1092	35

**Table 3:** Carbon dioxide sink data.

Sink	$X_{CO_2}$ (wt%)	Flow (t/d)	$C_s$ (\$/t CO <sub>2</sub> )	$C_e$ (t/t CO <sub>2</sub> )
Algae	6	344	-7	0.28
Storage	94	7500	9	0
Methanol	99	2072	-20	0.17
Urea	99	1488	-17	0.11
Acetic Acid	99	2500	-26	0.34

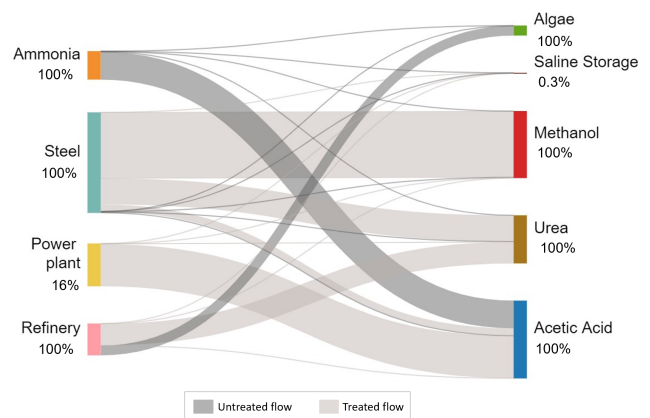
The optimal design achieved, through the implementation of CCUS, a carbon reduction of 1.5 million tons of CO<sub>2</sub> per year, while the remainder of the net reduction goal was released into the atmosphere and subsequently subjected to taxation. The total annual cost of the optimal design is 19 MM \$/yr, with a carbon tax cost amounting to 49 MM \$/yr. Table 4 delineates the composition of the net carbon dioxide reduction target, illustrating that a significant portion was achieved through the implementation of CCUS technology, while the remainder, not captured, was emitted, and subsequently subjected to

taxation. This underscores the efficacy of carbon taxation as a driving force behind the adoption of CCUS technologies.

**Table 4:** Analysis of carbon dioxide emission reduction target.

Element	Flow (MM t CO <sub>2</sub> /yr)
Net reduction target	2.48
Carbon captured	1.5
Carbon emitted	0.98

Optimal source to sink carbon allocations were identified and presented in Figure 4. The percentages on the left side of Figure 4 depict the CO<sub>2</sub> usage across different sources, whereas the percentages on the right show how well the CO<sub>2</sub> demands of each sink were satisfied. CO<sub>2</sub> from all sources was fully captured, utilized, or stored, except from the power plant as the low carbon quality limited utilization. To meet the government's targeted net reduction, CO<sub>2</sub> was directed towards economically profitable sinks such as algae, methanol, urea, and acetic acid, where the demand for carbon was completely satisfied, leveraging the economic benefits derived from these CO<sub>2</sub> utilization pathways. Conversely, the use of saline storage as a sink was not maximized due to its associated higher costs, including those for treatment, transport, and processing. Consequently, the CO<sub>2</sub> that was not sequestered or utilized was released into the atmosphere, resulting in it being subject to carbon taxation.



**Figure 4:** CO<sub>2</sub> source to sink allocation for the proposed system.

Most of the carbon demands for the algae sink are satisfied by the refinery's emissions, which do not require further treatment to align with the algae's carbon flow and quality specifications. Carbon from the steel-iron

sector is predominantly processed and then used to supply the necessary carbon flow for both methanol and urea production. The carbon requirements for the methanol sink are primarily fulfilled using processed carbon emitted by the steel-iron facility, whereas urea benefits from treated carbon emissions from both steel and refinery operations. Additionally, the carbon demands for acetic acid production are mainly met by emissions from ammonia and power generation plants. Notably, most of the carbon utilized in these processes is treated, except for the direct use of pure CO<sub>2</sub> from ammonia sources and emissions from refineries to algae sinks. This exception is due to the refinery's emissions directly meeting the algae's relatively low-quality carbon requirements.

Table 5 illustrates carbon emissions at various stages of the CCUS process, encompassing emissions from sources, capture units, transportation, and sinks. The emissions detailed on the source side reflect those discharged post-CCUS implementation. The predominant sources of emissions are identified as the main source of emissions, succeeded by carbon sinks, transportation, and capture facilities in terms of contribution. All emissions data are logged as transactions in the blockchain, facilitating the monitoring of CO<sub>2</sub> emissions throughout the CCUS supply chain.

**Table 5:** Carbon dioxide emissions from CCUS subsystems.

CCUS module	Flow (t CO <sub>2</sub> /d)
Sources	7883
Capturing unit	190
Transportation	245
Sinks	1462

Each transaction plays a crucial role in ensuring accountability, traceability, and reliability throughout the optimal CCUS network. The first transaction provides carbon dioxide generation data from sources, providing a foundation for mitigation strategies. The second transaction records emissions data from treatment units and the third transaction details emissions during compression and pumping, crucial for evaluating the environmental impact of transportation. The fourth transaction monitors sinks emissions, offering insights into sink's effectiveness. The final transaction summarizes financial aspects, including carbon tax payments, and total annual cost essential for assessing the project's economic viability and ensuring sustainable carbon management.

The blockchain-based optimal CCUS supply chain ensures transparency, accountability, and trust throughout the CCUS process, supporting environmental and economic goals. The study successfully meets the net emission capture target, and comprehensive results, encompassing both carbon dioxide flow rates and financial

implications, are securely recorded in the blockchain, ensuring tamper-proof integrity.

## CONCLUSIONS

This study highlights the significant potential of blockchain technology to enhance transparency, efficiency, and regulatory compliance in CCUS supply chains under carbon tax regimes. By developing and implementing the IdentityRegistration.sol and CO2Exchange.sol smart contracts, we established a framework for automating carbon emissions tracking and regulatory reporting, addressing key challenges such as emission verification and carbon leakage. The adoption of the Delegated Proof of Stake (DPoS) consensus mechanism further aligns our approach with environmental sustainability by minimizing the energy consumption of blockchain operations. Our case study demonstrates that a substantial portion of carbon reduction targets can be met through CCUS, with carbon taxation serving as a compelling incentive for technology adoption. This research contributes to the optimization of CCUS networks, offering a scalable and sustainable solution to climate change mitigation and highlighting the critical role of innovative technologies in achieving global sustainability goals.

## ACKNOWLEDGMENT

The authors would like to express their gratitude to Qatar Foundation (QF) for providing financial support and resources for this research. This study on the integration of CCUS, carbon taxes, and blockchain technology has been made possible through the funding received under the fellowship program supported by QF.

## REFERENCES

1. IPCC, "Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change," Geneva, Switzerland, 2023. doi: 10.59327/IPCC/AR6-9789291691647.
2. IEA, "Special Report on Carbon Capture Utilisation and Storage - CCUS in clean energy transitions," International Energy Agency, 2020. doi: 10.1787/208b66f4-en.
3. E. Liu, X. Lu, and D. Wang, "A Systematic Review of Carbon Capture, Utilization and Storage: Status, Progress and Challenges," *energies*, 2023, doi: 10.3390/en16062865.
4. L. DOU *et al.*, "Trend of global carbon dioxide capture, utilization and storage industry and challenges and countermeasures in China," *Pet. Explor. Dev.*, vol. 50, no. 5, pp. 1246–1260, 2023, doi: 10.1016/S1876-3804(23)60463-X.
5. D. Hagmann, E. H. Ho, and G. Loewenstein,

- "Nudging out support for a carbon tax," *Nat. Clim. Chang.*, vol. 9, no. 6, pp. 484–489, 2019, doi: 10.1038/s41558-019-0474-0.
6. N. Devi, P. Rani, and A. Gokul, *Introduction to Blockchain Technology, Blockchain Applications in Cybersecurity Solutions*. 2023. doi: 10.2174/9789815080599123010004.
  7. S. Namasudra and K. Akkaya, *Blockchain and its Applications in Industry 4.0*, vol. 119. 2023. doi: 10.1007/978-981-19-8730-4.
  8. G. Habib, S. Sharma, S. Ibrahim, I. Ahmad, S. Qureshi, and M. Ishfaq, "Blockchain Technology: Benefits, Challenges, Applications, and Integration of Blockchain Technology with Cloud Computing," *Futur. Internet*, vol. 14, no. 11, pp. 1–22, 2022, doi: 10.3390/fi14110341.
  9. K. Thakur, A.-S. K. Pathan, and S. Ismat, "Blockchain Technology BT - Emerging ICT Technologies and Cybersecurity: From AI and ML to Other Futuristic Technologies," K. Thakur, A.-S. K. Pathan, and S. Ismat, Eds. Cham: Springer Nature Switzerland, 2023, pp. 125–145. doi: 10.1007/978-3-031-27765-8\_4.
  10. D. Y. Shu, S. Deutz, B. A. Winter, N. Baumgärtner, L. Leenders, and A. Bardow, "The role of carbon capture and storage to achieve net-zero energy systems: Trade-offs between economics and the environment," *Renew. Sustain. Energy Rev.*, vol. 178, no. July 2022, 2023, doi: 10.1016/j.rser.2023.113246.
  11. S. Nie *et al.*, "Economic costs and environmental benefits of deploying CCUS supply chains at scale: Insights from the source–sink matching LCA–MILP approach," *Fuel*, vol. 344, no. February, p. 128047, 2023, doi: 10.1016/j.fuel.2023.128047.
  12. J. Han *et al.*, "Coal-fired power plant CCUS project comprehensive benefit evaluation and forecasting model study," *J. Clean. Prod.*, vol. 385, no. September 2022, 2023, doi: 10.1016/j.jclepro.2022.135657.
  13. L. Yang, H. Lv, N. Wei, Y. Li, and X. Zhang, "Dynamic optimization of carbon capture technology deployment targeting carbon neutrality, cost efficiency and water stress: Evidence from China's electric power sector," *Energy Econ.*, vol. 125, no. July 2022, 2023, doi: 10.1016/j.eneco.2023.106871.
  14. L. Sun and W. Chen, "Impact of carbon tax on CCUS source–sink matching: Finding from the improved ChinaCCS DSS," *J. Clean. Prod.*, vol. 333, p. 130027, 2022, doi: 10.1016/j.jclepro.2021.130027.
  15. K. Zhang, H. C. Lau, and Z. Chen, "The contribution of carbon capture and storage to Canada's net-zero plan," *J. Clean. Prod.*, vol. 404, no. December 2022, 2023, doi: 10.1016/j.jclepro.2023.136901.
  16. G. Leonzio, P. U. Foscolo, E. Zondervan, and I. D. L. Bogle, "Scenario Analysis of Carbon Capture, Utilization (Particularly Producing Methane and Methanol), and Storage (CCUS) Systems," *Ind. Eng. Chem. Res.*, vol. 59, no. 15, pp. 6961–6976, 2020, doi: 10.1021/acs.iecr.9b05428.
  17. S. Zhang, Y. Zhuang, L. Liu, L. Zhang, and J. Du, "Risk management optimization framework for the optimal deployment of carbon capture and storage system under uncertainty," *Renew. Sustain. Energy Rev.*, vol. 113, no. 2, 2019, doi: 10.1016/j.rser.2019.109280.
  18. W. Tang, Z. Peng, and J. J. Sheng, "Evaluation of CO<sub>2</sub> utilization and storage potential in the Jimsar shale play from an optimization study," *Geoenergy Sci. Eng.*, vol. 224, no. February, 2023, doi: 10.1016/j.geoen.2023.211607.
  19. A. Al Sadawi and M. Ndiaye, "Blockchain-based carbon trading mechanism to elevate governance and smartness," *ACM Int. Conf. Proceeding Ser.*, pp. 34–43, 2021, doi: 10.1145/3494193.3494198.
  20. A. Richardson and J. Xu, *Carbon Trading with Blockchain*. Springer International Publishing, 2020. doi: 10.1007/978-3-030-53356-4\_7.
  21. Y. Bai, T. Song, Y. Yang, O. Bocheng, and S. Liang, "Construction of carbon trading platform using sovereignty blockchain," *Proc. - 2020 Int. Conf. Comput. Eng. Intell. Control. ICCEIC 2020*, pp. 149–152, 2020, doi: 10.1109/ICCEIC51584.2020.00037.
  22. D. Patel, B. Britto, S. Sharma, K. Gaikwad, Y. Dusing, and M. Gupta, "Carbon Credits on Blockchain," *2020 Int. Conf. Innov. Trends Inf. Technol. ICITIIT 2020*, pp. 4–8, 2020, doi: 10.1109/ICITIIT49094.2020.9071536.
  23. P. Linke and D. M. Al-mohannadi, "Towards systematic approaches for the development of carbon dioxide management policy and strategy," vol. 2016, p. 5339, 2016.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Optimal Membrane Cascade Design for Critical Mineral Recovery Through Logic-based Superstructure Optimization

Daniel Ovalle<sup>a</sup>, Norman Tran<sup>a</sup>, Carl D. Laird<sup>a</sup>, and Ignacio E. Grossmann<sup>a\*</sup>

<sup>a</sup> Carnegie Mellon University, Department of Chemical Engineering, Pittsburgh, Pennsylvania 15213, United States

\* Corresponding Author: [grossmann@cmu.edu](mailto:grossmann@cmu.edu).

## ABSTRACT

Critical minerals and rare earth elements play an important role in our climate change initiatives, particularly in applications related with energy storage. Here, we use discrete optimization approaches to design a process for the recovery of Lithium and Cobalt from battery recycling, through membrane separation. Our contribution involves proposing a Generalized Disjunctive Programming (GDP) model for the optimal design of a multistage diafiltration cascade for Li-Co separation. By solving the resulting nonconvex mixed-integer nonlinear program model to global optimality, we investigated scalability and solution quality variations with changes in the number of stages and elements per stage. Results demonstrate the computational tractability of the nonlinear GDP formulation for design of membrane separation processes while opening the door for decomposition strategies for multicomponent separation cascades. Future work aims to extend the GDP formulation to account for stage installation and explore various decomposition techniques to enhance solution efficiency.

**Keywords:** Critical Minerals, Lithium Recovery, Diafiltration Cascade, Superstructure Optimization, Generalized Disjunctive Programming, Mixed-Integer Nonlinear Programming.

## INTRODUCTION AND RELATED WORK

The U.S. relies on the import of rare earth elements (REE) and critical minerals (CM) which are central to our climate change initiatives, particularly in applications such as electric vehicles and energy storage [1]. With an increasing demand for REE and CM, restrictions by competitors on exports have disrupted the U.S. supply chain and pose a risk to the national economy [2]. Projections by the IEA indicate a staggering 60% surge in renewable energy power capacity from 2020 to 2026 [3]. As a result, recycling REE and CM has become one of the main objectives of the Department of Energy (DOE) since 2014 [1].

One potential source of Lithium recovery is Lithium-ion battery recycling [4]. This process addresses the lack of domestic CM production [5], as it can recover battery-grade Lithium and Cobalt for a fraction of the materials extracted from either brine or ores [6]. However, concerns arise from improper recycling, which affects human

health and the environment [7], prompting opportunities to employ energy efficient processes with minimal environmental impact like membrane separation.

Diafiltration membranes offer significant advantages in the efficiency of recycling CM over existing battery recycling pathways, reducing energy needs and chemical use and cost. In a diafiltration membrane, the process employs a dilute solution, called diafiltrate, to reduce the solubility limit effect that leads to fouling. This approach allows the staging of membrane units into cascades, facilitating the extraction of lithium and cobalt from leach liquors.

There is a growing interest in utilizing membranes for the recovery and extraction of CM, as evident in recent literature. A comprehensive review and feasibility assessment for membrane-based technology in lithium recovery were provided by Li et al. [8]. Similarly, Alvarez et al. [9] demonstrated the potential to enhance water security by utilizing membranes to filter out metals during water treatment. In the realm of lithium recovery



methods, Bae et al. [10] conducted a study that investigated and compared various proposed techniques, including chemical extraction and selective membrane processes, with a focus on quantitative efficiency and purity based on existing literature. Razmjou et al. [11] delved into the physical construction of a nanofiltration membrane for lithium recovery, proposing design principles to enhance selectivity. The prevailing trend in the current literature emphasizes a primary understanding and performance evaluation of membrane processes, while rigorous mathematical optimization is less commonly addressed.

In a prior study conducted by Wamble et al. [12], a superstructure formulation was introduced to determine the optimal configuration of a membrane cascade. This superstructure selects design variables (e.g., flow, concentration, stage length) while maximizing Cobalt recovery and adhering to a minimum Lithium recovery fraction, employing the epsilon constraint approach. This superstructure was posed as a nonlinear program (NLP), which was solved using IPOPT.

In the earlier method, the determination of stream connectivity was performed continuously, allowing for the division of feed, diafiltrate, and refluxed streams. However, the authors noted that this continuous approach failed to consistently generate "physically sensible cascade designs" [12]. To address this issue, the authors introduced a second optimization step. In our proposed extension of this approach, we suggest modeling stream connectivity as a discrete decision, ensuring the singular allocation of side streams. As the model incorporates bilinear mixing and nonlinear performance constraints, the proposed method requires solving a nonconvex mixed-integer nonlinear programming (MINLP) problem. Additionally, we propose formulating the MINLP superstructure as a Generalized Disjunctive Program (GDP), a widely employed approach in the literature for superstructure optimization. The motivation behind this is the fact that GDP effectively avoids singularities in nonlinear expressions, particularly when variables become zero, a phenomenon recognized in the literature as zero-flow issues [13]. This approach allows for the identification of the globally optimal superstructure and facilitates future extensions for the optimal determination of the number of separation stages.

## Generalized Disjunctive Programming

Generalized Disjunctive Programming (GDP) corresponds to a mathematical optimization framework designed for modeling and solving problems characterized by embedded logic. Within GDP, the feasible region is represented by the intersection of disjunctions of sets, employing Boolean variables as indicators for each set. In this context, a *True* value for a Boolean variable indicates that the solution resides within the corresponding set

[14]. The general formulation of a GDP is as follows:

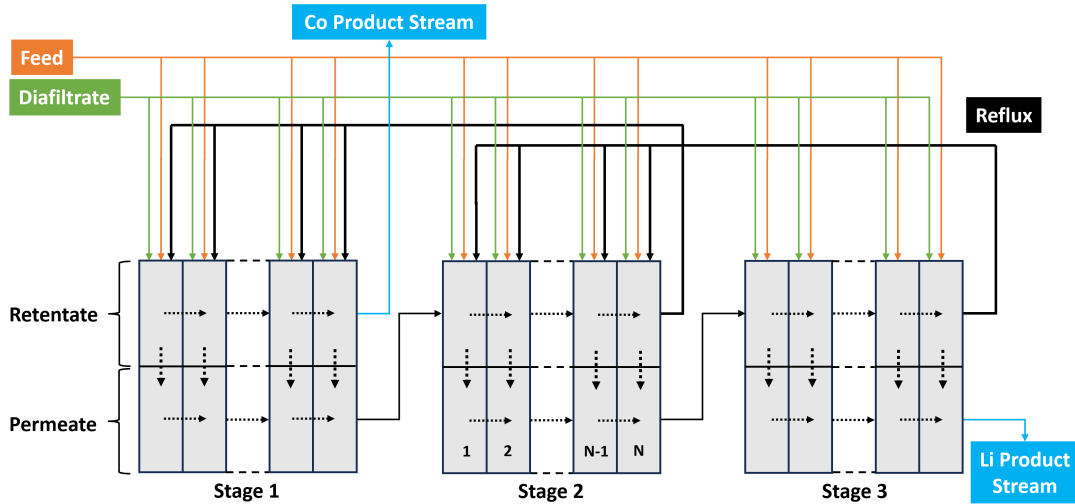
$$\begin{aligned} & \min f(x) \\ & \text{s. t. } g(x) \leq 0 \\ & \quad \bigvee_{i \in D_k} \left[ r_{ik}(x) \leq 0 \right], \quad k \in K \\ & \quad \Omega(Y_{i,k}) = \text{True} \\ & \quad x \in \mathbb{R}^n \\ & \quad Y_{ik} \in \{\text{True}, \text{False}\}, i \in D_k, k \in K \end{aligned}$$

where, an objective function  $f(x)$  is to be minimized over a set of continuous real variables  $x$  subject to a set of global constraints  $g(x) \leq 0$ . Boolean variable  $Y_{ik}$  acts as the indicator for the set implied by constraints  $r_{ik} \leq 0$ , standing for the  $i^{\text{th}}$  disjunct set of the  $k^{\text{th}}$  disjunction. Each of the  $K$  disjunctions are related with an exclusive OR operator ( $\bigvee$ ), which can be interpreted as an *exactly-One* operator when  $|D_k| > 2$ . The set of logical propositions  $\Omega$  is composed of logical clauses connected with logical operators such as AND ( $\wedge$ ), OR ( $\vee$ ), XOR ( $\bigoplus$ ), negation ( $\neg$ ), implication ( $\Rightarrow$ ), and equivalence ( $\Leftrightarrow$ ). Here,  $\Omega$  is required to be *True* to indicate that all the propositions must be satisfied.

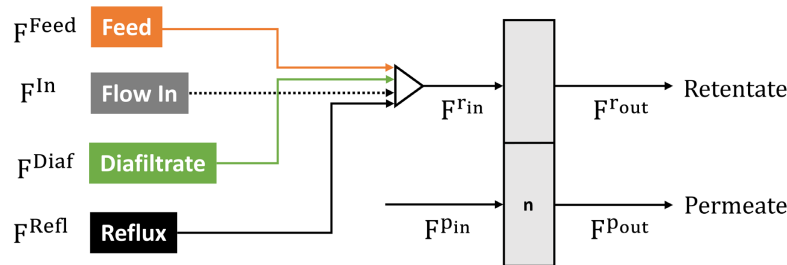
GDP offers a dual advantage by providing an intuitive means to model the logic inherent in problems and presenting a diverse array of solution methods that can be broadly categorized in two groups. One major category involves reformulating the GDP as a mixed-integer (non)linear program through various transformations. These transformations translate the logical structure of the problem into a mathematical formulation suitable for traditional MINLP optimization algorithms. Examples of such transformations include Big-M [15], Hull [16], and Hybrid Planes [17], among others. The second category encompasses logic-based decomposition methods that directly operate on the logical structure of the problem. Tailored methods within this category include Logic-based Outer Approximation [18], Logic-based Branch and Bound [16], and the Logic-based Discrete-Steepest Descent Algorithm [19], providing specialized approaches to address the unique complexities associated with the logic of the problem at hand.

## PROBLEM STATEMENT

Given is the following information for the optimization of a multistage diafiltration cascade aimed at separating Lithium from Cobalt. The provided parameters include the predetermined number of separation stages and the number of discretization elements per stage. Additionally, the model involves the flows and concentrations associated with both the feed and diafiltrate streams. Relevant membrane parameters, such as the stage width, solvent flux across the membrane, and the



**Figure 1:** Superstructure sketch for a three-stage ( $|K| = 3$ ) diafiltration membrane cascade.



**Figure 2:** Sketch of the flow structure of a single discretized element.

sieving coefficients for both components, are also specified. Furthermore, the model includes the performance relationships that govern the system's behaviour as provided in Wable et al. [12].

Design variables that specify the optimal diafiltration membrane cascade superstructure need to be determined. Firstly, the optimal stream connections within the superstructure must be established. This involves identifying the position (in terms of stage and discretized element) for the feed, the diafiltrate, and the recycle from upstream stages. Additionally, the length of each stage in the diafiltration cascade needs to be determined to achieve an efficient separation process. Moreover, comprehensive determination of all flows and concentrations in regard to both components (Lithium and Cobalt) throughout the entire system must be determined. This comprehensive set of determinations forms the basis for configuring an optimized and well-functioning multistage diafiltration system for the separation of Lithium and Cobalt.

The primary objective of the optimization model is to maximize the recovery of Cobalt within a Cobalt-rich stream extracted from the retentate flow of the initial stage in the diafiltration cascade. This goal is set against the constraint of ensuring a minimum recovery of Lithium

within a Lithium-rich stream exiting the permeate flow of the final stage. In essence, the goal is to manage the inherent multi-objective nature of recovering both Lithium and Cobalt. In this work, we consider the system to be isotropic, meaning that all stages share identical lengths. The configuration of the superstructure is presented in Figure 1.

## PROPOSED GDP SUPERSTRUCTURE

In this section, we illustrate the modifications made to the original model, primarily focusing on incorporating the necessary logic into the GDP framework to represent installing a singular position for the feed ( $F^{Feed}$ ), diafiltrate ( $F^{Diaf}$ ), and reflux per stage ( $F_n^{Refl}$ ). To accomplish this objective, we model each discrete element individually, as illustrated in Figure 2, wherein the incoming side flows are handled in a disaggregated manner. Through this approach, the following disjunctions are implemented, allowing for the activation or deactivation of incoming streams, thereby determining the existence of an incoming side stream. Here, the Boolean variables  $Y_{nk}^{Feed}$ ,  $Y_{nk}^{Diaf}$ , and  $Y_{nk}^{Refl}$  indicate the existence of a feed, diafiltrate or reflux in a particular position respectively.

$$\left[ \begin{array}{c} Y_{nk}^{Feed} \\ F_{nk}^{Feed} = \hat{F}^{Feed} \\ x_{cnk}^{Feed} = \hat{x}_c^{Feed} \quad \forall c \in C \end{array} \right] \vee \left[ \begin{array}{c} \neg Y_{nk}^{Feed} \\ F_{nk}^{Feed} = 0 \\ x_{cnk}^{Feed} = 0 \quad \forall c \in C \end{array} \right],$$

$$\forall n \in N, k \in K \quad (1)$$

$$\left[ \begin{array}{c} Y_{nk}^{Diaf} \\ F_{nk}^{Diaf} = \hat{F}^{Diaf} \\ x_{cnk}^{Diaf} = \hat{x}_c^{Diaf} \quad \forall c \in C \end{array} \right] \vee \left[ \begin{array}{c} \neg Y_{nk}^{Diaf} \\ F_{nk}^{Diaf} = 0 \\ x_{cnk}^{Diaf} = 0 \quad \forall c \in C \end{array} \right],$$

$$\forall n \in N, k \in K \quad (2)$$

$$\left[ \begin{array}{c} Y_{nk}^{Refl} \\ F_{nk}^{Refl} = F_{|N|, \{k+1\}}^{r_{out}} \\ x_{cnk}^{Refl} = x_{c, |N|, \{k+1\}}^{r_{out}} \quad \forall c \in C \end{array} \right] \vee \left[ \begin{array}{c} \neg Y_{nk}^{Refl} \\ F_{nk}^{Refl} = 0 \\ x_{cnk}^{Refl} = 0 \quad \forall c \in C \end{array} \right],$$

$$\forall n \in N, k \in K \setminus \{K\} \quad (3)$$

where  $\hat{F}^j$  and  $\hat{x}^j$  represent the known flow and composition of the incoming side stream  $j \in \{Feed, Diaf\}$ . The first two disjunctions establish connections between the incoming stream and known parameter values. In contrast, the third disjunction, which models the recycle, links a discretized element to the stream coming from the retentate of the subsequent stage. While these disjunctions exclusively represent the presence of a stream at a given position, we also ensure a singular feed and diafiltrate throughout the entire structure, as well as a singular recycle per stage. To account for this logic, the following cardinality clauses are introduced.

$$exactlyOne([Y_{nk}^{Feed} \quad \forall n \in N, k \in K]) \quad (4)$$

$$exactlyOne([Y_{nk}^{Diaf} \quad \forall n \in N, k \in K]) \quad (5)$$

$$exactlyOne([Y_{nk}^{Refl} \quad \forall n \in N]), \quad k \in K \setminus \{K\} \quad (6)$$

Equations (1-6) represent the modeling logic of the superstructure. Next, with the introduced side stream disaggregation, the per-element mass balances are outlined as follows:

$$F_{nk}^{In} + F_{nk}^{Feed} + F_{nk}^{Diaf} + F_{nk}^{Refl} = F_{nk}^{r_{in}}, \quad \forall n \in N, k \in K \quad (7)$$

$$F_{nk}^{In} x_{cnk}^{In} + F_{nk}^{Feed} x_{cnk}^{Feed} + F_{nk}^{Diaf} x_{cnk}^{Diaf} + F_{nk}^{Refl} x_{cnk}^{Refl} = F_{nk}^{r_{in}} x_{cnk}^{r_{in}}, \quad \forall c \in C, n \in N, k \in K \quad (8)$$

$$F_{nk}^{r_{in}} + F_{nk}^{p_{in}} = F_{nk}^{r_{out}} + F_{nk}^{p_{out}}, \quad \forall n \in N, k \in K \quad (9)$$

$$F_{nk}^{r_{in}} x_{cnk}^{r_{in}} + F_{nk}^{p_{in}} x_{cnk}^{p_{in}} = F_{nk}^{r_{out}} x_{cnk}^{r_{out}} + F_{nk}^{p_{out}} x_{cnk}^{p_{out}}, \quad \forall c \in C, n \in N, k \in K \quad (10)$$

The permeate mass balance is a function of the membrane flux ( $J$ ), the width of the stage ( $w$ ) and the element

length ( $\hat{L}$ ) as:

$$F_{nk}^{p_{in}} + Jw\hat{L}_{nk} = F_{nk}^{p_{out}}, \quad \forall n \in N, k \in K \quad (11)$$

In the previous study, the performance equation was examined in terms of a log transform. However, we suggest maintaining the equation in its exponential form and rearrange the expression to remove the fraction. This reformulation ensures a well-defined expression within the domain of our variables, mitigating the risk of evaluation errors and contributing to the overall stability and well-behaved nature of the NLP formulation. The resulting expression where,  $S$  is the sieving coefficient per component is as follows:

$$x_{cnk}^{r_{out}} = x_{cnk}^{r_{in}} \left( \frac{F_{nk}^{r_{out}}}{F_{nk}^{r_{in}}} \right)^{S_c - 1}, \quad \forall c \in C, n \in N, k \in K \quad (12)$$

The elements can now be interconnected to create stages, where the key concept is to establish connections between the flow and composition exiting the permeate of one stage with the retentate entering the subsequent stage. These connections can be expressed as:

$$F_{n\{k-1\}}^{p_{out}} = F_{nk}^{r_{in}}, \quad \forall n \in N, k \in K \setminus \{1\} \quad (13)$$

$$x_{cn\{k-1\}}^{p_{out}} = x_{cnk}^{r_{in}}, \quad \forall c \in C, n \in N, k \in K \setminus \{1\} \quad (14)$$

Moreover, the permeate and the flow-in entering the first element of each membrane are both assigned a value of zero. Similarly, in the last stage, there is no recycle entering.

$$F_{1k}^{p_{in}} = F_{1,1}^{r_{in}} = 0, \quad \forall k \in K \quad (15)$$

$$x_{c1k}^{p_{in}} = x_{c1k}^{r_{in}} = 0, \quad \forall c \in C, k \in K \quad (16)$$

$$F_{n|K|}^{Refl} = 0, \quad \forall n \in N \quad (17)$$

$$x_{cn|K|}^{Refl} = 0, \quad \forall c \in C, k \in K \quad (18)$$

A minimum Lithium recovery ( $R_{\min}^{Li}$ ) from the lithium-rich stream in the permeate exiting the last stage must be satisfied.

$$F_{|N||K|}^{p_{out}} x_{Li, |N||K|}^{p_{out}} \geq R_{\min}^{Li} (\hat{F}^{Feed} \hat{x}_{Li}^{Feed} + \hat{F}^{Diaf} \hat{x}_{Li}^{Diaf}) \quad (19)$$

The objective function is to maximize the recovery of Cobalt in the Cobalt-rich stream exiting the permeate of the first stage. Considering that the cobalt entering the system remains constant, it suffices to maximize the following expression:

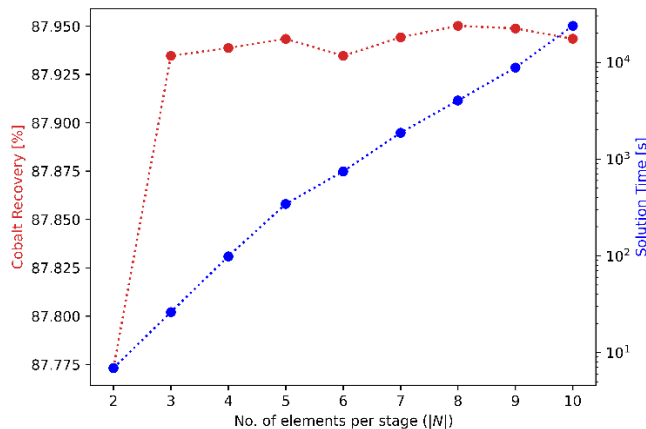
$$\max F_{|N|,1}^{r_{out}} x_{Co, |N|,1}^{r_{out}} \quad (20)$$

For the remaining aspects of the model formulation and parameter values along with a deeper understanding of the system's physical intuition, we direct the readers to Wamble et al. [12].

## RESULTS

The proposed GDP superstructure model was transformed into a nonconvex MINLP problem via a Big-M transformation. To solve this MINLP, the solver SCIP v8.0 was used, implemented through GAMS 40.4.0. The computation was performed on a Linux machine featuring 8 Intel Xeon Gold 6234 CPUs operating at 3.30 GHz, with a single hardware thread, and equipped with 1 TB of RAM, all within the Ubuntu environment. It is worth mentioning that we attempted solving the same problem using BARON v22.7.23 but the solver, incorrectly, found the problem to be infeasible.

We conducted a study of the scalability and solution quality of the proposed superstructure. For this, we opted for a three-stage ( $|K| = 3$ ) membrane cascade superstructure, varying the number of discretized elements. All superstructures were solved to global optimality (up to 0.001% gap). Figure 3 encapsulates the outcomes, specifically summarizing the results for the case where a minimum Lithium recovery ( $R_{\min}^{Li}$ ) of 60% was required.



**Figure 3:** Cobalt recovery and solution time of a three-stage superstructure with a minimum Lithium recovery of 60% for different number of discretized elements per stage.

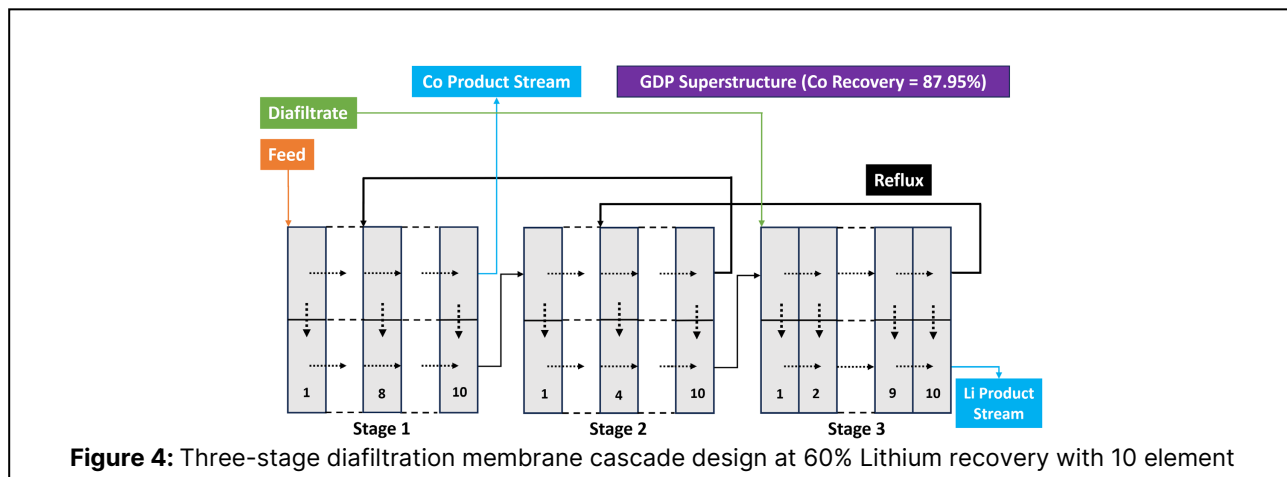
As anticipated, the solution time exhibits rapid growth with the increasing size of the superstructure, as a larger number of elements results in a larger model. This poses a significant limitation when compared to the previous methodology that could solve large superstructures within seconds. The challenge arises from the fact that achieving global optimality in solving a nonconvex

MINLP is computationally much more demanding than solving an NLP to local optimality. However, despite this computational complexity, the proposed approach demonstrates a slightly better solution compared to the previous methods. Figure 4 shows the resulting three-stage superstructure (with  $|N| = 10$ ) for a lithium recovery of 60%. It is noteworthy that the locally optimal superstructure obtained by Wamble et al. [12] exhibits the feed stream split at two different locations, while our superstructure is characterized by a single side stream allocation. Nevertheless, our approach yields a similar Cobalt recovery (0.25% increase) while also ensuring the construction of a physically sensible membrane cascade. Considering that we are dealing with a membrane design problem that requires a one-time solution, it may be preferable to tackle the problem to achieve global optimality. Furthermore, we want to extend this approach to use GDP methods for multicomponent cascade recycling.

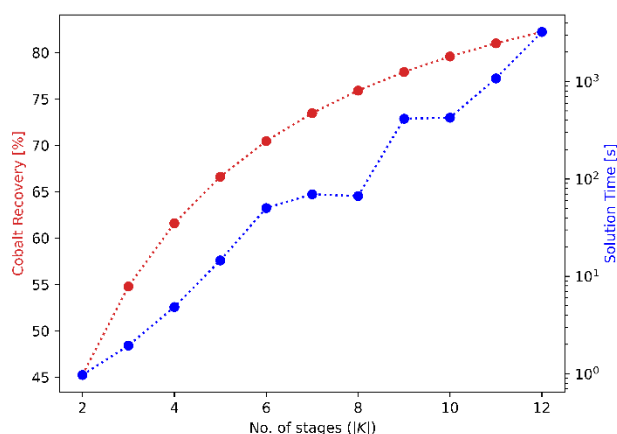
We proceed to assess the scalability of the number of stages, opting for a discretization of only two elements per stage. This choice was informed by the examination of objective values presented in Figure 3, revealing a relatively stable Cobalt recovery. Notably, this value does not exhibit a monotonic trend in relation to the number of discretized elements. Therefore, for this specific membrane system and its performance equations, the resolution achieved by increasing the number of discretization elements does not substantially impact the objective. Consequently, a coarser discretization can be employed. The outcomes of this scalability analysis can be found in Figure 5.

As expected, the Cobalt recovery rises as we include more stages. Interestingly, and in contrast to the scenario where discrete elements were added, an increase in the number of stages leads to a substantial rise in the recovery. It can be appreciated that the recovery values almost double when progressing from 2 stages to 12. This captures the trade-off between the profit of recovering Cobalt in kg/hr and the expense associated with installing an additional membrane stage. Although current solution approaches cannot easily address this problem, within the GDP framework, tackling this trade-off is straightforward. To achieve this, a Boolean variable can be introduced to represent the presence or absence of a stage. When the Boolean variable is *True*, the performance equations for the stage are considered. Conversely, when the Boolean variable is *False*, the stage operates as a bypass, and no side streams can be installed.

Regarding the solution time, it is evident that reducing the number of discretized elements has diminished the model's size, enabling the solution of larger structures. Although this reduction may lead to a loss in resolution, potentially impacting the quality of the solution, as discussed earlier, we anticipate the recovery values to remain within a similar order for a higher discretization.



Furthermore, the application of decomposition strategies will be investigated to tackle large-scale conceptual design problems.



**Figure 5:** Cobalt recovery and solution time for a superstructure with a minimum Lithium recovery of 95% for different number of stages.

## CONCLUSIONS

This study introduces a Generalized Disjunctive Programming formulation for optimizing the superstructure design of a multi-stage diafiltration cascade aimed at separating Lithium and Cobalt. The proposed superstructure ensures a unique side stream allocation for the feed, fresh diafiltrate, and recycle streams. We successfully solved the model to global optimality and investigated how the solution time and Cobalt recovery varied with different numbers of stages and discretized elements per stage. The results demonstrate the value of pursuing solutions with global optimality and highlight a trade-off between capitalizing on recovered Cobalt and the installation of additional membranes. Future research directions include expanding the GDP formulation to account for stage existence, addressing the aforementioned trade-

off. Moreover, given the computational expense of solving the superstructure directly, exploration of decomposition techniques to improve solution time is warranted. In this context, the Logic-based Steepest-Descent Algorithm (LD-SDA) [19] emerges as a promising alternative, particularly due to its suitability for handling spatially ordered Boolean decisions inherent in the membrane superstructure.

## ACKNOWLEDGEMENTS

This effort was funded by the US Department of Energy's Process Optimization and Modeling for Minerals Sustainability (PrOMMiS) Initiative, supported by the Office of Fossil Energy and Carbon Management's Office of Resource Sustainability.

## DISCLAIMER

This paper was prepared as an account of work sponsored by an agency of the United States Government through a support contract. Neither the United States Government nor any agency thereof, nor any of its employees, nor the support contractor, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## REFERENCES



1. Critical Minerals and Materials Program. National Energy Technology Laboratory (2021)
2. Lee Y, Dacass T. Reducing the United States' risks of dependency on China in the rare earth market. *Resources Policy* 77 (2022)
3. IEA. The role of critical minerals in clean energy transitions. OECD. (2021)
4. Zhang, X., Li, L., Fan, E., Xue, Q., Bian, Y., Wu, F., & Chen, R. Toward sustainable and systematic recycling of spent rechargeable batteries. *Chemical Society Reviews*, 47(19), (2018)
5. Rankin WJ. Minerals, metals and sustainability: meeting future material needs. CSIRO publishing (2011)
6. National Blueprint for Lithium Batteries. Office of Energy Efficiency and Renewable Energy (2021)
7. Kang DH, Chen M, Ogunseitan OA. Potential environmental and human health impacts of rechargeable lithium batteries in electronic waste. *Environmental science & technology* 47.10 (2013)
8. Li X, Mo Y, Qing W, Shao S, Tang CY, Li J. Membrane-based technologies for lithium recovery from water lithium resources: A review. *Journal of Membrane Science* (2019)
9. Alvarez PJ, Chan CK, Elimelech M, Halas NJ, Villagrán D. Emerging opportunities for nanotechnology to enhance water security. *Nature nanotechnology* 13.8 (2018)
10. Bae H, Kim Y. Technologies of lithium recycling from waste lithium ion batteries: a review. *Materials advances* 2.10 (2021).
11. Razmjou A, Asadnia M, Hosseini E, Habibnejad Korayem A, Chen V. Design principles of ion selective nanostructured membranes for the extraction of lithium ions. *Nature communications* 10.1 (2019)
12. Wamble NP, Eugene EA, Phillip WA, Dowling AW. Optimal diafiltration membrane cascades enable green recycling of spent lithium-ion batteries. *ACS Sustainable Chemistry & Engineering* 10.37 (2022)
13. Chen Q, Liu Y, Seastream G, Siirola JD, Grossmann IE. Pyosyn: a new framework for conceptual design modeling and optimization. *Computers & Chemical Engineering* 153 (2021)
14. Grossmann IE, Trespacios F. Systematic modeling of discrete-continuous optimization models through generalized disjunctive programming. *AIChE Journal* 59.9 (2013)
15. Raman R, Grossmann IE. Modelling and computational techniques for logic based integer programming. *Computers & Chemical Engineering* 18.7 (1994)
16. Lee S, Grossmann IE. New algorithms for nonlinear generalized disjunctive programming. *Computers & Chemical Engineering* 24.9-10 (2000)
17. Sawaya NW, Grossmann IE. A cutting plane method for solving linear generalized disjunctive programming problems. *Computers & chemical engineering* 29.9 (2005)
18. Türkay M, Grossmann IE. Logic-based MINLP algorithms for the optimal synthesis of process networks. *Computers & Chemical Engineering* 20.8 (1996)
19. Bernal, D.E., Ovalle, D., Liñán, D.A., Ricardez-Sandoval, L.A., Gómez, J.M. and Grossmann, I.E. Process Superstructure Optimization through Discrete Steepest Descent Optimization: a GDP Analysis and Applications in Process Intensification. *Computer Aided Chemical Engineering*. Vol. 49. Elsevier, (2022)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Membrane-based carbon capture process optimization using CFD modeling

Hector A. Pedrozo<sup>a</sup>, Cheick Dosso<sup>a</sup>, Lingxiang Zhu<sup>b,c</sup>, Victor Kusuma<sup>b,c</sup>, David Hopkinson<sup>b</sup>, Lorenz T. Biegler<sup>a</sup>, and Grigorios Panagakos<sup>a,b,c,\*</sup>

<sup>a</sup> Dept. of Chem. Eng., Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, United States

<sup>b</sup> National Energy Technology Laboratory, 626 Cochran Mills Rd, Pittsburgh, PA, 15236, United States

<sup>c</sup> NETL Support Contractor, 626 Cochran Mills Rd, Pittsburgh, PA, 15236, United States

\* Corresponding Author: [gpanagak@andrew.cmu.edu](mailto:gpanagak@andrew.cmu.edu)

## ABSTRACT

Carbon capture is a promising option to mitigate CO<sub>2</sub> emissions from existing coal-fired power plants, cement and steel industries, and petrochemical complexes. Among the available technologies, membrane-based carbon capture presents the lowest energy consumption, operating costs, and carbon footprint. In addition, membrane processes have important operational flexibility and response times. On the other hand, the major challenges to widespread application of this technology are related to reducing capital costs and improving membrane stability and durability. To upscale the technology into stacked flat sheet configurations, high-fidelity computational fluid dynamics (CFD) that describes the separation process accurately are required. High-fidelity simulations are effective in studying the complex transport phenomena in membrane systems. In addition, obtaining high CO<sub>2</sub> recovery percentages and product purity requires a multi-stage membrane process, where the optimal network configuration of the membrane modules must be studied in a systematic way. In order to address the design problem at process scale, we formulate a superstructure for the membrane-based carbon capture, including up to three separation stages. In the formulation of the optimization problem, we include reduced models, based on rigorous CFD simulations of the membrane modules. Numerical results indicate that the optimal design includes three membrane stages, and the capture cost is 45.4 \$/t-CO<sub>2</sub>.

**Keywords:** Carbon Capture, Membranes, Computational Fluid Dynamics, Process Design, Optimization

## INTRODUCTION

Climate change remains a relevant environmental concern with the global average temperature of Earth steadily rising, and this phenomenon is associated with the high concentration of carbon dioxide (CO<sub>2</sub>) in the atmosphere. The current levels of carbon dioxide are over 50% greater than those observed prior to the industrial era, achieving values of 424 ppm of CO<sub>2</sub> concentration [1]. CO<sub>2</sub> takes the lead as the primary contributor to climate change, and human activities, especially CO<sub>2</sub> emissions from the energy and industrial sectors, are significant drivers.

In this context, carbon capture approaches offer hope in reducing greenhouse emissions since these processes can isolate carbon dioxide at point sources. Using

this strategy, existing facilities can be retrofitted with capture plants, without changing the current industrial process significantly. Carbon capture technologies are crucial to achieving a net-zero emission scenario that limits the global average temperature increase below 1.5 °C [2]. Current CO<sub>2</sub> capture alternative pathways include absorption, adsorption, membrane-based, and cryogenic processes [3].

In contrast to adsorption and absorption, membrane-based carbon capture does not need a regeneration process, indicating potential capital and operating savings. Recent advancements in polymeric materials for gas separation, coupled with its straightforward plant design and operation, compactness, lightweight nature, and mobility, all combine to position membrane technology as a promising and feasible alternative for CO<sub>2</sub> capture [4].

Polymeric membranes present a range of advantages across diverse applications. Their straightforward fabrication simplifies technology adoption, ensuring accessibility and cost-effectiveness. Moreover, they exhibit reduced material requirements for module construction, offering both flexibility and cost savings. Additionally, their inherent processability empowers customization and adaptability, rendering them a versatile and appealing choice for various industries. For instance, the National Energy Technology Laboratory (NETL) has developed a rubbery thin film composite (TFC) membrane that exhibited high CO<sub>2</sub> permeance, high CO<sub>2</sub>/N<sub>2</sub> selectivity, and stable performance in the presence of water vapor and non-aging behavior [5].

Several types of membrane modules are investigated in experimental work for flue gas separation [6]. Researchers frequently employ plate-and-frame (flat sheet), tubular, and hollow fiber membrane modules. In particular, the plate-and-frame configuration offers the benefit of a low-pressure drop, which holds greater significance in high-volume, low-pressure applications. In addition, this kind of membrane is easy to clean and to replace. On the other hand, additional supports are required for the membrane, and it offers low area-to-volume ratios.

When compared to modeling approaches, conducting CFD simulations can be a cost-effective and efficient means to simultaneously study variable effects while ensuring accuracy in the outputs [7]. Specifically, CFD techniques provide the capability to analyze polymeric membranes under conditions that surpass the limits of experiments, enabling the generation of diverse trend analyses for better comprehension of the model results. CFD analysis proves valuable for assessing the viability and performance of specific membrane modules for industrial applications.

Regarding the implementation of the technology at process scale, it was concluded that a single-stage membrane process is unable to achieve high capture rates and high purity simultaneously [8,9]. Therefore, multi-stage membrane processes have been proposed to overcome this issue. In order to address the design problem using a holistic approach, membrane systems engineering has evolved as an important research area to understand the trade-offs among operating variables at plant scale, through mathematical modeling, simulations, and optimization techniques. In this context, superstructure optimization plays a crucial role in the formulation of the optimization problems since it involves creating a comprehensive network that includes all possible components, units, and interconnections relevant to the case study. Qi and Henson [10] addressed the separation of acid gases (CO<sub>2</sub> and H<sub>2</sub>S) from crude natural gas, by introducing a superstructure with potential network configurations obtained through a strategic arrangement of

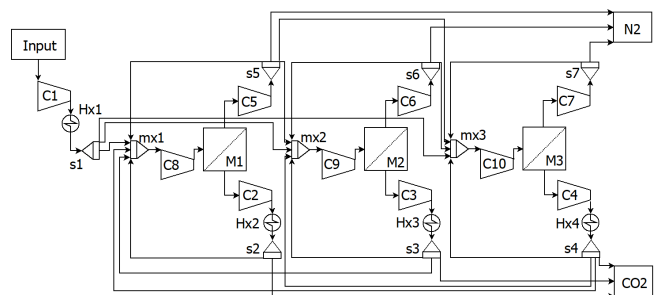
mixers, splitters, and compressors. Arias et al. [11] used mathematical programming and superstructure-based optimization to design multi-stage CO<sub>2</sub> capture membrane systems, to find the optimal configuration of membrane stages, areas, power requirements, recycle stream placement, and operational conditions.

In this work, we formulate a superstructure optimization for the optimal design of a membrane-based carbon capture process, where we use rigorous CFD simulations to build the models for the membrane modules.

## METHODOLOGY

Based on the literature [10,11], we formulate the superstructure shown in Fig. 1 for the optimal design of the multi-stage membrane process for carbon capture. In the present work, we demonstrate the potential of the technology using three separation stages. However, based on the application and the end goals, we could increase the number of stages as needed.

The input stream can be conditioned by increasing its pressure and/or a cooling operation. Then, the input stream could be sent to any of three possible membrane stages. For each membrane stage, the input stream, the retentate product, and the permeate product stream could be pressurized if this operation contributes to increasing the objective function. In addition, the permeate product could be cooled if required. It should be noted that this superstructure configuration embeds the potential process configurations, and the membrane order is not fixed beforehand. On the contrary, the order is determined based on the configuration that optimizes the objective function of the problem.



**Figure 1.** Superstructure for the optimal design of the multi-stage membrane process for carbon capture. M: membrane separation stage. C: compressor. Hx: heat exchangers. s: splitters. mx: mixers.

In order to model the superstructure of Fig. 1, we develop mathematical models for each equipment unit, and we consider the connections between them. Our modeling framework is mainly based on the literature [12,13] for the general process units, including the compressors, the heat exchangers, the mixers, and the splitters. For the calculation of the compound properties,

we fit polynomials based on the data from Aspen Plus, and we consider ideal gas thermodynamics calculations.

The objective function is the minimization of the capture cost, and it is defined as follows

$$CapCost = \frac{(\phi + \varpi) \cdot CAPEX + OPEX + M_{REP}}{F_{CO_2}} \quad (1)$$

where  $CAPEX$  and  $OPEX$  are the capital and operating costs, respectively;  $\phi$  is the capital recovery factor (0.154),  $\varpi$  is a factor to consider the annual maintenance cost (0.045),  $M_{REP}$  is the annualized replacement cost of the membrane modules, and  $F_{CO_2}$  is the flowrate of the captured  $CO_2$  in the plant. The  $CAPEX$  is calculated from the purchased equipment cost of individual equipment units, and we use the cost correlations reported in [14], while the capital costs are updated to 2023 dollars by using a Chemical Engineering Plant Cost Index (CEPCI) of 803.3.

Regarding the membrane modules, we formulate a model including mass and enthalpy balances for the membrane modules, and we consider the operation of them in parallel. To calculate the  $CO_2$  recovery and the purity product in the permeate we also develop surrogate models based on a rigorous CFD model, which is described in the next section.

### Rigorous CFD simulation

The representation of the CFD model for the membrane module is shown in Fig. 2. We consider a 3D geometry that includes two main parts: the retentate side and the permeate side. The input  $CO_2$ -rich stream enters the feed side, and carbon dioxide efficiently crosses through the selective boundary to the permeate side, where a concentrated  $CO_2$  stream is recovered. This technology utilizes thin film composite membranes consisting of rubbery polymer blends [5], without the inclusion of carriers.

We implement the CFD model in Comsol Multiphysics to describe the separation process using membranes. For this, we use the Navier-Stokes equation to calculate the velocity distribution, considering compressible laminar flow with Mach number less than 0.3, as shown in Eq. (2):

$$(\mathbf{u} \cdot \nabla) \mathbf{u} - \nu \nabla^2 \mathbf{u} + \frac{1}{\rho} \nabla p = 0 \quad (2)$$

where  $\mathbf{u}$  is the velocity,  $p$  is the pressure,  $\nu$  is the kinematic viscosity and  $\rho$  is the density.

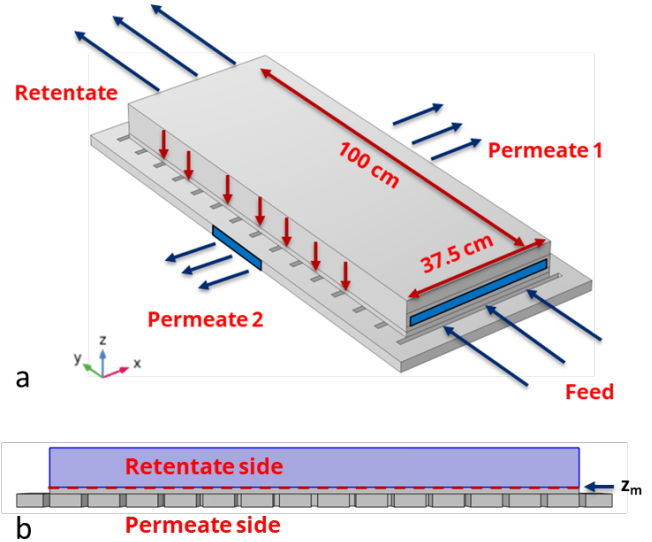
Regarding the mass transfer phenomena, we use the physics for transport of concentrated species, by considering the following expression:

$$\rho \mathbf{u} \cdot \nabla \omega_i + \nabla \cdot (-\rho D_i \nabla \omega_i) = 0 \quad (3)$$

where  $\omega_i$  is the mass fraction of species  $i$ , and  $D_i$  is the diffusion coefficient of species  $i$ .

For both retentate and permeate sides, we consider the coupled phenomena, which include the physics of the

laminar flow and the mass transport. However, these domains have different reference pressures since the retentate side operates near atmospheric pressure and the permeate side usually operates under vacuum. The driving force for the species flow through the selective boundary is the  $CO_2$  partial pressure difference; increasing this pressure difference increases the stage cut of the membrane process.



**Figure 2.** CFD model for the membrane module. a) Main configuration and flux directions. b) Domains and selective boundary

In order to model the membrane separation process, we define the selective boundary layer as a boundary condition for both physics. This selective layer is located at a height  $z_m$  in the membrane module, as shown in Fig. 2. The species fluxes ( $J_{x,i}, J_{y,i}, J_{z,i}$ ) are given by Eq. (3) in the transport phenomena, considering experimental data for the permeance of  $CO_2$  and  $N_2$  as pure gases ( $Q_i$ ). In this way, the modeling approach considers the different mass transfer resistances due to the membrane. Therefore, we do not have to include the information associated with the porosity of the support in the model. Regarding the velocity profile calculation, we consider a leaking wall as a boundary condition, calculating the equivalent velocity to satisfy momentum balances (see Eq. (4)).

$$z = z_m \rightarrow \begin{cases} J_{x,i} = 0 \\ J_{y,i} = 0 \\ |J_{z,i}| = Q_i M_i (p^{ret} x_i^{ret} - p^{per} x_i^{per}) \end{cases} \quad (3)$$

$$z = z_m \rightarrow \begin{cases} u_x = 0 \\ u_y = 0 \\ u_z = -\frac{1}{\rho} \sum_{i \in \{CO_2, N_2\}} J_{z,i} \end{cases} \quad (4)$$

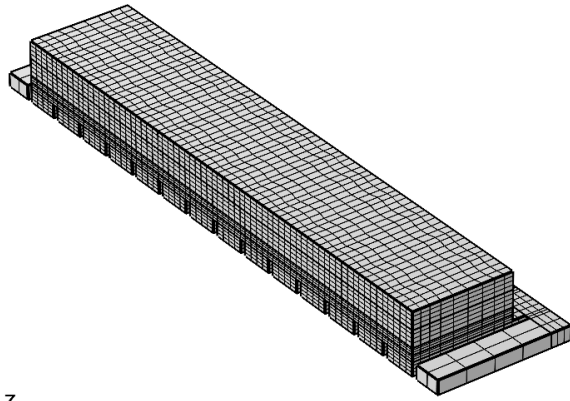
where  $Q_i$  is the permeance of species  $i$ , and  $M_i$  is the molar mass of species  $i$ .

Regarding the boundary conditions for the inlet and outlet streams, we use the conventional approach to specify them, as explained in [15]. In particular, we run the CFD model several times to generate data for the reduced model generation, as we explain below. In this matter, we consider inlet mass flowrates between 40-220 mg/s, inlet CO<sub>2</sub> concentrations between 0.1 and 0.7, retentate outlet pressures between 1.02-2 atm, and permeate outlet pressures between 0.09-0.6 atm.

### Model simplification and mesh studies

To enhance the computational tractability of the CFD model, a symmetry plane has been introduced for the membrane module. This modification involves the incorporation of a yz plane, resulting in a halved geometry, as depicted in Fig. 3. In this way, the number of cells required for the mesh is also halved.

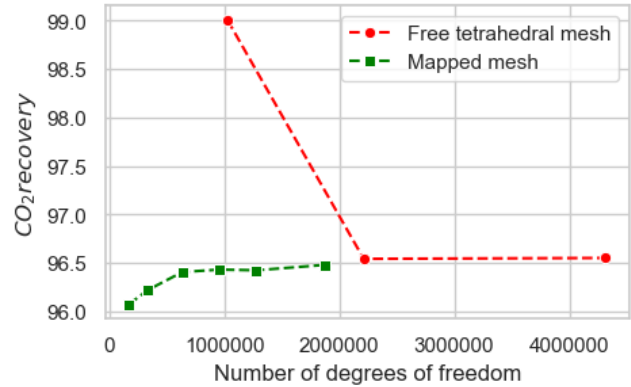
Regarding the mesh, we have developed a mapped mesh for the geometry as shown in Fig. 3. This meshing technique has shown a superior performance for mesh generation, offering the possibility of reducing the required number of cells to achieve accurate results.



**Figure 3.** Mesh for the membrane module (637,460 degrees of freedom)

For the present case study, we perform a mesh independence study to select an appropriate number of mesh elements. This analysis is shown in Fig. 4, where we show the influence of degrees of freedom, directly proportional to the number of mesh nodes, on the CO<sub>2</sub> recovery of the membrane module for the mapped mesh and the free tetrahedral mesh. The results indicate that as the mesh becomes finer, the CO<sub>2</sub> recovery decreases for the free tetrahedral mesh and increases for the mapped mesh. In addition, it is observed that the number of degrees of freedom to obtain accurate results is lower for the mapped mesh. It is worth noting that beyond 637,460 degrees of freedom, there is no significant change in the CO<sub>2</sub> recovery. The relative difference between this point and the one with the highest number

of nodes (which is also the most accurate) is only 0.076%. This suggests that a mesh containing 637,460 degrees of freedom provides sufficient computational accuracy.



**Figure 4.** Variation of the CO<sub>2</sub> recovery with the degrees of freedom in the mesh. Input flow: 1250 SCCM;  $x_{CO_2}$ : 0.3;  $p^{ret}$ : 1.2 bar;  $p^{per}$ : 0.2 bar.

### Surrogate model for the membrane process

The CFD for the membrane module was validated using experimental data, as described in the literature [16], for a bench-scale experimental implementation. In the present work, we have scaled up the membrane module to a length of 1 m and a width of 0.375, as shown in Fig. 2. Then, from this model, we generate sampling data and build a reduced model by using kriging-based methodology available in the IDAES computational framework [17]. This modeling approach considers a correlation function between different sampling points, which is usually called kernel. In the present case, we use the Gaussian kernel ( $k(\mathbf{x}_i, \mathbf{x}_j)$ ), which is shown in Eq. (5).

$$k(\mathbf{x}^i, \mathbf{x}^j) = \exp\left(-\sum_{r=1}^d \theta_r |x_r^i - x_r^j|^2\right) \quad (5)$$

where  $\mathbf{x}^i$  and  $\mathbf{x}^j$  are sampling points,  $\boldsymbol{\theta} = \{\theta_1, \theta_2, \dots, \theta_d\}^T$  is vector associated with the length scale hyperparameters, and  $d$  is the dimension of the input variable vector. The values of output variables are calculated through the following expression:

$$y_* = \mu + \mathbf{k}_*^T (\mathbf{K} + \sigma^2 \mathbf{I}_n)^{-1} (\mathbf{y} - \mathbf{1}\mu) \quad (6)$$

where  $y_*$  is the expected value of the output variable,  $\mu$  is the mean,  $n$  is the number of sample points,  $\mathbf{I}_n$  is the identity matrix of dimension  $n$ ,  $\sigma$  is the regularization parameter associated with the noise of the data and  $\mathbf{y}$  is the vector of the sampled output variable. The vector  $\mathbf{k}_*$  and the matrix  $\mathbf{K}$  are calculated through Eqs. (7) and (8), considering the Gaussian correlation function (see Eq. (5)).

$$[\mathbf{K}]_{i,j} = k(\mathbf{x}^i, \mathbf{x}^j) \quad (7)$$

$$[\mathbf{k}_*]_i = k(\mathbf{x}^i, \mathbf{x}^*) \quad (8)$$

In particular, we build reduced models that consider the



following input variables: i) inlet flowrate; ii) pressure in the retentate; iii) pressure in the permeate; iv) inlet CO<sub>2</sub> concentration. On the other hand, we consider the CO<sub>2</sub> recovery and the CO<sub>2</sub> purity in the permeate side as output variables for the model. In this way, we can exploit the information of the rigorous CFD simulation to create a reduced model, which is implemented in the superstructure for the selection of the multistage membrane process.

**Table 1.** Main statistics metrics of the Kriging model for the CO<sub>2</sub> recovery (output variable)

CO <sub>2</sub> recovery	Training data	Validation data
R <sup>2</sup>	1.0000	0.9992
Mean squared error	2.6·10 <sup>-6</sup>	4.7·10 <sup>-5</sup>
Mean rel. error (%)	0.2819	1.5116
Max. rel. error (%)	2.1905	5.9386

The main performance metrics for our built-reduced models are shown in Table 1. It is observed that the performance of the built reduced models is appropriate for their implementation in the optimization model since the R<sup>2</sup> coefficients are close to one, for both output variables. In addition, the statistics metrics for the training and the validation data do not change significantly, indicating a low risk of overfitting.

**Table 2.** Main statistics metrics of the Kriging model for the CO<sub>2</sub> purity in the retentate (output variable)

Purity	Training data	Validation data
R <sup>2</sup>	1.0000	0.9998
Mean squared error	4.8·10 <sup>-7</sup>	8.0·10 <sup>-6</sup>
Mean rel. error (%)	0.1266	0.4301
Max. rel. error (%)	0.7907	1.9357

## RESULTS

In the present work, we consider the optimal design of a plant to treat a flue gas of 30 % CO<sub>2</sub> -70% N<sub>2</sub> and a molar flow rate of 1000 mol/s, as a case study. This CO<sub>2</sub> concentration is representative of steel plants but our model can account for any set of inlet and operating conditions. We set the lower bounds for CO<sub>2</sub> recovery and the purity at 80 % and 90 %, respectively.

The mathematical programming model is implemented in Pyomo and solved using IPOPT [18]. As commercial solvers for global optimization turn out to be inefficient in solving the present problem design, we use a multi-start approach to increase the possibility of finding the global optimal solution.

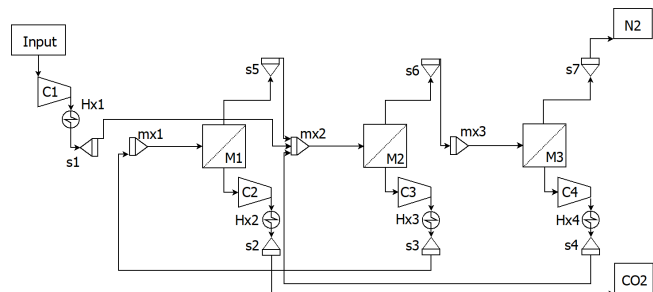
The optimization problem includes 539 variables and 50 optimization variables. We note that the use of

reduced models reduced the problem complexity significantly since the rigorous simulation of each membrane separation stage includes 637,460 variables. Therefore, solving the full optimization model for optimal design represents computational challenges, which would be the bottleneck of any decision-making. It is only through surrogate models that we can seek a practical design optimization in an acceptable timeline.

The main economic indicators are shown in Table 3. The objective function is the capture cost, for which we obtain a value of 45.4 \$/t-CO<sub>2</sub>. This value is of the expected order for a multi-stage membrane process [6]. The total annual cost for the capture plant is also shown in Table 1. The main contributor to this economic indicator is the annualized capital cost (51 %), while the operating cost represents 29 % of the total annual cost. Our results indicate that the capital cost associated with compressors accounts for 65 % of the CAPEX, while the membrane cost is 33%. It should be noted that the purity and the CO<sub>2</sub> recovery are at their lower bounds for the optimization problem.

**Table 3.** Economic indicators associated with the optimal design of the membrane-based carbon capture

	Value
Capture cost (\$/t-CO <sub>2</sub> )	45.69
Total annual cost (MM\$/y)	13.89
Capital cost (MM\$)	46.25
Operating cost (MM\$/y)	4.00
Purity (%)	90
CO <sub>2</sub> recovery (%)	80



**Figure 5.** Optimal configuration for the membrane process. M: membrane separation stage. C: compressor. Hx: heat exchangers. s: splitters. mx: mixers.

The optimal configuration for the membrane-based carbon capture process is shown in Fig. 5. It comprises a three-stage membrane separation setup. The inlet flue gas is introduced into the second membrane module, M2. The retentate product from M2 is fed to the third module, M3, where a N<sub>2</sub>-rich stream is recovered in the retentate stream, while the permeate side is recirculated back to M2. The permeate stream originating from M2 is then

sent to the first module, M1, which yields the purified CO<sub>2</sub> stream. The retentate product from M1 is also recycled to M2 for further processing.

It should be noted that the optimal network for the separation process using membranes is not trivial, and there is a trade-off between the recovery and purity in each separation stage. At process level, there are also trade-offs related to recycled streams; high recirculation rates increase the operating costs but they have the potential to reduce the required membrane area. In this context, the proposed superstructure optimization is a valuable tool for finding a solution, considering these process interactions in the design.

Regarding the operating condition and separation performance of the membrane stages, Table 4 shows the main results. It is observed that the membrane M2 includes a higher area than M1 and M3, and this result is associated with its high inlet flow (recycling streams for M1 and M3 are sent to M2 in the process network). In addition, we observe that M1 operates at high inlet CO<sub>2</sub> concentration to achieve the target purity of the CO<sub>2</sub> product (90 %), while M3 operates with a CO<sub>2</sub>-dilute stream (15 %) to recover the greenhouse gas before expelling the N<sub>2</sub>-rich stream from the retentate of this membrane stage. The permeate side for the membrane stages is at vacuum pressure, while the retentate is at atmospheric pressure. M2 and M3 operate at low pressure (0.1 atm) in the permeate side to promote CO<sub>2</sub> transport through the selective layer; while the pressure of the permeate side of M1 is 0.4 atm. For each membrane stage, we also report the dimensionless feed flow and the retentate recovery, which are associated with the capital cost and the energy requirement [19], respectively.

In order to verify the accuracy of the reduced models in the environment of the optimal solution, we have run the rigorous CFD simulation considering the input variables reported in Table 4 for each membrane stage. These input variables include the inlet flow, the inlet CO<sub>2</sub> molar fraction, and the pressure of the retentate and the permeate sides. The results associated with this analysis are shown in Table 5. In general, we obtained relative differences between the reduced model and the rigorous CFD simulations lower than 1 %. The sole exception is the prediction of the reduced model for the CO<sub>2</sub> purity for M3, which is 5 % lower than the value obtained by the CFD simulation. This means that the objective function of the optimization problem could slightly improve by updating the reduced model with this information. Future work will focus on using the trust-region filter method [20,21] to solve the problem and obtain highly accurate solutions

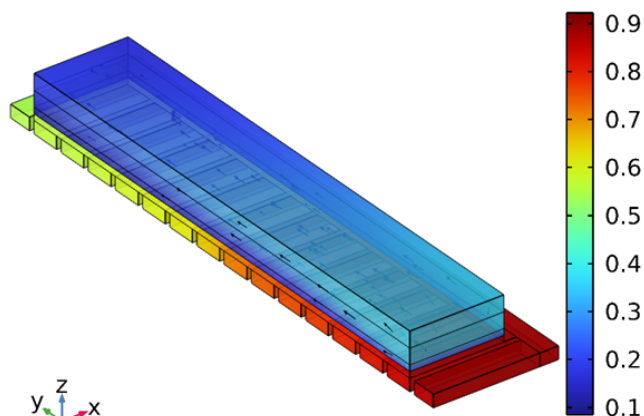
**Table 4.** Operating conditions and results for the membrane stages

	M1	M2	M3
Membrane area (m <sup>2</sup> )	30,849	107,800	54,887
Inlet flow (mol/s)	565	1465	901
Inlet CO <sub>2</sub> molar fraction	0.70	0.36	0.15
Retentate pressure (atm)	1.02	1.02	1.02
Permeate pressure (atm)	0.40	0.11	0.10
CO <sub>2</sub> recovery	0.61	0.74	0.55
CO <sub>2</sub> purity	0.90	0.70	0.42
Dimensionless feed flow	0.33	0.25	0.30
Retentate recovery	0.53	0.62	0.81

**Table 5.** Output variables from the rigorous CFD simulations and the NLP model (calculated from the Kriging approach), using the input variables shown in Table 4

	M1	M2	M3
CO <sub>2</sub> recovery (CFD model)	0.608	0.740	0.551
CO <sub>2</sub> recovery (NLP model)	0.614	0.745	0.552
CO <sub>2</sub> purity (CFD model)	0.906	0.701	0.445
CO <sub>2</sub> purity (NLP model)	0.900	0.700	0.423

The CO<sub>2</sub> molar fraction profile ( $x_{CO_2}$ ) of membrane stage M2 is shown in Fig. 6. Initially, the inlet  $x_{CO_2}$  is 0.36, while the retentate outlet  $x_{CO_2}$  is 0.15. The decrease in  $x_{CO_2}$  is attributed to the selective CO<sub>2</sub> flow through the membrane selective layer during the separation process. In the permeate region close to the feed,  $x_{CO_2}$  reaches high purities of up to 0.90. However, as we move closer to the permeate side near the retentate outlet, the CO<sub>2</sub> purity decreases to around 0.50. Consequently, the average CO<sub>2</sub> purity of the outlet permeate flow is estimated to be approximately 0.70.



**Figure 6.** Profile of CO<sub>2</sub> molar fraction ( $x_{CO_2}$ ) for M2. Input flow: 86.06 mg/s;  $x_{CO_2}$ :0.36;  $p^{ret}$ : 1.02 atm;  $p^{per}$ : 0.11 atm.

### Effect of the CO<sub>2</sub> concentration in the feed stream

We also explore the effects of changing the CO<sub>2</sub> molar fraction in the feed stream on key parameters, such as the objective function (capture cost), the membrane area, and the energy demand. For this purpose, we run the optimal design problem by considering different inlet CO<sub>2</sub> molar fractions in the interval of 0.2-0.3, while keeping the input molar flowrate of CO<sub>2</sub> constant.

The results associated with this analysis are shown in Table 6. In all of the cases, the purity and the CO<sub>2</sub> recovery were found at their lower bounds, 90 % and 80 %, respectively. It is observed that the capture cost decreases significantly as the inlet CO<sub>2</sub> molar fraction increases. These results are in agreement with the literature [14]. This trend can be explained by the simultaneous reduction in both membrane area and energy demand as the inlet CO<sub>2</sub> molar fraction rises, as detailed in Table 6.

**Table 6.** Effect of the inlet CO<sub>2</sub> concentration of the feed in the capture cost, membrane area, and energy demand

Inlet CO <sub>2</sub> molar fraction	Capture cost (\$/t-CO <sub>2</sub> )	Membrane area (m <sup>2</sup> )	Energy demand (kWh/t-CO <sub>2</sub> )
0.2	73.13	371,880	284.5
0.225	64.77	320,059	252
0.25	57.90	277,819	231.4
0.275	52.01	244,103	206.51
0.3	45.69	193,536	182.8

### CONCLUSIONS AND FUTURE WORK

In this work, we address the optimal design of a multi-stage membrane process for carbon capture. For this task, we develop a CFD model of a membrane module, from which we generate a Kriging reduced model to

be implemented in an optimization problem. In order to design the separation network, we formulate a superstructure to consider a three-stage membrane process. As the main case study, we consider an inlet flue gas stream with a CO<sub>2</sub> molar fraction of 30 % and a molar flowrate of 1000 mol/s. However, our computational framework is flexible and versatile and can therefore be adapted to accommodate various inlet and operating conditions to simulate any post-combustion or industrial plant and real-life gas stream compositions. Our work could also be extended to inform both new plants and to retrofit existing ones.

Our results show that the optimal design includes the three membrane stages and the capture cost of the process is 45.69 \$/t-CO<sub>2</sub>. Determining the resulting network configuration and optimal operating conditions for the membrane stages is a non-trivial task. In this context, the optimization formulation shows its value as it determines optimal trade-offs between the decision variables.

Future work will focus on including additional design variables associated with the membrane module, i.e., geometric variables. In addition, we will include the effect of temperature on the transport phenomena for the CFD model. Regarding the solution strategy, we will solve the problem by using trust region methods, to guarantee the optimality of the solution.

### ACKNOWLEDGEMENTS

This project was funded by the United States Department of Energy, National Energy Technology Laboratory, in part, through a site support contract. Neither the United States Government nor any agency thereof, nor any of their employees, nor the support contractor, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

### REFERENCES

1. National Oceanic and Atmospheric Administration. Broken record: Atmospheric carbon dioxide levels jump again. (2023)
2. Shu DY, Deutz S, Winter BA, Baumgärtner N,

- Leenders L, Bardow A. The role of carbon capture and storage to achieve net-zero energy systems: Trade-offs between economics and the environment. *Renew Sustain Energy Rev.* 178:113246 (2023)
3. Kazemifar F. A review of technologies for carbon capture, sequestration, and utilization: Cost, capacity, and technology readiness. *Greenh Gases Sci Technol.* 12(1):200–30 (2022)
  4. Pasichnyk M, Stanovsky P, Polezhaev P, Zach B, Šyc M, Bobák M, Jansen JC, Přebyl M, Bara JE, Friess K, Havlica J. Membrane technology for challenging separations: Removal of CO<sub>2</sub>, SO<sub>2</sub> and NO<sub>x</sub> from flue and waste gases. *Sep Purif Technol.* 323:124436 (2023)
  5. Zhu L, Kusuma V, Hopkinson D. Highly Permeable Thin-Film Composite Membranes of Rubbery Polymer Blends for CO<sub>2</sub> Capture. In United States; (2022).
  6. Chatziasteriou CC, Kikkinides ES, Georgiadis MC. Recent advances on the modeling and optimization of CO<sub>2</sub> capture processes. *Comput Chem Eng.* 165:107938 (2022)
  7. Sithamparam M, Lai LS, Tay WH. Computational fluid dynamics simulation for carbon dioxide gas transport through polydimethylsiloxane membrane with gyroid structure. *Mater Today Proc.* 46:1922–1928 (2021)
  8. Khalilpour R, Abbas A, Lai Z, Pinnau I. Modeling and parametric analysis of hollow fiber membrane system for carbon capture from multicomponent flue gas. *AIChE J.* 58(5):1550–61 (2012)
  9. Brunetti A, Scura F, Barbieri G, Drioli E. Membrane technologies for CO<sub>2</sub> separation. *J Memb Sci.* 359(1):115–25 (2010)
  10. Qi R, Henson MA. Optimal design of spiral-wound membrane networks for gas separations. *J Memb Sci.* 148(1):71–89 (1998)
  11. Arias AM, Mussati MC, Mores PL, Scenna NJ, Caballero JA, Mussati SF. Optimization of multi-stage membrane systems for CO<sub>2</sub> capture from flue gas. *Int J Greenh Gas Control.* 53:371–90 (2016)
  12. Pedrozo HA, Rodriguez Reartes SB, Vecchiotti AR, Diaz MS, Grossmann IE. Optimal Design Of Ethylene And Propylene Coproduction Plants With Generalized Disjunctive Programming And State Equipment Network Models. *Comput Chem Eng.* 149:107295 (2021)
  13. Dowling AW, Biegler LT. A framework for efficient large scale equation-oriented flowsheet optimization. *Comput Chem Eng.* 72:3–20 (2015)
  14. Hasan MMF, Baliban RC, Elia JA, Floudas CA. Modeling, Simulation, and Optimization of Postcombustion CO<sub>2</sub> Capture for Variable Feed Concentration and Flow Rate. 1. Chemical Absorption and Membrane Processes. *Ind Eng Chem Res.* 51(48):15642–64 (2012)
  15. Comsol Multiphysics®. User's Guide. Version 6.2 (2023).
  16. Dosso C, Zhu L, Kusuma VA, Hopkinson D, Biegler LT, Panagakos G. CFD Modeling of High-Flux Plate-and-Frame Membrane Modules for Post-Combustion Carbon Capture. In: AIChE Annual Meeting. Orlando, FL; (2023)
  17. Lee A, Ghouse JH, Eslick JC, Laird CD, Sirola JD, Zamarripa MA, Gunter D, Shinn JH, Dowling AW, Bhattacharyya D, and Biegler LT. The IDAES process modeling framework and model library—Flexibility for process simulation and optimization. *J Adv Manuf Process.* 3(3):e10095 (2021)
  18. Wächter A, Biegler LT. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math Program.* 106(1):25–57 (2006)
  19. Sun L, Panagakos G, Lipscomb G. Effect of Packing Nonuniformity at the Fiber Bundle-Case Interface on Performance of Hollow Fiber Membrane Gas Separation Modules. *Membranes* 12(11):1139 (2022)
  20. Eason JP, Biegler LT. Advanced trust region optimization strategies for glass box/black box models. *AIChE J.* 64(11):3934–3943 (2018)
  21. Pedrozo HA, Panagakos G, Biegler LT. Including CFD rigorous models in the optimal design of carbon capture plants through trust-region methods. *Chemical Engineering Science.* 286, 119646 (2024)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Enhancing PHAs Production Sustainability: Biorefinery Design through Carbon Source Diversity

Fernando D. Ramos<sup>a,b</sup>, Matías H. Ramos<sup>a,b</sup>, Vanina Estrada<sup>a,b</sup> and M. Soledad Diaz<sup>a,b\*</sup>

<sup>a</sup> Planta Piloto de Ingeniería Química (PLAPIQUI CONICET-UNS), Camino La Carrindanga km. 7, Bahía Blanca, Argentina

<sup>b</sup> Departamento de Ingeniería Química, Universidad Nacional del Sur (UNS), Bahía Blanca, Argentina

\* Corresponding Author: [sdiaz@plapiqui.edu.ar](mailto:sdiaz@plapiqui.edu.ar).

## ABSTRACT

In this work, we propose a Mixed Integer Nonlinear Programming (MINLP) model to determine the optimal sustainable design of a poly(hydroxyalkanoate)s (PHAs) production plant configuration and its heat exchanger network (HEN). The superstructure-based optimization model considers different carbon sources as raw material: glycerol (crude and purified), corn starch, cassava starch, sugarcane sucrose and sugarcane molasses. The PHA extraction section includes four alternatives: the use of enzymes, solvent, surfactant-NaOCl or surfactant-chelate. Model constraints include detailed capital cost for equipment, mass and energy balances, product specifications and operating bounds on process units. To assess the feasibility of the PHA plant, we considered the Sustainability Net Present Value (SNPV) as the objective function, a multi-criteria sustainability metric that considers economic, environmental and social pillars. The Net Present Value (NPV) was also calculated. SNPV metric provides useful insights on sustainable PHA production, as the optimal technological route results in the sugarcane-surfactant-chelate option, rather than the sugarcane-enzyme pathway which proves more economically profitable, but with higher environmental impacts. Moreover, inclusion of HEN design significantly improves the objective function value, mainly due to a 24% carbon footprint impact reduction.

**Keywords:** Biomass, Environment, Modelling, Optimization, Process Design

## INTRODUCTION

Poly(hydroxyalkanoate)s (PHAs) constitute a family of biodegradable and biocompatible polymers emerging as promising alternatives to substitute conventional synthetic plastics such as polypropylene and polyethylene [1]. PHAs can be intracellularly stored as a carbon and energy reservoir in the cytoplasm of different microorganisms capable of metabolizing various substrates [2].

Koller and Mukherjee (2022) [3] documented the current global companies actively engaged in the commercialization of PHAs. However, its industrial scale production is constrained by high production costs mainly attributed to the feedstock used as carbon source [4]. Consequently, competing with petroleum-based plastics market price becomes challenging. Nevertheless, embracing an integral process optimization framework [5] and a circular economy action plan emerges as a favorable strategy for biopolymers, envisioning sustainability by

addressing not only economic aspects but also environmental and social considerations.

Previous work shows that employing sugarcane as carbon source for PHAs production and enzymes for PHAs extraction, results in a biopolymer production cost round to 3 USD/kg of PHA [6]. In this context, an extension of this study focusing on sustainability design, could offer valuable insights on whether the chosen carbon source and extraction method align with optimal sustainable approaches for PHAs production, as well as simultaneously design the HEN.

In this work, a superstructure is presented for the simultaneous optimal sustainable design of a PHA production process and its heat exchanger network (HEN). The objective function is the maximization of the sustainability net present value (SNPV) proposed by Zore *et al.* (2018) [7]. The model has into account several alternative carbon sources as substrates and includes different technological pathways for the biosynthesis, extraction



and purification stages.

## PROCESS DESCRIPTION

The PHA production process has three main stages: upstream processing (raw material pretreatment), biosynthesis and downstream processing (biopolymer extraction and purification). Figure 1 presents the proposed superstructure, including different technological alternatives. For a more comprehensive description of the technological pathways, refer to Ramos *et al.* (2019) [6].

### Upstream processing

Three different raw materials and some of their derivatives are considered as carbon source, namely, glycerol, starch and sugarcane. Raw glycerol can be employed as an economical carbon source for several microbiological processes such as PHAs production [8]. Crude glycerol can be fed to the fermentation step after being sterilised, or it can be purified to improve downstream biomass growth. If glycerol purification step is selected, methanol can be sold as a co-product after recovery. It is worth mentioning that selling purified glycerol also is set as a possible alternative, instead of being used as carbon source for PHAs production.

Starch is one agro-industrial product that holds a promising outlook as substrate for biomaterial production. Three potential alternatives are considered for starch production in the superstructure presented in Fig. 1. The first one consists of including the process of starch production from corn. The second option involves the direct purchase of corn starch, while the third one contemplates the possibility of using cassava starch [9]. The superstructure also includes the possibility of producing

glucose, through starch liquefaction and saccharification, to be used as substrate for the microbial biosynthesis in the PHA production stage.

The last alternative is the use of sucrose from sugarcane as a carbon source for the production of biopolymers [10]. In this sense, sugarcane processing is embedded within the proposed superstructure for PHA production. Sugarcane bagasse constitute an important residue from the sugar industry, which can be processed for the production of thermal and electrical energy. In the proposed superstructure, a typical electrical energy production system is considered as a destination for this sub-product [11]. As alternatives to the processing of sugarcane, the possibility of directly buying processed sucrose or sugarcane molasses is included in the superstructure [12].

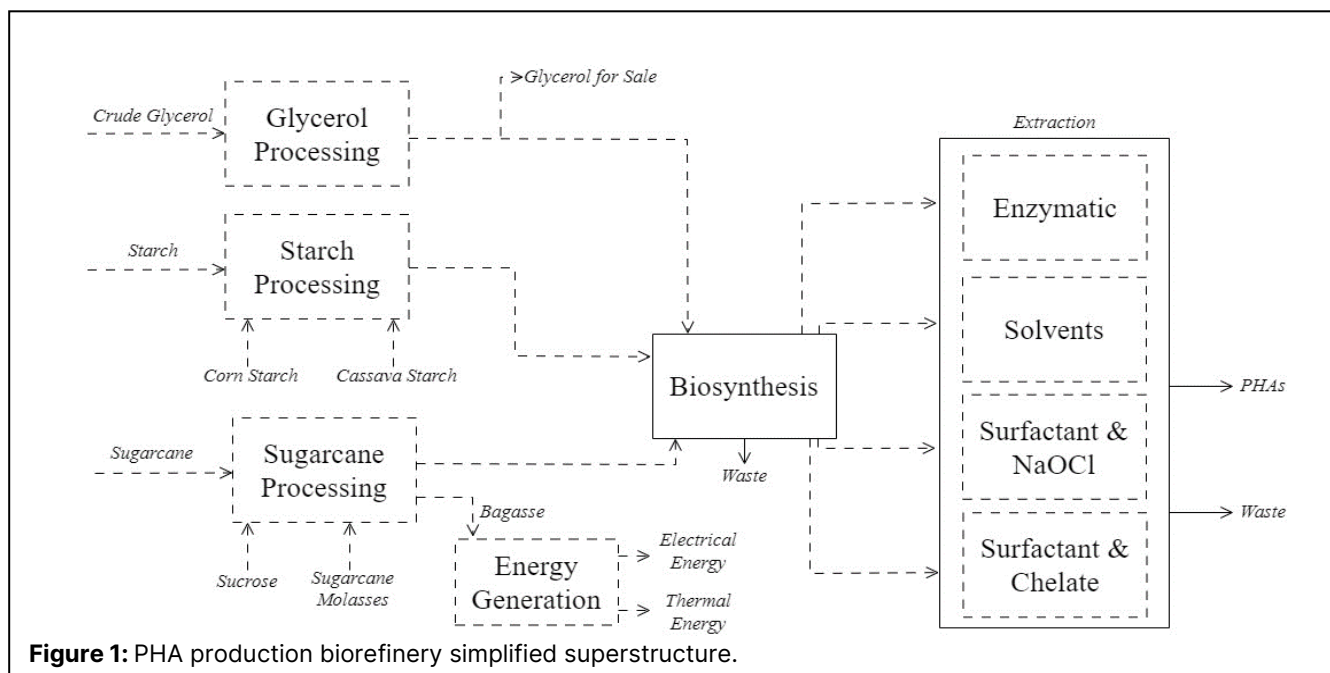
### Biosynthesis

PHA production takes place in the biosynthesis stage by the biopolymer intracellular accumulation in a microbial strain, through an excess supply of a carbon source and the limitation of another growth essential nutrient like nitrogen or phosphorus.

According to the selected carbon source, several operational alternatives are presented for the bioreactors used in this stage of the production process. Available technologies involve two fermentation stages. The first stage is included for biomass optimal growth without nutrient limitation imposed, while in the second, PHAs production takes place and its accumulation is triggered by the limitation of essential nutrient sources.

### Downstream processing

The extraction of biopolymers from the



**Figure 1:** PHA production biorefinery simplified superstructure.

microorganism's cytoplasm is a crucial step in the production process, as its quality directly impacts the final product. Therefore, an appropriate selection of the extraction technology is essential to establish a viable process in terms of sustainability. As shown in Fig. 1, four extraction alternatives are included in the PHA production process superstructure, the use of enzymes, solvent, surfactant- sodium hypochlorite (NaOCl) or surfactant-chelate.

The enzymatic extraction alternative is based in a digester, where the cell wall lysis is carried out by the use of enzyme pancreatin in a 2 wt % [13]. The process that employs solvent for PHA extraction uses diethyl succinate (DES) and includes a homogenizer, which operates at a high pressure to enable cell lysis [14]. The third option involves the use of NaOCl and a surfactant for biopolymer extraction. The last technological route embedded in the superstructure comprises surfactant and chelate as extraction agents. This method is based on the process proposed by Chen *et al.* (2001) [15], and is presented as a promising alternative because it is environmentally friendly having a higher quality of the final product and minimizing the requirement of chemicals.

## MATHEMATICAL MODEL

The proposed superstructure is formulated as a mixed-integer nonlinear programming (MINLP) problem and implemented in GAMS 35.2.0 [16] in order to determine the optimal sustainable design of a PHAs production plant and its heat exchanger network (HEN). The objective function is to maximize the monetary-based sustainability metric, Sustainability Net Present Value (SNPV) [7]. The proposed superstructure includes mass and energy balances, equipment design equations and capital cost correlations for the PHAs production plant processes, as well as its HEN design and connection equations to link process design variables with HEN variables. Binary variables are associated to potential units and to heat exchanger matches.

Nonreactive units mass balances within the superstructure are formulated as shown in Eq. (1).

$$\sum_{u1 \in UI_u} f_{u1,u,j} = \sum_{u2 \in UO_u} f_{u,u2,j} \quad j \in J, u \in U^{NR} \quad (1)$$

where  $f_{u1,u,j}$  is the flowrate from unit  $u1$  to  $u$  of component  $j$ ;  $UI_u$  and  $UO_u$  are subsets that define the connection between units. Nonreactive units splitting are represented by Eq. (2) as it is shown below.

$$f_{u,u2,j} = \alpha_{u,u2,j} \cdot \sum_{u1 \in UI_u} f_{u1,u,j} \quad j \in J, u \in U^{NR}, u2 \in UO_u \quad (2)$$

where  $\alpha_{u,u2,j}$  is a splitting parameter for component  $j$  that

leaves unit  $u$  and leads to unit  $u2$ . It is worth mentioning that Eq. (3) should be verified in order to fulfill mass conservation.

$$\sum_{u2 \in UO_u} \alpha_{u,u2,j} = 1 \quad j \in J, u \in U^{NR} \quad (3)$$

Reactive unit mass balances are described by Eq. (4).

$$\sum_{u2 \in UO_u} f_{u,u2,j} = \sum_{u1 \in UI_u} f_{u1,u,j} + \sum_{r \in R_u} \left( \xi_{u,r,j} \sum_{u1 \in UI_u} f_{u1,u,r} \right) \quad (4)$$

$j \in J, u \in U^R$

where  $\xi_{u,r,j}$  denotes the mass coefficient for component  $j$  from the main reactant  $r$  in reactive unit  $u$ . This coefficient is positive for products and negative for reactants. Reactive unit splitting is represented by Eq. (5) as follows.

$$f_{u,u2,j} = \alpha_{u,u2,j} \cdot \left[ \sum_{u1 \in UI_u} f_{u1,u,j} + \sum_{r \in R_u} \left( \xi_{u,r,j} \sum_{u1 \in UI_u} f_{u1,u,r} \right) \right] \quad (5)$$

$j \in J, u \in U^R, u2 \in UO_u$

The electrical energy consumption ( $EC_u$ ) is calculated through linear functions, as shown in Eq. (6).

$$EC_u = ECR_u \cdot \left( \sum_{u1 \in UI_u} \sum_{j \in J} f_{u1,u,j} \right) \quad u \in \{U^{NR} \cup U^R\} \quad (6)$$

where  $ECR_u$  is the energy consumption ratio per unit of mass flowrate relative to unit  $u$ . Conditional units proposed in the superstructure are associated with binary variables through Big-M formulations in order to determine the optimal technological route. The corresponding mixed integer constraints are formulated as it is shown in Eq. (7)

$$\sum_{u2 \in UO_u} \sum_{j \in J} f_{u,u2,j} - BM \cdot y_u \leq 0 \quad u \in U^C \quad (7)$$

where  $y_u$  is the binary variable related to the selection of conditional unit  $u$ ; and BM is a big M parameter [28].

Regarding equipment design equations and capital cost correlations, we include standard capital cost models from the literature [17-18].

The objective function Sustainability Net Present Value (SNPV) presented by Zore *et al.* (2018) [7], includes economic ( $NPV^{Economic}$ ), environmental ( $NPV^{Environmental}$ ) and social ( $NPV^{Social}$ ) net present values (Eq. (8)). It is worth mentioning that, in order to assess sustainability from a combined point of view of the company and government, we consider the SNPV macroeconomic

perspective [7].

$$SNPV = NPV^{Economic} + NPV^{Environmental} + NPV^{Social} \quad (8)$$

The economic pillar of the objective function is represented by  $NPV^{Economic}$  and is calculated as shown in Eq. (9). The annual cash flows are discounted from its year to the present time by the use of a corresponding factor that involves the discounted rate,  $i=10\%$ , and the project lifetime,  $n=20$  years.  $R$  accounts for annual revenues from products and by-products sales,  $E$  includes annual manufacturing and general costs.  $I$  is the total capital investment cost for a turnkey plant and involves the sum of equipment cost, considering tax contingency factor ( $\varphi_1=1.18$ ) and a grass-root factor ( $\varphi_2=1.3$ ), land cost, piping and instrumentation cost, and the working capital. In addition,  $V_{rec}$  is the recovery value, an income estimation from selling the equipment and recovering the working capital at the end of the project lifespan.

$$NPV^{Economic} = (R - E) \cdot \frac{(1+i)^n - 1}{i \cdot (1+i)^n} - I + V_{rec} \cdot (1+i)^n \quad (9)$$

Furthermore,  $NPV^{Environmental}$  stands for the environmental pillar of sustainability and is defined by Eq. (10.1). The annual environmental benefit ( $EB$ ) represents, in monetary terms, the unburdening effect of raw materials, products, technologies, transport, energy and waste on the environment (Eq. (10.2)). On the other hand, the annual environmental cost ( $EC$ ) corresponds to their burdening effect on the environment (Eq. (10.3)). Both terms are calculated by the eco-cost coefficients [7],  $c_{i,tech}$  and  $c_{j,tech}$ , for annual raw materials and products, respectively. Superindexes  $R^{UNB}$  and  $R^B$  denote the unburdening and burdening effect of the annual raw materials required by process  $tech$ , respectively. Similarly,  $P^{UNB}$  and  $P^B$  denote the unburdening and burdening effect of the annual products produced by unit  $tech$ , respectively. These environmental effects are proportional to annual raw material and product mass flowrates,  $q_{m_i}$  and  $q_{m_j}$ , respectively. Also, a substitution factor,  $f_{j,tech}^{S/P^{UNB}}$ , is taken into account to represent the ratio between the annual amount of product to be substituted and the annual amount of product obtained in the bioprocess.

$$NPV^{Environmental} = (EB - EC) \cdot \frac{(1+i)^n - 1}{i \cdot (1+i)^n} \quad (10.1)$$

$$EB = \sum_{tech \in T} \sum_{ieR^{UNB}} f_{i,tech}^{R^{UNB}} \cdot c_{i,tech}^{R^{UNB}} + \sum_{tech \in T} \sum_{jeP^{UNB}} f_{j,tech}^{P^{UNB}} \cdot S f_{j,tech}^{S/P^{UNB}} \cdot c_{j,tech}^{P^{UNB}} \quad (10.2)$$

$$EC = \sum_{tech \in T} \sum_{ieR^B} f_{i,tech}^{R^B} \cdot c_{i,tech}^{R^B} + \sum_{tech \in T} \sum_{jeP^B} f_{j,tech}^{P^B} \cdot c_{j,tech}^{P^B} \quad (10.3)$$

Referring to the environmental pillar of the objective function, it is important to note that environmental burden can be classified into a set of impact categories as midpoints or endpoints. The former are considered to be links in the cause-effect chain of an impact category at which characterization factors can be derived to reflect the relative importance of emissions or extractions [19]. Midpoints that are considered in this framework are global warming (*Globwarm*), acidification (*Acid*), eutrophication (*Eutro*), ecotoxicity (*Ecotox*), photochemical oxidants (*Sums*), human toxicity carcinogens (*Htoxc*), human toxicity non carcinogens (*Htoxcnc*), fine dust (*Fined*), metals scarcity (*Mets*), fossil products (*Foss*), uranium (*Ura*), land-use (*Lanu*) and baseline water stress (*Basew*). On the other hand, the endpoints are indicators that show the environmental impact on higher aggregation levels, being considered carbon footprint (*Globwarm*), ecosystems (*Acid*, *Eutro*, *Ecotox*), human health (*Sums*, *Htoxc*, *Htoxcnc*, *Fined*) and resource scarcity (*Mets*, *Foss*, *Ura*, *Lanu*, *Basew*).

The social term of the objective function,  $NPV^{Social}$ , is represented by Eq. (11.1) and considers social security contributions ( $SS$ ), social unburdening effect due to new jobs creation ( $SU$ ) and social cost ( $SC$ ). Social security contributions (Eq. (11.2)) are defined as the difference between gross and net annual salaries ( $S^{gross}$  and  $S^{net}$ , respectively) for the number of jobs created ( $N^{jobs}$ ). This variable is included in the model through the manufacturing cost term, which takes into account direct costs such as operating labor associated with the attention required by an operator to run the equipment [18]. Social unburdening effect is described by Eq. (11.3) and contemplates the state social transfer needed to annually support an unemployed person ( $C_s^{UNE,state}$ ) who is now employee ( $N^{jobs}$ ). Eq. (11.4) describes the annual social support of the state and the company for each employee ( $N^{jobs}$ ) considering a state social transfer ( $C_s^{EMP,state}$ ) and a company social charge ( $C_s^{EMP,company}$ ) parameter.

$$NPV^{Social} = (SS + SU - SC) \cdot \frac{(1+i)^n - 1}{i \cdot (1+i)^n} \quad (11.1)$$

$$SS = N^{jobs} \cdot (S^{gross} - S^{net}) \quad (11.2)$$

$$SU = N^{jobs} \cdot C_s^{UNE,state} \quad (11.3)$$

$$SC = N^{jobs} \cdot (C_s^{EMP,state} + C_s^{EMP,company}) \quad (11.4)$$

In this work, we address a simultaneous design of the bioprocess and its heat exchanger network (HEN),

following Pedrozo *et al.* [17]. It is worth mentioning that, since we performed a simultaneous process and HEN design, heat balance equations at each stage are nonlinear, as both process stream flowrates and stage temperatures are variables, as presented in the previous publication. The heat exchanger network capital cost ( $HENC$ ) is calculated as shown in Eq. (12), while the total utility cost ( $TUC$ ) is given by Eq. (13), as follows,

$$HENC = \sum_{k \in ST} \sum_{c \in C} \sum_{h \in H} (\alpha_0 z_{h,c,k} + HX_{h,c,k}) + \sum_{h \in H} (\alpha_0 z_h^{CU} + HX_h^{CU}) + \sum_{c \in C} (\alpha_0 z_c^{HU} + HX_c^{HU}) \quad (12)$$

$$TUC = \beta_{CU} \sum_{h \in H} q_h^{CU} + \beta_{HU} \sum_{c \in C} q_c^{HU} \quad (13)$$

where  $\alpha_0$  corresponds to fixed cost associated with heat exchanger units;  $HX_{h,c,k}$  is the heat exchanger cost corresponding to the hot stream  $h$  and cold stream  $c$  match at stage  $k$ ;  $HX_h^{CU}$  and  $HX_c^{HU}$  are the heat exchanger cost of the hot stream  $h$  and cold stream  $c$  with their respective external utilities;  $z_{h,c,k}$  is the binary variable that is equal to one when the hot stream  $h$  matches the cold stream  $c$  at stage  $k$  and equal to zero otherwise;  $z_h^{CU}$  and  $z_c^{HU}$  are binary variables that are equal to one when the hot stream  $h$  and the cold stream  $c$  require extreme external utilities, respectively, and equal to zero otherwise.  $\beta_{CU}$  and  $\beta_{HU}$  stand for the cooling utility and the heating utility costs, respectively.  $q_h^{CU}$  and  $q_c^{HU}$  are cooling and heating utilities for hot and cold streams, respectively. It is worth noting that the continuous elements of the model, involves mass flowrates, equipment cost, sustainability indices, and energy variables. Conversely, the discrete component comprises unit selection variables ( $y_u$ ) and binary variables involved in the HEN model ( $z_{h,c,k}$ ,  $z_h^{CU}$  and  $z_c^{HU}$ ). For further details regarding HEN design, please refer to Pedrozo *et al.* [17].

## NUMERICAL RESULTS AND DISCUSSION

The MINLP model formulated for sustainability optimization and simultaneous process and HEN design includes 5,412 discrete variables, 30,228 continuous variables and 41,632 constraints. It was solved to an objective function value of  $SNPV = -111.40$  M USD using DICOPT, with CONOPT and CPLEX as nonlinear and linear subsolvers, respectively [20]. The optimal technological route in terms of sustainability, includes the use of sugarcane (0.21 Mt/y) as carbon source for biopolymer production, selecting the surfactant-chelate pathway as the PHA extraction method. A PHA capacity is fixed at 10,000 t/y is achieved, which is similar to numerous currently industrial production plants [21].

A conventional techno-economic assessment was

carried out for the optimal configuration, considering an interest rate of 10 % and a project lifetime of 20 years, providing a Net Present Value (NPV) of 131.93 M USD, which is consistent with the results already presented in Ramos *et al.* (2019) [6].

With respect to the SNPV value, a deeper analysis for comparing purposes was performed, considering four different suboptimal technological routes. Table 1 shows for the different alternatives, the SNPV and NPV values. The alternatives were based on sugarcane (*sc*) or sugarcane molasses (*m/s*) as carbon source, as these resulted to be the more sustainable options. Also, the extraction methods for PHAs that prove to be the more sustainable, were enzymatic extraction (*enz*), NaOCl and surfactant extraction (*surf*) and surfactant and chelate as extraction (*surf & ch*).

Alternative 1, which is based on sugarcane as a carbon source and whose extraction method is enzymatic, yielded more economical profitability in comparison with the optimal configuration. Alternative 2, also based on sugarcane as carbon source, with surfactant and NaOCl as option for PHA extraction method, resulted in nearly the same economical profitability in comparison to the optimal scheme. Although the economical profitability is higher or nearly the same for alternatives 1 and 2, respectively, there is a significant detriment in the SNPV value. This is mainly due to the ecotoxicity environmental impact of NaOCl emissions [22], chemical that is used for dissolution improvement of the cell wall for the enzymatic extraction in alternative 1, and for chemical digestion in surfactant and NaOCl extraction in alternative 2. Considering alternatives 3 and 4, where is employed sugarcane molasses as carbon source, there is a worsening in both SNPV and NPV. Hence, processing of sugarcane leads to enhanced sustainability in PHAs production plant design.

**Table 1:** Set of potential technological routes for PHAs production.

Alternative	SNPV (MUSD)	VPN (MUSD)
Optimal	-111.40	131.93
1 ( <i>sc</i> + <i>enz</i> )	-217.86	187.88
2 ( <i>sc</i> + <i>surf</i> )	-608.76	131.69
3 ( <i>m/s</i> + <i>enz</i> )	-340.43	117.31
4 ( <i>m/s</i> + <i>surf &amp; ch</i> )	-244.06	58.40

Regarding heat integration, additional runs were carried out without heat integration and numerical results show that the optimal integrated scheme with HEN increased the number of heat exchangers required, as compared to the optimal base case without HEN (16 vs 15) and it also reduced the cost of external thermal utilities by 35.60 % (from  $123.65 \times 10^6$  kJ/h to  $79.64 \times 10^6$  kJ/h). Moreover, the reduction of utilities matches from 15 to 13 accounts for significant utility savings that improve the sustainability objective function (see Table 2), that led to

an improvement of 48.58 % in the objective function. This aligns with the results reported by Ramos *et al.* (2023) [27].

According to the results shown in Table 2, a meaningful enhancement of 15.14 % is found in the economic pillar between the two proposed cases. This is due to the energy savings, that result economically more impactful than the increase of the heat exchanger capital cost (1 extra heat exchanger for HEN case). Furthermore, there is a significant improvement of 19.42 % in the environmental pillar, while only a difference of 0.89 % can be reached in the social pillar. Regarding the economic pillar, a further analysis demonstrates that, while the utility cost reduction with heat integration represented a 34.14 % (from 8.46 M USD to 5.57 M USD), the heat exchanger capital cost increased in a 19.72 % (from 1.20 M USD to 1.43 M USD). The social pillar difference is associated with the increase in the number of heat exchangers after heat integration, representing more workers' contracts. A decrease of fossil fuels requirements (external utilities), positively benefits the environmental pillar in the objective function.

**Table 2:** Sustainability Net Present Value pillars comparison for heat and non-heat integration scenarios.

Sustainability metric (MUSD)	With HEN	Without HEN	Difference
SNPV	-111.40	-216.64	48.58 %
NPV <sub>economic</sub>	208.11	180.74	15.14 %
NPV <sub>environmental</sub>	-322.91	-400.75	19.42 %
NPV <sub>social</sub>	3.39	3.36	0.89 %

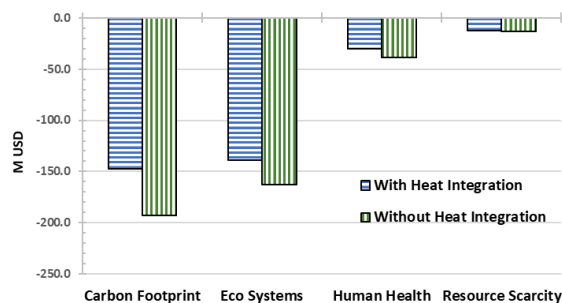
In order to provide a comprehensive insight into the environmental pillar of the objective function, Fig. 2 displays a breakdown in endpoint impact categories, for the heat and non-heat integration scenarios, both for the optimal configuration of the PHAs production plant.

The highest environmental burden is attributed to carbon footprint, as a consequence of the thermal energy required in the surfactant and chelate extraction method, which is the extraction method with the higher thermal energy demand. Significant reductions in the carbon footprint impact category can be achieved with heat integration, as it reduces this burden by 23.65 %. This reduction is mainly attributed to the lower thermal energy consumption required at the biorefinery, leading to less energy production from conventional sources.

In the case of eco systems environmental burden, impact that follows carbon footprint in intensity, the major contribution to this category resulted from the background environmental impact of the acquisition of raw sugarcane. This background accounts for all the environmental burdens related to eco systems category, primarily linked to ecotoxicity and acidification impacts, that were generated in order to bring along the carbon source

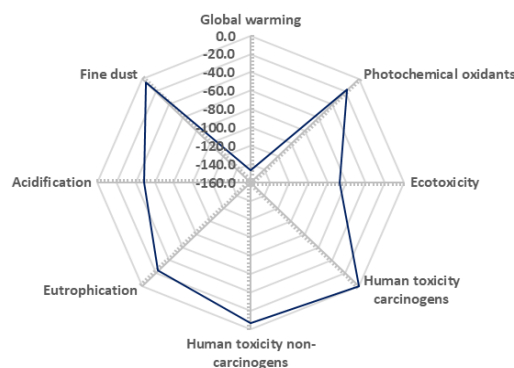
into the Biorefinery.

Human health and resource scarcity resulted in the least negatively impacted categories, where the minor contributions were the background process environmental burden from energy requirements. Overall, numerical results reinforce the need for including simultaneous HEN models within the superstructure framework [23].



**Figure 2.** SNPV environmental pillar breakdown.

In order to provide major insight into environmental impacts due to foreground and background processes emissions, Fig. 3 presents the midpoints radar chart expressed in M USD, for the optimal configuration of the biorefinery (HEN scenario).



**Figure 3.** Midpoints radar chart for heat integration case.

Also, we performed a local sensitivity analysis in order to determine most influential parameters on SNPV, that could be improved to achieve a higher sustainability profit on the biorefinery, considering a variation of  $\pm 10\%$  from base value of the parameters (Figure 4). It can be seen that PHA's selling price resulted in the parameter with the largest effect on SNPV. For instance, a 10 % increase in this parameter (from 5.46 USD/kg<sub>PHA</sub> [24] to 6.01 USD/kg<sub>PHA</sub>) represents an objective function enhancement of 41.74 % (from -111.40 M USD to -64.90 M USD). On sensitivity effects, it is followed by PHA yield, CEPCI index, sugarcane raw material price.

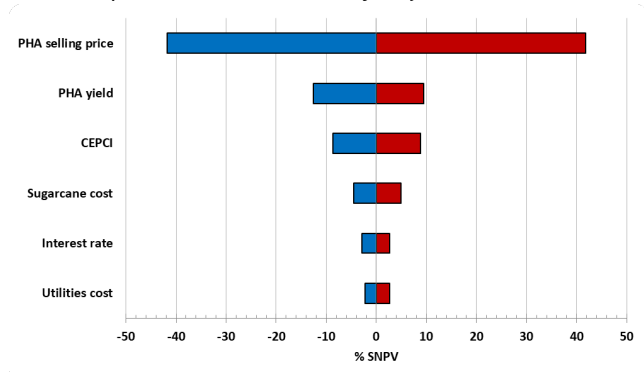
PHA yield [11] resulted in a meaningful effect on SNPV, mainly due to the fact that, if this yield is higher in the biosynthesis section, this leads to a minor requirement of sugarcane as carbon source in order to achieve



the PHA demand. As it was aforementioned, sugarcane background processes' environmental burden resulted in a major contribution in the ecosystems impact category. Also, reduction of sugarcane results in a decrease of capital costs.

It is worth noting that SNPV is not particularly sensitive to interest rate and utility service's cost. This last outcome, is due to the fact that the major impact on SNPV for the optimal configuration of the biorefinery, is the environmental pillar. Hence, the external utilities cost is not as relevant as diminishing the utilities requirements.

Based on these results, some of the above-mentioned parameters can be modified to potentially improve plant economics as proposed by Aui *et al.* [25] and Rizwan *et al.* [26], and subsequently enhance the proposed objective function for the biorefinery. However, some of these parameters are intrinsically connected with the market, such as product selling price, CEPCI index or interest rate. Therefore, special attention should be devoted to the technology-dependent parameters, such as the biological parameter PHA yield, since it is crucial to improve the sustainability objective function.



**Figure 4.** Sensitivity analysis for the optimal scheme.

## CONCLUSIONS

In this work, we have addressed the production of biopolymers, namely PHAs, by a mixed integer nonlinear programming (MINLP) model for the optimal sustainable design of a biorefinery and its heat exchanger network (HEN). A superstructure that considered the processing of different carbon sources and PHAs extraction methods was developed. Optimization results point out that the optimal technological route is the one that employs sugarcane as a carbon source for PHAs production and surfactant-chelate for PHAs extraction. Heat integration was proved to be valuable since it enhanced significantly the SNPV objective function. Also, various technological alternatives were explored, and although they exhibited economic parity or superiority, they were notably inferior in terms of sustainability in comparison with the optimal scheme. Moreover, a breakdown of the environmental pillar was conducted, in order to provide insights into the

environmental impacts related to the biorefinery design. It was clearly demonstrated that the reduction of external utilities through heat integration, resulted in a significant decrease in carbon footprint impact category. Furthermore, sensitivity analysis shows the potential aspects that should be taken into account to increase the process sustainability.

## ACKNOWLEDGEMENTS

Authors greatly acknowledge the National Research Council CONICET, ANPCyT, and Universidad Nacional del Sur for supporting their work through grants PICT 2019-03444, PICT 2019-03053, PGI 24/M173 and PGI 24/M161, respectively.

## REFERENCES

1. Celik PA, Barut D, Blaise ME, Kubra EG, Belma NY, Mehmet BM, Ahmet C, A novel higher polyhydroxybutyrate producer *Halomonas halmophila* 18H with unique cell factory attributes. *Bioresour. Technol*: 372, 128669 (2023)
2. Zhou W, Bergsma S, Colpa DI, Euverink GJ, Krooneman J, Polyhydroxyalkanoates (PHAs) synthesis and degradation by microbes and applications towards a circular economy. *J. Environ. Manage*: 341, 118033 (2023)
3. Koller M, Mukherjee A, A New Wave of Industrialization of PHA Biopolyesters. *Bioeng*: 9, 72 (2022)
4. González-Rojo S, Díez-Antolínez, R, Production of polyhydroxyalkanoates as a feasible alternative for an integrated multiproduct lignocellulosic biorefinery. *Bioresour. Technol*: 386, 129493 (2023)
5. Ramos, F.D., Villar, M.A., Diaz, M.S., 2017. Optimal design of poly(3-hydroxybutyrate) production using alternative carbon sources. *Comput. Aided Chem. En.* 40, 877–882.
6. Ramos, F. D.; Delpino, C. A.; Villar, M. A.; Diaz, M. S. Design and Optimization of Poly(Hydroxyalkanoate)s Production Plants Using Alternative Substrates. *Bioresour. Technol.* 2019, 289 (May).
7. Zore, Ž.; Čuček, L.; Širovnik, D.; Novak Pintarič, Z.; Kravanja, Z. Maximizing the Sustainability Net Present Value of Renewable Energy Supply Networks. *Chem. Eng. Res. Des.* 2018, 131, 245–265
8. López, J.A., Naranjo, J.M., Higuera, J.C., Cubitto, M.A., Cardona, C.A., Villar, M.A., 2012. Biosynthesis of PHB from a new isolated *Bacillus megaterium* strain: outlook on future developments with endospore forming bacteria. *Biotechnol.*

- Bioprocess Eng. 17, 250–258.
9. Poomipuk, N., Reungsang, A., Plangklang, P., 2014. Poly- $\beta$ -hydroxyalkanoates production from cassava starch hydrolysate by *Cupriavidus* sp. KKU38. *Int. J. Biol. Macromol.* 65, 51–64.
  10. Nonato, R., Mantelatto, P., Rosell, C., 2001. Integrated production of biodegradable plastic, sugar and ethanol. *Appl. Microbiol. Biotechnol.* 57 (1–2), 1–5.
  11. Rincón, L.E., Becerra, L.A., Moncada, J., Cardona, C.A., 2014. Techno-economic analysis of the use of fired cogeneration systems based on sugar cane bagasse in south eastern and mid western regions of Mexico. *Waste Biomass Valor.* 5, 189–198.
  12. Moncada, J., El-Halwagi, M.M., Cardona, C.A., 2013. Techno-economic analysis for a sugarcane biorefinery: Colombian case. *Bioresour. Technol.* 135, 533–543.
  13. Kapritchkoff, F.M., Viotti, A.P., Alli, R.C.P., Zuccolo, M., Pradella, J.G.C., Maiorano, A.E., Miranda, E.A., Bonomi, A., 2006. Enzymatic recovery and purification of polyhydroxybutyrate produced by *Ralstonia eutropha*. *J. Biotechnol.* 122, 453–462.
  14. Posada, J.A., Naranjo, J.M., López, J.A., Higueta, J.C., Cardona, C.A., 2011. Design and analysis of PHB production processes from crude glycerol. *Process Biochem.* 46, 310–317.
  15. Chen, J., Gu, G., 2001. Cleaner recovery of poly(3-hydroxybutyric acid) synthesized in *Alcaligenes eutrophus*. *Process Biochem.* 36, 501–506
  16. McCarl, B.; Meeraus, P.; Van Der Eijk, M.; Bussieck, S.; Dirkse, F. McCarl GAMS User Guide Table of Contents. *GAMS Dev. Corp.* 2012.
  17. Pedrozo, H. A.; Casoni, A. I.; Ramos, F. D.; Estrada, V.; Diaz, M. S. Simultaneous Design of Macroalgae-Based Integrated Biorefineries and Their Heat Exchanger Network. *Comput. Chem. Eng.* 2022, 164, 107885.
  18. Ulrich, G. d.; Palligarnai T., V. *Chemical Engineering Process Design and Economics: A Practical Guide*; Process Publishing, 2004.
  19. Bare, J. C.; Hofstetter, P.; Pennington, D. W.; Udo de Haes, H. A. Life Cycle Impact Assessment Workshop Summary. Midpoints versus Endpoints: The Sacrifices and Benefits. *Int. J. Life Cycle Assess.* 2000, 5 (6), 319–326.
  20. Grossmann, I. E., Viswanathan, J., Vecchiotti, A., Raman, R., Kalvelagen, E., 2003. GAMS/DICOPT: A Discrete Continuous Optimization Package. Washington, DC, USA
  21. Levett, I., Birkett, G., Davies, N., Bell, A., Langford, A., Laycock, B., Lant, P., Pratt, S., 2016. Techno-economic assessment of poly-3-hydroxybutyrate (PHB) production from methane-The case for thermophilic bioprocessing. *J. Environ. Chem. Eng.* 4, 3724–3733.
  22. Delft Universtiy of Technology. The Model of the Eco-Costs/Value Ratio (EVR). <https://www.ecocostsvalue.com/>
  23. Kang, L.; Liu, Y.; Liang, X. Multi-Objective Optimization of Heat Exchanger Networks Based on Analysis of Minimum Temperature Difference and Accumulated CO<sub>2</sub> Emissions. *Appl. Therm. Eng.* 2015, 87, 736–748.
  24. Acharjee, S. A.; Bharali, P.; Gogoi, B.; Sorhie, V.; Walling, B.; Alemtoshi, 2022. PHA-Based Bioplastic: a Potential Alternative to Address Microplastic Pollution. *Water, Air, & Soil Pollution* 21, 234.
  25. Aui, A.; Li, W.; Wright, M. M. Techno-Economic and Life Cycle Analysis of a Farm-Scale Anaerobic Digestion Plant in Iowa. *Waste Manag.* 2019, 89, 154–164.
  26. Rizwan, M.; Saif, Y.; Almansoori, A.; Elkamel, A. Optimal Processing Route for the Utilization and Conversion of Municipal Solid Waste into Energy and Valuable Products. *J. Clean. Prod.* 2018, 174, 857–867.
  27. Ramos, M.; Lasry Testa, R.; Ramos, F. D.; Estrada, V.; Diaz, M. S., 2023. Simultaneous design of integrated cyanobacteria-based biorefinery and its heat exchanger network. *Comput. Aided Chem. Eng.* 52, 2075–2080.
  28. Nemhauser, G.; Wolsey, L. *Integer and Combinatorial Optimization*; Graham, R. L., Lenstra, J. K., Tarjan, R. E., Eds.; John Wiley & Sons, Inc, 1988.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Computer-Aided Mixture Design Using Molecule Superstructures

Philipp Rehner, Johannes Schilling, André Bardow\*

ETH Zurich, Energy and Process Systems Engineering, Zurich, Switzerland

\* Corresponding Author: [abardow@ethz.ch](mailto:abardow@ethz.ch).

## ABSTRACT

Computer-aided molecular and process design (CAMPD) tries to find the best molecules together with their optimal process. If the optimization problem considers two or more components as degrees of freedom, the resulting mixture design is challenging for optimization. The quality of the solution strongly depends on the accuracy of the thermodynamic model used to predict the thermophysical properties required to determine the objective function and process constraints. Today, most molecular design methods employ thermodynamic models based on group counts, resulting in a loss of structural information of the molecule during the optimization. Here, we unlock CAMPD based on property prediction methods beyond first-order group-contribution methods by using molecule superstructures, a graph-based molecular representation of chemical families that preserves the full adjacency graph. Disjunctive programming is applied to optimize molecules from different chemical families simultaneously. The description of mixtures is enhanced with a recent parametrization of binary group/group interaction parameters. The design method is applied to determine the optimal working fluid mixture for an Organic Rankine cycle.

**Keywords:** Molecular Design, Energy Conversion, Process Design, Optimization, Exergy Efficiency

## INTRODUCTION

Computer-aided molecular and process design (CAMPD) [1] determines optimal molecules jointly with their optimal process. Evaluating molecules for their performance in a process enables a comprehensive assessment. The influence of different molecular characteristics is combined in a process-level target function that can quantify the thermodynamic, economic, or environmental performance of the process.

Molecules occur as degrees of freedom for various energy conversion processes like heat pumps [2] and organic Rankine cycles (ORC) [3-5] or separation processes like CO<sub>2</sub> capture [6]. The design space of molecules is vast but can even be extended exponentially by considering blends of multiple molecular species [7,8].

To optimize a molecule, its structure must be featurized, i.e., expressed in a machine-readable way [9]. Common features used in molecular design are group counts that indicate the number of occurrences of predefined groups in a molecule. The group counts can be used as degrees of freedom in an optimization algorithm.

Additional molecular constraints can ensure that the optimization only generates valid molecules.

The group counts are used in group-contribution methods to determine the thermophysical properties required for the process model [10]. A disadvantage of using group counts as features is that some of the structural information of the molecule is lost in the featurization. Therefore, the optimization does not necessarily allow the identification of an optimal molecule but rather a set of optimal isomers that share the same group counts.

The limitation of group counts is alleviated by maintaining structural information during the optimization. This can be achieved by optimizing the coefficients of an adjacency matrix [11]. However, the number of degrees of freedom becomes large, which is a limiting factor in the presence of expensive model evaluations that occur in an integrated molecular and process design. As an alternative, we recently introduced molecule superstructures [12]. The superstructures can encode all molecules of a particular chemical family using a significantly reduced number of binary variables.

The molecular design space superstructures offer is

smaller than first-order group-contribution methods. However, the molecular representation should not only be picked for computational convenience but also based on the expected accuracy of property prediction methods. Often, extrapolating a group contribution method to molecules significantly different from the molecules used for parameter adjustment (e.g., polyfunctional molecules) leads to low accuracies that undermine the results of a molecular design.

Evaluating a process model requires an equation of state or an equivalent model. For mixture design, the model should predict a wide range of fluids and their mixtures. Molecular equations of state like those based on statistical associating fluid theory (SAFT) [13] leverage molecular insights to describe mixtures accurately based on pure-component parameters. If a higher accuracy is required, binary interaction parameters can further enhance the description of mixtures.

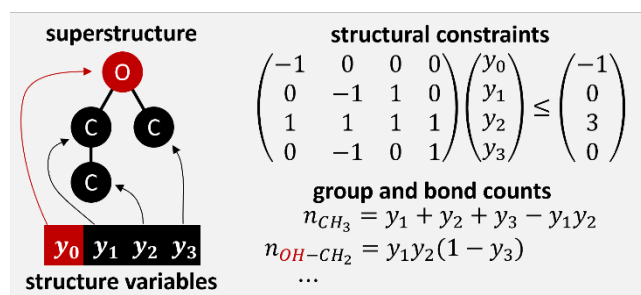
In a recent study, the authors and co-workers [14] extended the group contribution method for PCP-SAFT [15,16] by Sauer et al. [17] to mixtures by accounting for hydrogen bonding between polar molecules (aldehydes, ketones, ethers, formates, and esters) and self-associating molecules (alcohols and amines), and by parametrizing a group contribution method for binary interaction parameters. The study compares a homosegmented and a heterosegmented group-contribution method. The homosegmented approach determines PCP-SAFT parameters from group counts and group-specific parameters. In the heterosegmented approach, segments replace molecules as species in the equation of state. The heterosegmented approach was determined to be more accurate for pure components [17] and for mixtures [14], but to parametrize the model, the number of bonds between different groups is required in addition to the group counts. The molecule superstructures [12] can be used to infer both group counts and bond information.

In this work, the more accurate heterosegmented group-contribution method for PCP-SAFT is unlocked for application in an integrated molecular and process design using molecule superstructures. We demonstrate the approach in a case study determining the optimal working fluid mixture for an organic Rankine cycle (ORC). Due to the combinatorial complexity of possible working fluid mixtures, a molecular design is necessary to determine the optimal working fluid systematically.

## MOLECULE SUPERSTRUCTURES

To optimize molecules in a computer-aided molecular and process design, the structure of the molecule needs to be represented in a format that is accessible to the optimization algorithm. Molecule superstructures are a graph-based representation of a molecular family [12]. Fig. 1 shows an example of a small molecule structure. It

consists of a graph with four nodes and three edges. Every node corresponds to an atom type and a binary structure variable. Structural constraints are introduced to ensure that only valid molecules are generated (in this case water, methanol, ethanol, or dimethyl ether). The structure variables are converted to group and bond counts that are required to parametrize the equation of state. The structural constraints and group/bond counts can be derived generically for arbitrarily large superstructures. We refer to the original publication [12] for the detailed expressions.



**Figure 1:** Example for an ether/alcohol superstructure including the corresponding structural constraints and resulting group and bond counts.

Previously, optimal molecules were determined for every chemical family individually [12]. For mixtures, the number of possible combinations of chemical families becomes large. Therefore, we extend the superstructure approach to optimize any number of chemical families simultaneously using methods from disjunctive programming [18].

The superstructure converts binary structure variables  $y_k$  into molecular features used in the property prediction method. Linear inequality constraints are required to ensure that only valid molecules are found. The specific constraints are elucidated in the original publication [12]. In general, the constraints can be expressed as

$$\sum_k a_{ijk} y_k \leq b_{ij} \quad (1)$$

Here,  $i$  enumerates the chemical families and their respective molecule superstructures,  $j$  indicates different constraints, and  $k$  enumerates the structure variables  $y_k$ . The number of constraints  $N_{c,i}$  and structure variables  $N_{y,i}$  depends on the considered molecule superstructures.

To combine the different chemical families in one structure, a common set of structure variables is defined by using  $N_y = \max_i N_{y,i}$  binary variables and adding additional constraints that force the unused variables to 0.

$$\sum_{k=N_{y,i}+1}^{N_y} y_k \leq 0 \quad (2)$$

The constraints associated with each molecule superstructure can then be formally expressed as disjunctions in the optimization problem.

$$v_i [\sum_k a_{ijk} y_k \leq b_{ij}] \quad (3)$$

The Big-M method and convex hulls have been established as standard methods to convert disjunctive programs into MILPs or MINLPs [18]. The convex hull has a tighter feasible region than the Big-M formulation but requires more additional variables. Therefore, the Big-M formulation is better suited for the integrated process design task with expensive target function evaluations. Applying the Big-M method introduces additional binary variables  $c_i$ , one for each disjunction and hence chemical family. Only one disjunction can be active at the same time, leading to the additional constraint:

$$\sum_i c_i = 1 \quad (4)$$

The constraints for every molecule superstructure are rewritten by introducing the parameter  $M_{ij}$ .

$$\sum_k a_{ijk} y_k + M_{ij} c_i \leq b_{ij} + M_{ij} \quad (5)$$

In general,  $M_{ij}$  must be large enough to ensure that the constraint is turned off if  $c_i = 0$ , but as small as possible to ensure tight relaxation bounds. For the binary variables  $y_k$ , the optimal value for  $M_{ij}$  can be derived generically, as

$$M_{ij} = \sum_k \max(a_{ijk}, 0) - b_{ij} \quad (6)$$

The original publication on molecule superstructures [12] describes a generic method to infer the group and bond counts from the structure variables. The same method is used for every individual superstructure. The resulting group counts  $n_{i,\alpha}$  and bond counts  $b_{i,\alpha\beta}$  are then weighted with the chemical family indicator  $c_i$  of every superstructure to give the total group and bond counts.

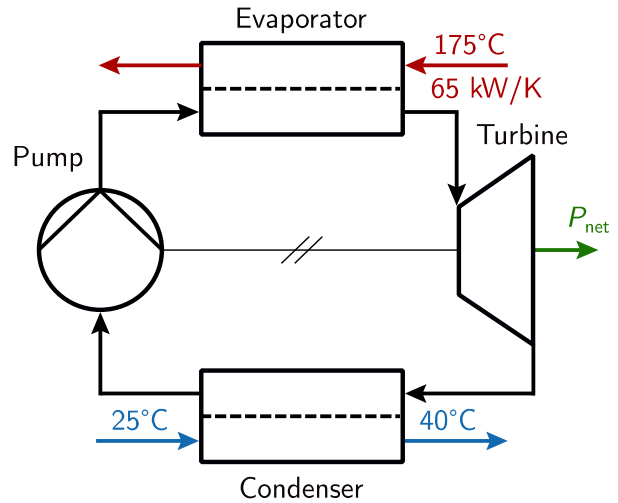
$$n_\alpha = \sum_i c_i n_{i,\alpha}, \quad b_{\alpha\beta} = \sum_i c_i b_{i,\alpha\beta} \quad (8)$$

The group and bond counts are used to parametrize the heterosegmented group-contribution method for PCP-SAFT [17], which is used to determine phase equilibria and residual properties. To calculate caloric properties, the group counts are also used in the group-contribution method for the ideal gas heat capacity by Joback and Reid [19]. The resulting MINLP is solved using the branch-and-bound implementation in Artelys Knitro [20].

## CASE STUDY: ORGANIC RANKINE CYCLE

We apply the molecular design method using molecule superstructures to the integrated design of an organic Rankine cycle (ORC) and its working fluid. ORCs can convert waste heat into power, reducing energy losses and increasing the exergetic efficiency of processes. The process flowsheet of the standard ORC used in this case study is shown in Fig. 2. Analogous to a regular Rankine cycle, the working fluid of the ORC is evaporated and potentially superheated to be then fed into a

turbine that extracts power from the working fluid. Part of the power is used to pump the working fluid that leaves the condenser back to the pressure level of the evaporator.



**Figure 2:** Flowsheet and key specifications for the ORC.

Using an organic working fluid rather than water enables the operation above atmospheric pressures, even for low-temperature heat sources. This work optimizes the working fluid mixture to maximize the power output of the ORC. The design of an actual waste heat valorization unit needs to consider additional aspects regarding safety, stability, and costs that are outside the scope of this study.

ORC processes are promising targets for mixture design since replacing pure working fluids with zeotropic mixtures leads to non-isothermal evaporation and condensation. The resulting temperature glide in the condenser and evaporator can be tuned to the respective temperature glides in the heat source and cooling mediums. A more homogeneous temperature difference across the heat exchangers reduces the exergy loss during heat transfer and, in conclusion, increases the net power output of the process.

The case study is based on the analysis of Chys et al. [21] and the design study by Schilling et al. [8]. The most important specifications are shown in Fig. 2: The heat source has an inlet temperature of 175 °C and a heat capacity rate of 65 kW/K. The cooling medium is heated from 25 °C to 40 °C. The outlet temperature of the heat source medium is only indirectly constrained via the pinch condition. The heat exchangers are modeled as isobaric. The full specification of the process is given in Tab. 1. This thermodynamic optimization aims to maximize the net power  $P_{net}$  of the cycle.



**Table 1:** ORC case study specifications from [8,21]

Component	Parameter	Value
Heat source	Inlet temperature	175 °C
	Heat capacity rate	65 kW/K
	Min. approach temp.	20 K
Turbine	Isentropic efficiency	0.65
	Min. vapor quality (outlet)	1
Pump	Isentropic efficiency	0.8
Condenser	Min. absolute pressure	1 bar
	Min. relative pressure	1e-3
Evaporator	Max. absolute pressure	50 bar
	Max. relative pressure	0.8
Cooling	Inlet temperature	25 °C
	Outlet temperature	40 °C
	Min. approach temp.	10 K

## RESULTS

As a reference, the molecule superstructures with disjunctions are first used to determine the optimal pure working fluid for the ORC. A ranking of the ten molecules that yield the highest net power output is shown in Tab. 2. The ranking consists of alkanes, alkenes, and propyne.

**Table 2:** Ranking of the best pure component working fluids based on the net power output of the ORC.

#	SMILES	Name	Net power [kW]
1	CCC	propane	472
2	C=CC	propene	435
3	C#CC	propyne	425
4	CC(C)C	isobutane	412
5	CC(C)(C)C	neopentane	409
6	CCCC	butane	399
7	C=C(C)C	isobutene	396
8	C=CCC	butene	395
9	C=CC(C)(C)C	neohexene	395
10	CCC(C)C	isopentane	394

The molecular design space also contains polar (aldehydes, ketones, ethers, formates, and esters) and self-associating molecules (alcohols and amines). However, these molecular families are unsuitable for the application due to their lower vapor pressures and higher critical temperature. The ranking is topped by the C3 hydrocarbons propane, propene, and propyne, with the highest achievable net power output being 472.25 kW using propane as the working fluid.

Higher net power outputs can be achieved by considering mixed working fluids. To obtain a ranking of working fluid candidates, we use the molecule

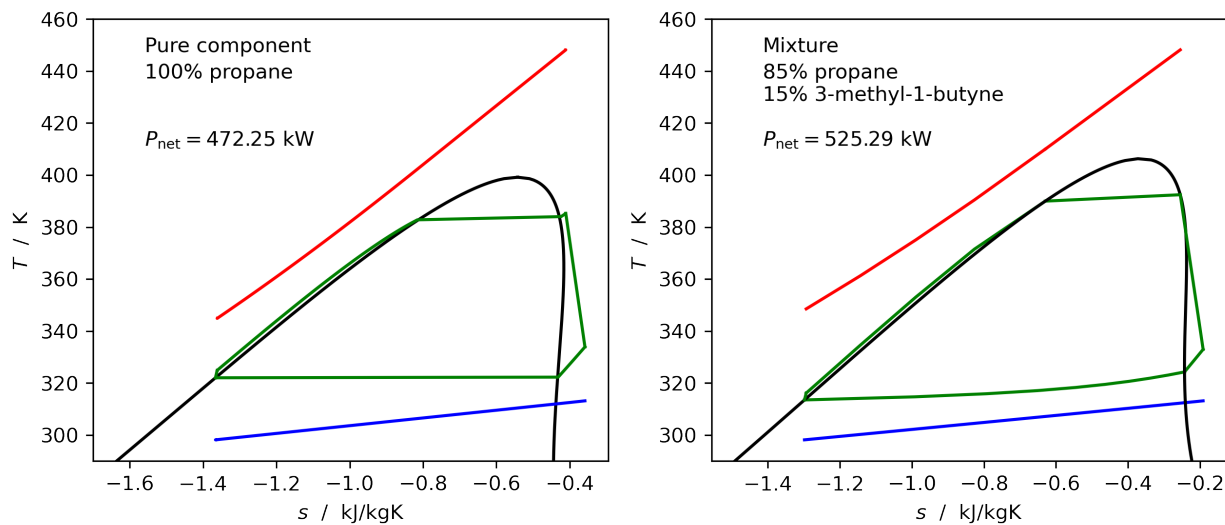
superstructure concept to find optimal additives for each of the ten best-performing pure components. The ranking is shown in Tab. 3. The ten best mixtures all consist of mainly propane or propene with traces of larger, mostly non-polar molecules. The mixtures containing one of the other well-performing fluids from the pure-component case consistently lead to lower power outputs than those based on propane or propene. This observation strongly implies that optimizing additives can determine the optimal mixture for this case study. However, both molecules must be optimized in a single optimization problem for a more comprehensive result. Schilling et al. [8] demonstrate how the approach can significantly reduce the number of function evaluations compared to a screening of mixtures using a CAMD method based on group counts. The design setup in our study allows the simultaneous optimization of both constituent molecules, but the prevalence of local optima impedes an efficient direct determination of optimal working fluid mixtures. Therefore, it was not possible yet to determine a ranking comparable to Tab. 3 by optimizing both fluids rather than an additive in a reasonable timeframe. A MINLP solution algorithm better suited for the specificity of integrated molecular and process design is required to improve the solution times and generate more robust candidate mixtures [22]. The non-convexity of the MINLP also suggests using global optimization [23], however, the expensive function evaluations that include multiple phase equilibrium calculations with the heterosegmented group-contribution method for PCP-SAFT are a limiting factor.

Fig. 3 compares the process for the optimal pure working fluid (propane) with the best mixed working fluid candidate. In the pure-component case, the temperature of the working fluid is constant during evaporation and condensation. Therefore, the optimization mainly determines process conditions in which the temperature glide of the heat source medium aligns well with the temperature during the preheating step in the evaporator. The mixed working fluid shows a temperature glide in the condenser. The resulting process conditions find a compromise between the exergy loss in the evaporator and the condenser. Consequently, the net power output of the ORC is increased by 11% compared to propane as the working fluid. This increase is partly driven by the temperature glide in the cooling medium, which is fixed to 15 K in this case study. For other process conditions, the benefit of using mixed working fluids can be small [24], and a reduced heat transfer during condensation and evaporation must be considered for an economic assessment of the process [25,8].

Choosing the correct composition of the mixture is essential for the performance of the ORC. Fig. 4 shows the optimized net power output of the ORC using the four best mixtures from Tab. 3 over the entire composition

**Table 3:** Ranking of the best working fluid mixtures found in the integrated design.

#	Component 1		Component 2		Molar composition	Net power [kW]
	SMILES	Name	SMILES	Name		
1	CCC	propane	C#CC(C)C	3-methyl-1-butyne	85% / 15%	525
2	CCC	propane	CC(C)(C)C	neopentane	92% / 8%	524
3	CCC	propane	C#CCC	1-butyne	65% / 35%	522
4	CCC	propane	CCC(C)C	isopentane	94% / 6%	519
5	C=CC	propene	CCCC	butane	83% / 17%	519
6	C=CC	propene	CCC(C)C	Isopentane	93% / 7%	518
7	CCC	propane	C=C(C)CC	2-methyl-1-butene	95% / 5%	518
8	CCC	propane	CCCCC	pentane	97% / 3%	515
9	CCC	propane	CCCO	1-methoxypropane	96% / 4%	514
10	CCC	propane	COC(C)C	2-methoxypropane	95% / 5%	512

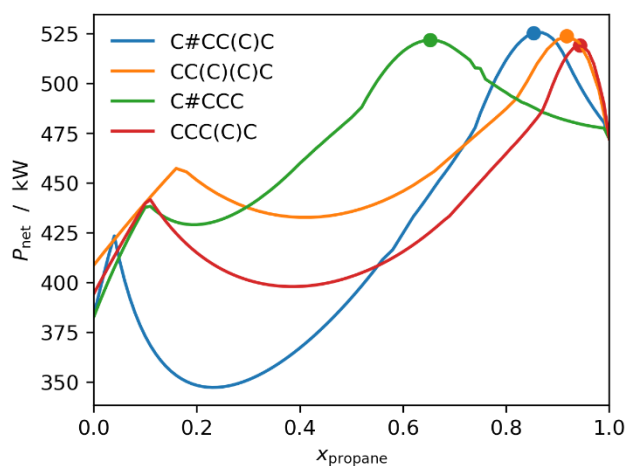
**Figure 3:** Ts-diagram for the ORC using propane (left) and the best mixture (right) as working fluid. The phase change of the working fluid (green) is isothermal for pure components but exhibits a temperature glide for mixtures. In the optimization the temperature glide is adjusted to the temperature profiles in the heat source medium (red) and the cooling water (blue).

range. The solutions from the mixture design align well with the maxima in the discretized curves. In all cases, adding the longer hydrocarbon to pure propane initially increases the net power due to the establishment of the temperature glide. Further increasing the amount of additive reduces the performance, with the curves showing local minima that, in one case, are even below the net power output of the pure additive.

The net power output is strongly related to the ratio of the temperature glides in the condenser and the cooling medium [8]. Changing the composition of the mixtures affects the glide in the condenser, with the largest glide occurring in the middle of the composition range. The glide also explains the shape of the maxima in the net power: A jump of the pinch from the condenser outlet to the inlet leads to a kink in the net power (cf. Fig. 4). If the temperature profile in the condenser is shaped in a

way that the pinch is located between the inlet and the outlet, the maximum in the net power is smooth.

A strong sensitivity to variations in the composition of the working fluid can be problematic in an actual ORC process when leakage is non-negligible, and the composition can change due to the constituents' different relative volatilities. With the abundance of near-optimal mixtures in the solution space of the molecular design, the robustness with respect to composition changes can be incorporated as an additional constraint in the optimization problem.



**Figure 4:** Optimized net power output of the ORC using a propane + X mixture with the specified composition. The marks represent the optimal composition found in the molecular design.

## CONCLUSION

Molecule superstructures are used to determine the optimal working fluid mixture for an ORC. Multiple chemical families represented by their own superstructure are optimized simultaneously using disjunctive programming. Due to the full structural information available during the molecular design, the accurate heterosegmented group-contribution method for PCP-SAFT can be used as a property prediction method. The model accuracy is further enhanced by a recent parametrization of binary group/group interaction parameters [14].

The optimization determines mixtures of propane or propene with larger hydrocarbons as optimal working fluids. The top-ranked mixtures deliver comparable performances. This observation suggests that additional material properties should be considered within the optimization to reduce the molecular design space. Amongst the mixtures with the best thermodynamic properties are two that contain alkynes, components that are unsuited for long-term application as working fluids due to their low chemical stability. If predictive models are available, adding constraints for the working fluid's safety, stability, and environmental impact avoids a posteriori filtering of the results and provides a more efficient design method. Finally, economic considerations were circumvented by empirical pinch constraints instead of a direct consideration by using more detailed, rate-based process models [8]. Also, in this case, the more detailed representation by molecule superstructures could enable advanced thermodynamic models, e.g., for the required transport properties.

## ACKNOWLEDGEMENTS

PR acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 497566159. AB acknowledges support by BRIDGE as part of the project “High-Efficiency High-Temperature Heat Pumps with Temperature Glide.” We thank the Swiss National Science Foundation SNSF and Innosuisse for their support.

## REFERENCES

1. Papadopoulos AI, Tsivintzelis I, Linke P, Seferlis P. Computer-aided Molecular Design: Fundamentals, Methods, and Applications. *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering* (2018)
2. Roskosch D, Atakan B. Reverse engineering of fluid selection for thermodynamic cycles with cubic equations of state, using a compression heat pump as example. *Energy* 81:202-212 (2015)
3. Papadopoulos A, Stijepovic M, Linke P. On the systematic design and selection of optimal working fluids for Organic Rankine Cycles. *Appl Therm Eng* 30:760-769 (2010)
4. Schilling J, Lampe M, Gross J, Bardow A. 1-stage CoMT-CAMD: An approach for integrated design of ORC process and working fluid using PC-SAFT. *Chem Eng Sci* 159:217-230 (2017)
5. Bowskill D, Tropp UE, Gopinath S, Jackson G, Galindo A, Adjiman CS. Beyond a heuristic analysis: integration of process and working-fluid design for organic Rankine cycles. *Mol Syst Des Eng* 5:493-510 (2020)
6. Papadopoulos A, Badr S, Chremos A, Forte E, Zarogiannis T, Seferlis P, Papadokostantakis S, Galindo A, Jackson G, Adjiman C. Computer-aided molecular design and selection of CO<sub>2</sub> capture solvents based on thermodynamics, reactivity and sustainability. *Mol Syst Des Eng* 1:313-334 (2016)
7. Papadopoulos A, Stijepovic M, Linke P, Seferlis P, Voutetakis S. Toward optimum working fluid mixtures for organic Rankine cycles using molecular design and sensitivity analysis. *Ind Eng Chem Res* 52:12116-12133 (2013)
8. Schilling J, Entrup M, Hopp M, Gross J, Bardow A. Towards optimal mixtures of working fluids: Integrated design of processes and mixtures for Organic Rankine Cycles. *Renew Sust Energ Rev* 135:110179 (2021)
9. Austin ND, Sahinidis NV, Trahan DW. Computer-aided molecular design: An introduction and review of tools, applications, and solution techniques. *Chem Eng Res Des* 116:2-26 (2016)

10. Marrero J, Gani R. Group-contribution based estimation of pure component properties. *Fluid Phase Equilib* 183:183-208 (2001)
11. Churi N, Achenie LEK. Novel mathematical programming model for computer aided molecular design. *Ind Eng Chem Res* 35(10):3788-3794 (1996)
12. Rehner P, Schilling J, Bardow A. Molecule superstructures for computer-aided molecular and process design. *Mol Syst Des Eng* 8:488-499 (2023)
13. Chapman WG, Gubbins KE, Jackson G, Radosz M. New reference equation of state for associating liquids. *Ind Eng Chem Res* 29(8):1709-1721 (1990)
14. Rehner P, Bardow A, Gross J. Modeling mixtures with PC-SAFT: Insights from large-scale parametrization and group-contribution method for binary interaction parameters. *Int J Thermophys* 44(12), 179 (2023)
15. Gross J, Sadowski G. Perturbed-chain SAFT: An equation of state based on a perturbation theory for chain molecules. *Ind Eng Chem Res* 40(4):1244-1260 (2001)
16. Gross J, Vrabec J. An equation-of-state contribution for polar components: dipolar molecules. *AIChE J* 52(3):1194-1204 (2006)
17. Sauer E, Stavrou M, Gross J. Comparison between a homo- and a heterosegmented group contribution approach based on the perturbed-chain polar statistical associating fluid theory equation of state. *Ind Eng Chem Res* 53(38):14854-14864 (2014)
18. Grossmann I. Review of nonlinear mixed-integer and disjunctive programming techniques. *Optim Eng* 3:227-252 (2002)
19. Joback KG, Reid RC. Estimation of pure-component properties from group-contributions. *Chem Eng Commun* 57:233-243 (1987)
20. Byrd RH, Nocedal J, Waltz RA. KNITRO: An integrated package for nonlinear optimization. In: Large-Scale Nonlinear Optimization. Springer (2006)
21. Chys M, van den Broek M, Vanslambrouck B, De Paepe M. Potential of zeotropic mixtures as working fluids in organic Rankine cycles. *Energy* 44(1):623-632
22. Jonuzaj S, Gupta A, Adjiman CS. The design of optimal mixtures from atom groups using Generalized Disjunctive Programming. *Comput Chem Eng* 116:401-421 (2018)
23. Boukouvala F, Misener R, Floudas CA. Global optimization advances in Mixed-Integer Nonlinear Programming, MINLP, and Constrained Derivative-Free Optimization, CDFO. *Eur J Oper Res* 252(3): 701-727 (2016)
24. Krempeus D, Bahamonde S, van der Stelt TP, Klink W, Colonna P, De Servi CM. On mixtures as working fluids of air-cooled ORC bottoming power plants of gas turbines. *Appl Therm Eng* 236(D):121730 (2024)
25. Shah MM. A method for predicting heat transfer during boiling of mixtures in plain tubes. *Appl Therm Eng* 89:812-821 (2015)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Internally Heated Crackers for Decarbonization and Optimization of Ethylene Production

Edwin A. Rodriguez-Gil and Rakesh Agrawal\*

Purdue University, Davidson School of Chemical Engineering, West Lafayette, IN, USA

\* Corresponding Author: [agrawalr@purdue.edu](mailto:agrawalr@purdue.edu)

## ABSTRACT

Ethylene is a crucial precursor for a diverse spectrum of products and services. As global production exceeds 150 million tons annually and is projected to surpass 255 million tons by 2035, the imperative for sustainable and efficient ethylene production becomes increasingly clear. Despite Externally Heated Crackers (EHCs) dominating ethylene production for over a century, they face intrinsic limitations that necessitate transformative solutions, including intense radial thermal gradients, high metal demand, and substantial CO<sub>2</sub> emissions. This study employs a robust combination of Computational Fluid Dynamics (CFD) coupled with detailed chemical kinetics to rigorously assess selected configurations of Internally Heated Crackers (IHCs) against the leading EHC designs. Our findings reveal that IHCs exhibit the potential to enhance ethylene output by a factor of 1.66 when compared to EHCs of the same length, diameter, and surface temperature. These results herald a promising era for developing more efficient cracking reactor designs, poised to redefine the landscape of sustainable chemical manufacturing towards achieving Net-Zero emissions. Embracing innovative technologies like IHCs presents an opportunity for the chemical industry to make significant strides in reducing its environmental footprint while meeting the growing global demand for ethylene and its derivatives.

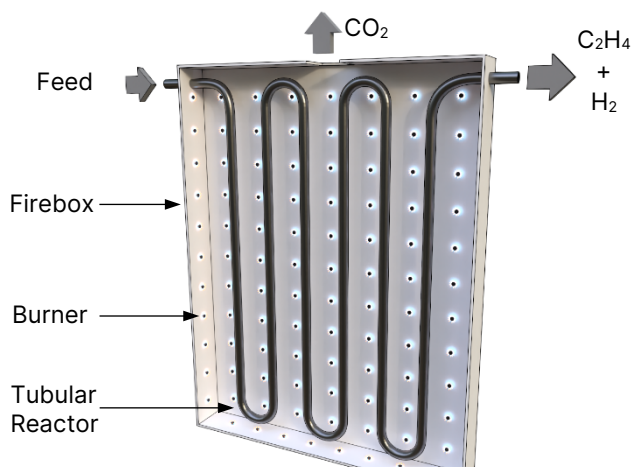
**Keywords:** Net-Zero, Decarbonization, Ethylene, Cracking, Reactor Design, Process Optimization, CFD

## INTRODUCTION

Ethylene (C<sub>2</sub>H<sub>4</sub>) may be one of the simplest hydrocarbons, yet it is a pillar of the global economy. Its value chain includes vital chemical products like polyethylene, ethylene oxide, and polyester. These, in turn, are used across multiple sectors, such as internet communications, medical sterilization, and textiles [1], [2]. The level of importance of ethylene is such that in 2017, the American Association of Fuel and Petrochemical Manufacturers (AAPM) referred to it as the "World's Most Important Chemical" [3].

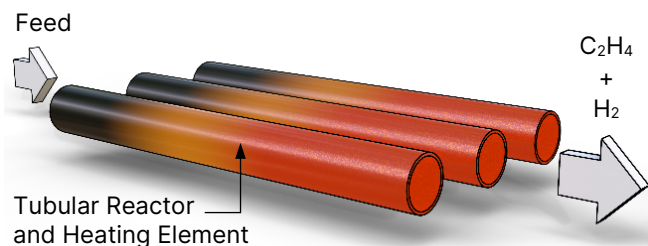
### 1.1 Externally Heated Crackers (EHCs)

Since their industrial debut in 1921, Externally Heated Crackers (EHCs) have been the dominant technology for ethylene production [4], [5]. These specialized units operate at temperatures of ~850°C to split or "crack" hydrocarbon feedstocks into ethylene [6]. Figures 1, 2, and 3 illustrate the main EHC configurations.

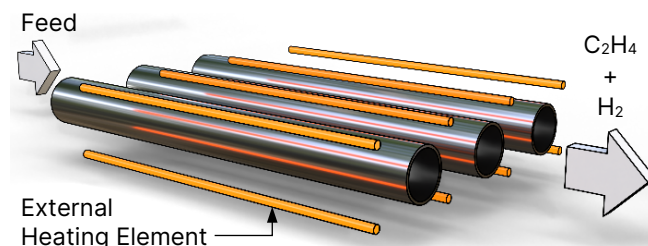


**Figure 1.** Illustration of a Flame-heated Tubular Cracker (FTC). External burners heat a series of long reaction tubes, facilitating cracking. This configuration is commonly called a steam cracker when steam serves as the diluent.





**Figure 2.** Illustration of an Impedance Tubular Cracker (ITC). Electricity flows directly through the reaction tubes generating heat. This configuration is also known as direct heating cracking [7].



**Figure 3.** Illustration of an External Resistance Tubular Cracker (ERTC). Electric heating elements outside the reaction tube provide heat to facilitate cracking. This configuration is also known as indirect heating cracking [8].

## 1.2 Intrinsic Limitations of EHCs

The efficiency of cracking reactors is intrinsically tied to their heat transfer performance. In the case of EHCs, one of the primary limitations stems from internal radial thermal gradients, reaching up to 100°C from the tube walls to the center [9]. These uneven reaction conditions result in reduced ethylene yields near the center and promote undesired reactions close to the wall. To mitigate this effect, tube diameters are kept remarkably small, sometimes as narrow as 3 cm [10]. However, this approach reduces the flow area, necessitating a large number of parallel tubes to scale up production. Depending on the specific design, a single EHC may require up to 200 reaction tubes, representing ~2500 tons of raw metal and costing around half a billion dollars due to the specialized alloys involved [11], [12].

Achieving precise temperature control in EHCs is also challenging due to the large number of reaction tubes and the difficulty in continuously monitoring internal temperatures. In practice, only the mean outlet temperatures are measured, resulting in sub-optimal ethylene outputs and reduced reactor lifespan. At typical reaction conditions, an increase of just 20°C in the metal temperature can halve its lifespan [13]. This situation is highly undesirable due to the high cost of the materials involved. Furthermore, when fossil fuels are used for heating, EHCs can release vast amounts of CO<sub>2</sub>.

## 1.3 EHCs and CO<sub>2</sub> emissions

Over the past century, fossil fuel combustion has been the default alternative for heating in EHCs. This is due to the historically lower cost of fossil fuels compared to other energy sources. However, this approach also entails a substantial climate cost. For every ton of ethylene produced, between 1.0 and 1.6 tons of CO<sub>2</sub> are released [6]. In 2021 alone, with a market of ~150 Million tons (Mt), EHCs emitted ~200 Mt of CO<sub>2</sub>, accounting for ~0.6% of global CO<sub>2</sub> emissions [14], [15]. The ethylene market is projected to increase ~73% from \$132B in 2022 to \$228B by 2032 [16]. In a business-as-usual scenario, emissions could surpass 335 Mt in the next decade, directly conflicting with global decarbonization efforts to reach a Net-Zero emissions economy before 2050 [17], [18].

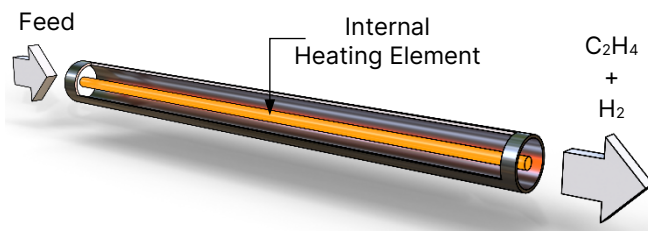
## 1.4 Electric & Hydrogen-powered EHCs

EHCs powered by electricity or hydrogen have recently emerged as alternatives to reduce CO<sub>2</sub> emissions [19], [20]. Notably, in September 2022, BASF, in collaboration with Linde and SABIC, began the construction of a pilot-scale Impedance Tubular Cracker (ITC) and External Resistance Tubular Cracker (ERTC) in Germany [21]. When coupled with low-carbon energy sources, these units can help reduce CO<sub>2</sub> emissions by up to 95% compared to FTCs [22]. However, since they still rely on heat transfer to the gas from the tube walls, they present the same intrinsic limitations that affect all EHC designs.

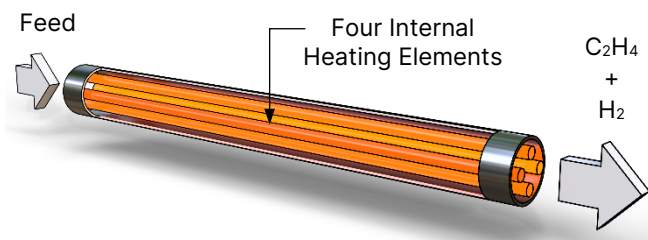
## 1.5 Internally Heated Crackers (IHCs)

The limitations of EHCs have prompted the quest for alternative reactor configurations with superior performance and lower carbon emission, leading to the emergence of Internally Heated Crackers (IHCs) [23], [24]. IHCs set themselves apart by embedding the heating elements directly within the reaction zone, where energy is needed, effectively reducing the radial thermal gradients. This innovative approach facilitates using larger diameter reaction tubes while increasing the heat transfer area, consequently reducing the required amount of metal during construction, and lessening the weight for transportation compared to EHCs [25]. Furthermore, as they are electrically heated, they retain the same benefits of decarbonization. Configurations featuring combined internal and external heating options offer added versatility [26], especially at small diameters where the tube's external surface contributes more to the total heat transfer area.

This study focuses on comparing external heating against solely internal heating designs. However, future simulations will delve into the effects of the combined approach. Figures 4 and 5 illustrate two IHC configurations introduced by Agrawal et al. [23] [26], which will be employed in this study.



**Figure 4.** Illustration of an Internally Heated Cracker (IHC) with one heating rod. An electrical element generates heat directly inside the reactor, in the reaction zone.



**Figure 5.** Illustration of an Internally Heated Cracker (IHC) with four heating rods. The additional heat transfer area enhances heat delivery, producing more ethylene.

## 1.6 IHC vs EHC: A Comparative Study

While IHCs hold promise over traditional EHCs, it is imperative to demonstrate their performance quantitatively through rigorous modeling and experimentation. This study employs a robust combination of Computational Fluid Dynamics (CFD) coupled with detailed chemical kinetics to compare selected lab-scale EHC and IHC configurations. This scale was selected to provide a basis for subsequent experimental work.

The following sections delve into the comparative analysis details and the results obtained, shedding light on the intricacies of ethylene production and the efficiency of these two distinct technologies.

## METHODS

### 2.1 The need for CFD simulations

Evaluating novel cracking reactor designs requires a modeling approach capable of handling fluid dynamics and predicting the heat transfer coefficient for arbitrary geometries. This model must seamlessly integrate with detailed chemical kinetics to rigorously account for free-radical reactions while addressing turbulence effects and heat transfer from radiation. Additionally, it should converge quickly enough and offer reasonable numerical stability to allow for timely simulations. Computational Fluid Dynamics (CFD) simulations provide a useful modeling tool to meet these requirements.

## 2.2 Types of CFD Simulations

Accurately selecting a simulation strategy for a specific project is crucial for ensuring reliable results and overall success. For example, Direct Numerical Simulation (DNS) is the most accurate CFD method, avoiding flow approximations. However, it is also characterized by high computational costs and is primarily reserved for specialized applications and benchmarking. In contrast, approaches like Large Eddy Simulation (LES) or Reynolds-Averaged Navier-Stokes (RANS) deliberately simplify physics across spatial scales, sacrificing some precision, but reducing the computational cost by orders of magnitude compared with DNS.

This research utilizes the RANS equations coupled with an Menter's Shear Stress (SST)  $k$ - $\omega$  model to account for turbulence and a Surface to Surface (S2S) model to address radiation. The SST  $k$ - $\omega$  model was chosen as it is widely accepted as an industry standard for turbulence. The S2S model although computationally expensive, was selected because it is particularly good for modeling the radiative heat transfer in closed domains.

## 2.3 Kinetic Model

The correct implementation of chemical kinetics is a critical element for the modeling of cracking reactors. While the concept of cracking can be traced back to the early 1800s [27], it was not until the 1970s that compelling kinetic models were readily available. This research implements Sundaram and Froment's (1978) kinetic model [28], recognized as some of the most comprehensive in the field, enabling the modeling of feedstocks containing ethane, ethylene, propane, propylene, n-butane, and iso-butane.

Olsvik and Billaud's model (1992) can be added to account for methane thermal decomposition [29]. However, at the reaction conditions for ethylene production, methane behaves mainly as inert, resulting in little differences compared with just Froment's model [30]. Combined, both models comprise a total of 170 chemical reactions and 41 chemical species, including free radicals.

## 2.4 Thermodynamic Properties

Given the temperature changes and chemical reactions involved, obtaining a comprehensive set of thermodynamic properties for each species is essential to model cracking reactors accurately. As the reactions occur in the gas phase, the main properties needed are heat capacity, enthalpy, entropy, and Gibbs free energy. This research employs the NASA 7-coefficient polynomial system with coefficient values sourced from the Active Thermochemical Tables (ATcT), maintained by Argonne National Laboratory [31].

## 2.5 Boundary Condition

Common boundary conditions for inlet/outlet ports in cracking reactors include mass flow rate, flow velocity, and flow pressure. Mass flow rate is often employed for closed domains like tubes, while flow velocity is typically for open environments. Pressure conditions can be used to complement both open and closed domains. This study implemented a combination of pressure boundary conditions at the reactor inlet and mass flow at the reactor outlet.

The inlet pressure was maintained at a constant value of 2 bar across all simulations, aligning with typical industrial ethylene cracker conditions. At the same time, the outlet mass flow was optimized for each configuration to maximize ethylene production. This setup ensures fair evaluation across all cases.

For the reactor walls, common boundary conditions include isothermal, constant heat, and convection cases. Isothermal conditions can be used to assess the maximum ethylene production capacity for a given configuration as it predicts the heat transfer capacity ceiling when the maximum metal temperature is used across all surfaces. Constant heat flow conditions can be used to simulate the effect of electric heating but require prior knowledge of the heat transfer profile. Convection boundary conditions can be used to model heat transfer in open environments. This study implements isothermal walls at 1100 °C for all simulations to compare the maximum ethylene production capacity across designs. The selected temperature mirrors the upper limit achievable with industry-standard nichrome alloy heating elements.

## 2.6 Feed Composition

Industrial ethylene crackers utilize diverse feedstocks, often based on local availability. In the United States, for example, gas feedstocks like ethane or shale gas are prevalent, while in Europe, liquid feedstocks like naphtha are more common [6].

Another crucial aspect is the inclusion of dilution agents, which enhance reaction conversion and reduce coke formation. The level of dilution is an economic optimization variable; higher dilution favors reaction but increases energy consumption [32]. Typical dilution ratios range between 0.3 - 0.7 kg of inert per kg of hydrocarbon [33]. Historically, steam has been the most common dilution agent, but the shale revolution has made methane an attractive alternative due to its abundance in shale gas deposits [34], [35].

Our simulation scheme can incorporate various dilution agents, including steam, methane, nitrogen, and carbon dioxide. However, for this study, no dilution agent was implemented; instead, pure ethane was used as feed in all simulations. This choice enables the isolation of geometry effects from those of dilution. Future simulations will explore the combined effects.

## 2.7 Feed Temperature

Another essential aspect is the inlet feed temperature. In Industrial EHCs, the feed is preheated using the flue gases from combustion or heat integration with the product, reaching ~650 °C inlet feed temperature [6]. In this case, all simulations use 350 °C, aligned with typical lab-scale cracking reactors.

## 2.8 Tube Material

As the material for the reactor tubes, all simulations employ SANDVIK 353 MA. This material is an austenitic chromium-nickel steel alloy with nitrogen and rare earth metals and is representative of the alloys utilized in industrial cracking reactors. Its maximum operating temperature is ~1175 °C and is characterized by a high resistance to oxidation and carburization environments. The main material properties required for simulation are density, heat capacity polynomial, and thermal conductivity polynomial, which can be found in its datasheet.

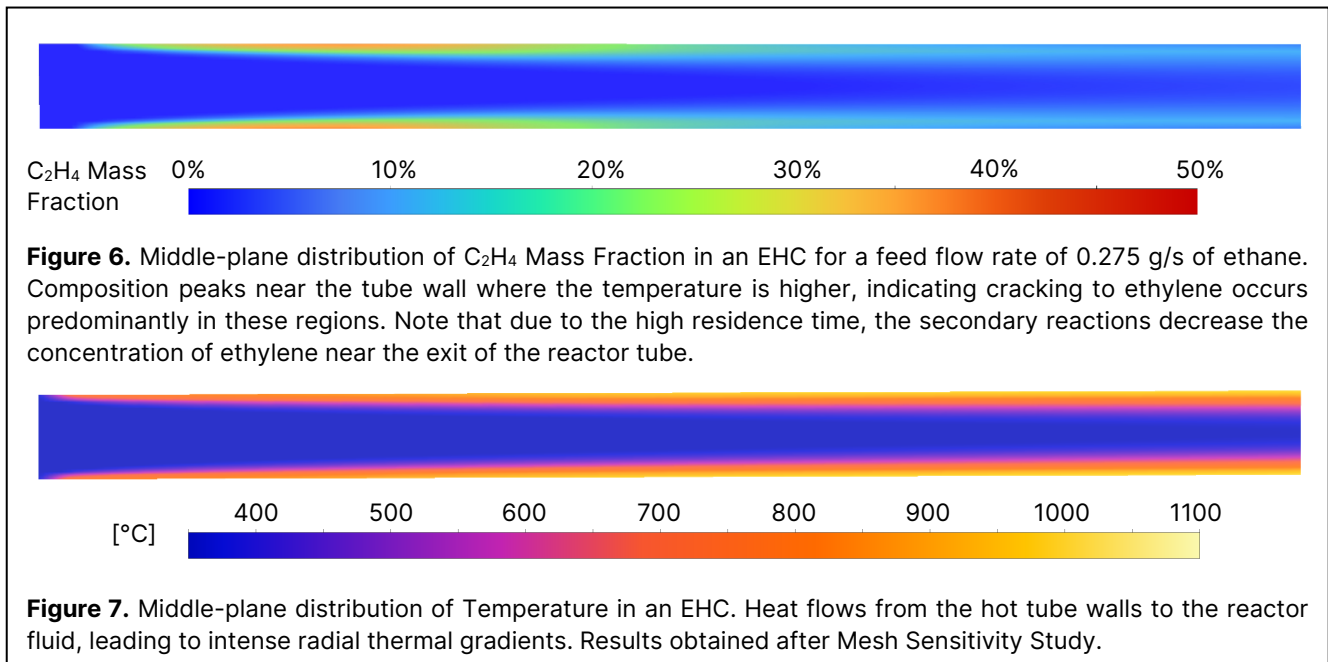
## 2.9 Convergence

In CFD simulations, the convergence criteria play a crucial role in determining the accuracy of the results. While residual functions are commonly used and provide a good indication of the simulation state, defining a set of parameters is advisable to determine when a simulation has reached a converged state.

This study employs the Mass-Weighted Average Gas Temperature and Mass-Weighted Average Ethylene Mass Fraction at the outlet as convergence criteria. When further iterations do not yield changes beyond the third decimal point, convergence is deemed achieved. Also, since this study focuses on the system's steady state, intermediate-state results are not saved, allowing the use of pseudo-transient solvers. Future simulations could explore the effects of reactor configuration on the system's dynamic response.

## 2.10 Mesh Sensitivity Study

In CFD simulation, selecting an appropriate mesh density is crucial for achieving accuracy and efficiency. Coarse meshes may introduce spatial discretization errors, while excessively fine meshes prolong simulation times unnecessarily. A Mesh Sensitivity Study (MSS) should be conducted to balance these factors. An MSS involves running the same simulation sequentially using higher mesh densities until the convergence parameters reach a stable value within a defined tolerance. This analysis should be performed for each configuration under consideration for best practices. In this research, the mesh densities were sequentially increased until a 2% change in convergence variables was no longer observed. At this point, the results were deemed independent of spatial resolution, ensuring acceptable accuracy for further analysis.



## 2.11 Simulation Workflow

This research implements the following workflow:

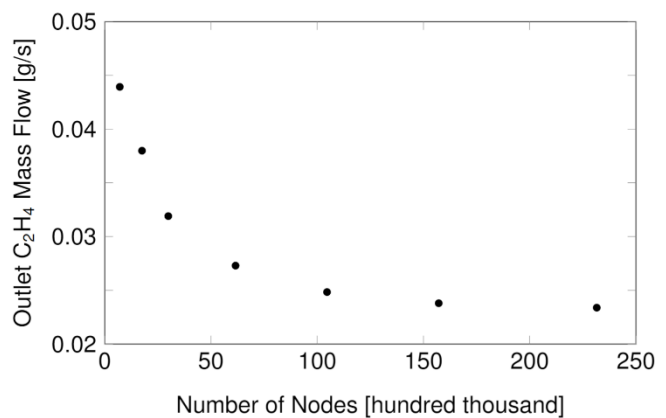
1. Define reactor geometry using Ansys Design Modeler.
2. Discretize the domain using the Ansys Mesh editor.
3. Load Fluent. Activate and set up the energy, viscous, species, and radiation models.
4. Load kinetic model and thermodynamic properties.
5. Establish boundary conditions, starting with an initial mass flow guess.
6. Define convergence criteria.
7. Run simulations until convergence is achieved.
8. Repeat the process from step 5 to determine a first approximation of the mass flow that maximizes ethylene production.
9. With a well-informed guess for the optimal mass flow, conduct a mesh sensitivity study, increasing mesh density by a factor in the range of 1.5–2.0 between cases.
10. Once the final mesh density has been established, optimize the mass flow again.
11. Save results.

## 2.12 Reactor Dimension

Simulations were performed considering lab-scale reactors to lay the groundwork for future experimental efforts. Typical lab-scale EHCs range from 1–2 m long by 5–13 mm diameter. However, to better showcase the potential of internal heating for more compact reactors with larger diameter tubes, a length of 0.5 m and internal diameter of 35 mm (Schedule 40 NPS 1-1/4 piping) were selected.

## 3. RESULTS

Figures 6 and 7 illustrate the outcomes obtained during step 10 of the simulation workflow, presenting a series of fields displaying the parameter distributions within the reaction domain. Analyzing these fields provides crucial insights into the system's performance. For instance, Figure 6 showcases the intense radial thermal gradients of EHCs. Also, since the peak ethylene composition is observed within the unit rather than at the outlet, it suggests the mass flow of 0.275 g/s is too low in this case. Moreover, radial temperature distribution results in higher reaction rates at the tube walls rather than evenly distributed across the cross-section. These Figures were generated using a mesh density of ~150k nodes. As shown in Figure 8, this resolution allows for mesh-independent results in this case. 50



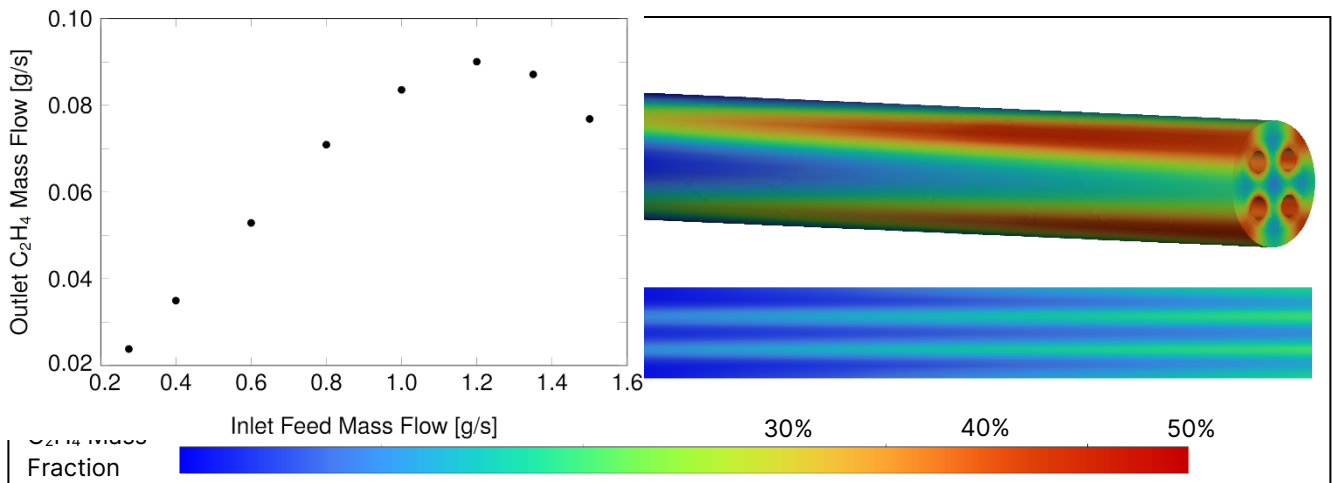
**Figure 8.** Mesh sensitivity study for the externally heated case. Results converge after ~150,000 nodes.



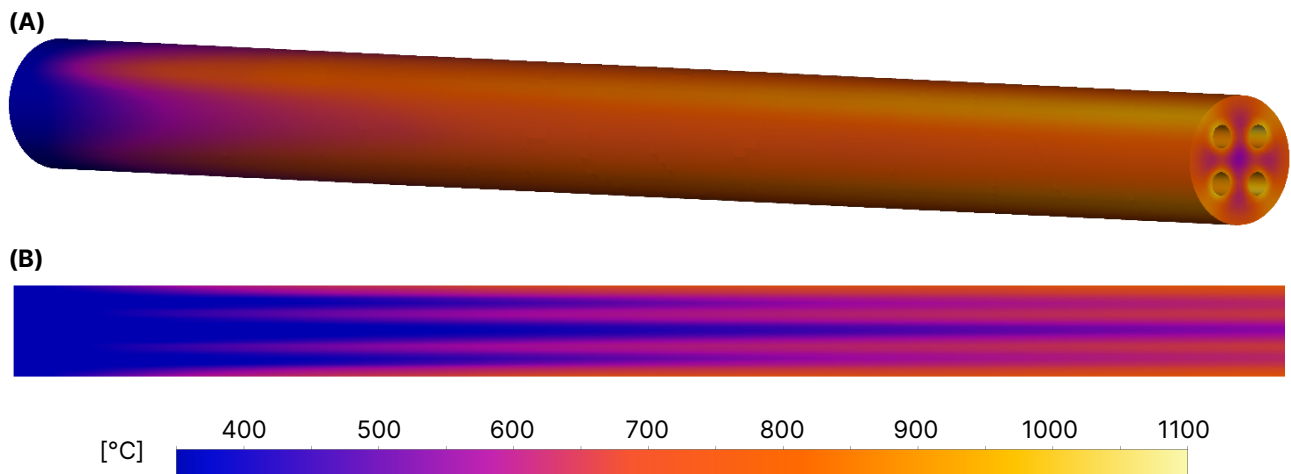
Figure 9 illustrates the optimization process for the inlet feed flow in the externally heated case. Each data point represents a fully converged CFD simulation after the MSS study. The analysis reveals that the optimal feed flow for this configuration is  $\sim 1.2$  g/s, accompanied by a corresponding outlet ethylene feed flow of  $\sim 0.09$  g/s. This specific value emerges as the optimal operating point for the externally heated case.

The same iterative process is systematically repeated for the other configurations as well. For instance, Figures 10 and 11 offer comprehensive 2D and 3D views of the composition and temperature profiles of one of the simulations for the internally heated case with four rods. In contrast to the externally heated configuration, the radial gradients are less pronounced, resulting in a more even radial temperature distribution and higher ethylene production.

**Figure 9.** Feed flow rate optimization for the externally heated case. The optimal point is at 1.2 g/s.



**Figure 10.** (A) External-face and (B) Middle-plane C<sub>2</sub>H<sub>4</sub> mass fraction distribution in an Internally Heated Cracker (IHC) with four rods.



**Figure 11.** (A) External-face and (B) Middle-plane temperature distribution in an Internally Heated Cracker (IHC) with four rods. Results obtained after Mesh Sensitivity Study.



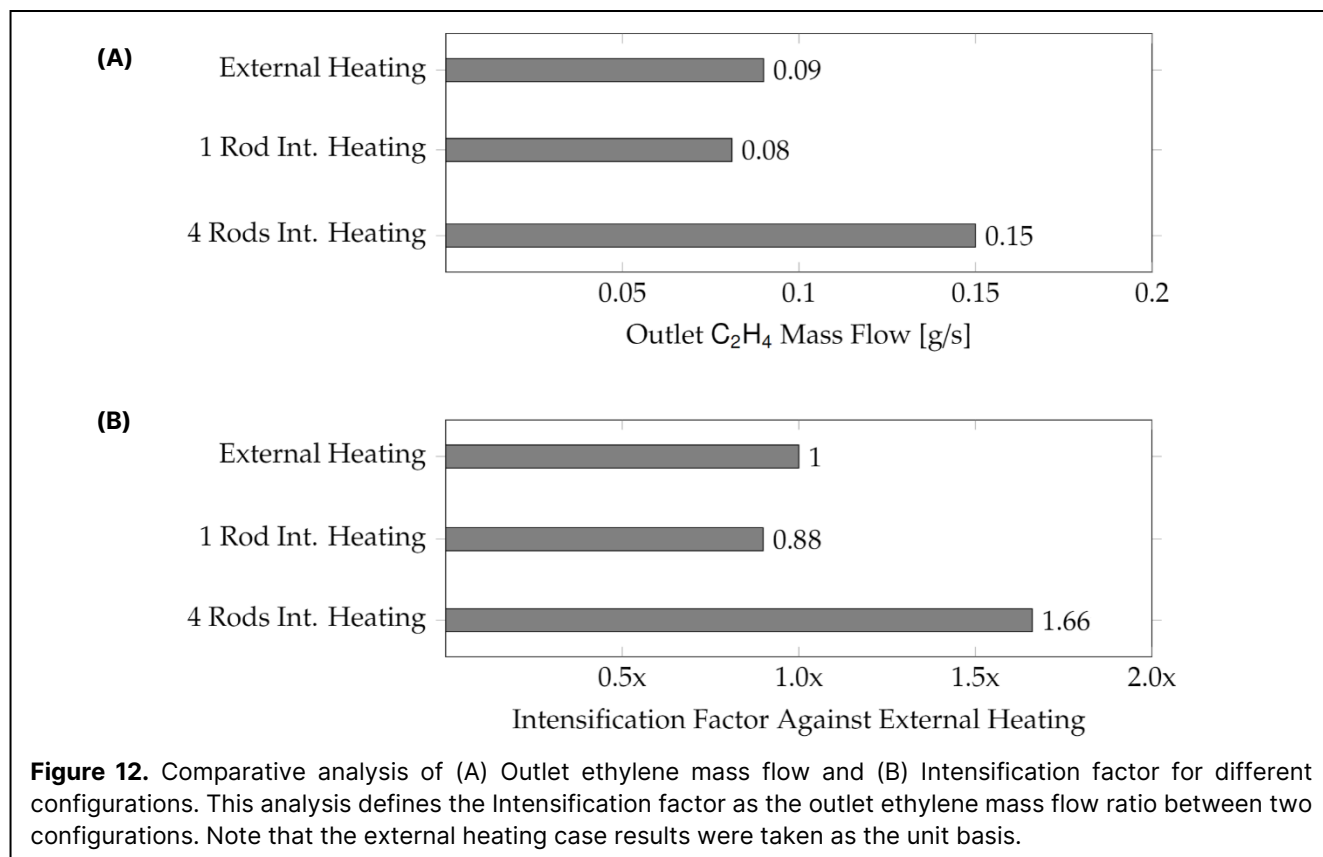


Figure 12 summarizes the results obtained for each configuration after step 11 of the workflow. The results reveal compelling insights. For example, the configuration featuring a single internally heated rod achieves a lower ethylene composition of ~0.08 g/s than the externally heated case, taken as a reference. This behavior is attributed to the limited heat transfer area offered by a single rod. In contrast, the scenario with four heated rods exhibits a remarkable outlet ethylene composition of ~0.15 g/s, surpassing the external heating case by a factor of ~1.66, showing the potential for achieving higher ethylene production than state-of-the-art externally heated designs. This behavior is attributed to more homogeneous set reaction conditions and increased transfer area. Furthermore, these findings encourage further exploration to discover new internally heated configurations capable of achieving even higher performance levels.

The analysis presented in this study sheds light on the intricacies of ethylene production and heralds a promising frontier for developing more efficient cracking reactor designs, poised to redefine the chemical manufacturing landscape toward achieving a Net-Zero future.

## ACKNOWLEDGEMENTS

We would like to acknowledge the National Science Foundation under cooperative agreement no. EEC-1647722 for funding.

## REFERENCES

1. K. Yamaguchi, H. Kishi, and T. Nagasawa, "On the use of Polyethylene Insulation in the Manufacture of Submarine Cables," *IEEE Trans. Electr. Insul.*, no. 1, pp. 75–79, 1977.
2. "Ethylene Oxide (EtO)," EPA. Environmental Protection Agency. [Online]. Available: <https://www.epa.gov/ingredients-used-pesticide-products/ethylene-oxide-eto>
3. "Ethylene: The 'World's most important chemical,'" *American Fuel & Petrochemical Manufacturers (AFPM)*. Aug. 2022. [Online]. Available: <https://www.afpm.org/newsroom/blog/ethylene-worlds-most-important-chemical>
4. H. Pierre and G. Curme, "PRODUCTION OF ETHYLENE," U.S. Patent 1,460,545 (1923)
5. "Landmark for the birthplace of the petrochemical industry," *American Chemical Society*. [Online]. Available: <https://www.acs.org/content/acs/en/education/whatischemistry/landmarks/petrochemical-industry-birthplace.html>
6. H. Zimmermann and R. Walzl, "Ethylene," *Ullmann's Encycl. Ind. Chem.*, 2000.
7. S. Sankaran and et al., "ELECTRICALLY POWERED FURNACES TO HEAT A FEED AND RELATED METHODS," WO 2023/006475 A1, Feb. 2023

8. S. Stevenson, "FURNACE INCLUDING HEATING ZONES WITH ELECTRICALLY POWERED HEATING ELEMENT AND RELATED METHODS." WO 2023/016968 A1, Feb. 2023.
9. K. M. Van Geem, G. J. Heynderickx, and G. B. Marin, "Effect of Radial Temperature Profiles on Yields in Steam Cracking," *AIChE J.*, vol. 50, no. 1, pp. 173–183, Jan. 2004, doi: 10.1002/aic.10016.
10. S. M. Sadrameli, "Thermal/catalytic cracking of hydrocarbons for the production of olefins: A state-of-the-art review I: Thermal cracking review," *Fuel*, vol. 140. Elsevier Ltd, pp. 102–115, Jan. 15, 2015. doi: 10.1016/j.fuel.2014.09.034.
11. "US ethane cracker construction costs rise 1-2 percent year on year," *Reuters Events*. [Online]. Available: <https://www.reutersevents.com/downstream/engineering-and-construction/infographic-us-ethane-cracker-construction-costs-rise-1-2-year-year>
12. "Shale Gas Fuels U.S. Chemical Plant Construction." [Online]. Available: <https://www.enr.com/articles/45920-shale-gas-fuels-us-chemical-plant-construction>
13. A. I. Foster and B. J. Cromarty, "The theory and practice of steam reforming," in *4th European Technical Seminar on Hydrogen Plants, Lisbon*, 2003.
14. "Global Ethylene Market Analysis," *Research and Markets*. [Online]. Available: [https://www.researchandmarkets.com/reports/5743468/global-ethylene-market-analysis-plant-capacity?utm\\_source=GNOM&utm\\_medium=PressRelease&utm\\_code=z75q4z&utm\\_campaign=1828184+-+Global+Ethylene+Market+Analysis+Report+2023%3A+Plant+Capacity%2C+Production%2C](https://www.researchandmarkets.com/reports/5743468/global-ethylene-market-analysis-plant-capacity?utm_source=GNOM&utm_medium=PressRelease&utm_code=z75q4z&utm_campaign=1828184+-+Global+Ethylene+Market+Analysis+Report+2023%3A+Plant+Capacity%2C+Production%2C)
15. Z. Liu, Z. Deng, S. J. Davis, C. Giron, and P. Ciais, "Monitoring global carbon emissions in 2021," *Nature Reviews Earth and Environment*, vol. 3, no. 4. Springer Nature, pp. 217–219, Apr. 01, 2022. doi: 10.1038/s43017-022-00285-w.
16. "Ethylene Market Size To Surpass Around USD 228.69 Billion By 2032." [Online]. Available: <https://www.precedenceresearch.com/ethylene-market>
17. J. KERRY and G. MCCARTHY, *THE LONG-TERM STRATEGY OF THE UNITED STATES Pathways to Net-Zero Greenhouse Gas Emissions by 2050*. 2021. [Online]. Available: <https://www.whitehouse.gov/wp-content/uploads/2021/10/US-Long-Term-Strategy.pdf>
18. "Climate change: what the EU is doing." [Online]. Available: <https://www.consilium.europa.eu/en/policies/climate-change/>.
19. Z. Chen, EA. Rodriguez-Gil, and R. Agrawal, "Toward Carbon Neutrality for Natural Gas Liquids Valorization from Shale Gas," *Ind. Eng. Chem. Res.*, vol. 61, no. 12, pp. 4469–4474, 2022.
20. S. T. Wismann *et al.*, "Electrified methane reforming: A compact approach to greener industrial hydrogen production," *Science (80-. )*, vol. 364, no. 6442, pp. 756–759, 2019.
21. "Joint News Release - BASF, Sabic, and Linde start construction of the world's first demonstration plant for large-scale electrically heated steam cracker furnaces," *BASF*. [Online]. Available: <https://www.basf.com/global/en/media/news-releases/2022/09/p-22-326.html>
22. P. E. Hardisty, T. S. Clark, and R. G. Hynes, "Life cycle greenhouse gas emissions from electricity generation: a comparative analysis of Australian energy sources," *Energies*, vol. 5, no. 4, pp. 872–897, 2012.
23. R. Agrawal, Z. Chen, and P. Oladipupo, "ELECTRICALLY HEATED DEHYDROGENATION PROCESS," U.S. Patent 11,578,019 B2 (2023)
24. V. Balakotaiah and R. R. Ratnakar, "Modular reactors with electrical resistance heating for hydrocarbon cracking and other endothermic reactions," *AIChE J.*, vol. 68, no. 2, Feb. 2022, doi: 10.1002/aic.17542.
25. EA. Rodriguez-Gil, and R. Agrawal, "Introducing Electric Cracking Towers (E-CRATOS) for Net-Zero Ethylene Production," *Ethylene Producers' Committee Vol 36*, 2024.
26. EA. Rodriguez-Gil, and R. Agrawal, "ELECTRICAL HEAT EXCHANGER AND REACTION TECHNOLOGY (E-HEART)," US Patent App. 63/561,869, 2024
27. J. Dalton, *A new system of chemical philosophy - Olefiant Gas*, vol. 2. Smithsonian Institution libraries, 1810.
28. K. M. Sundaram and G. F. Froment, "Modeling of thermal cracking kinetics. 3. Radical mechanisms for the pyrolysis of simple paraffins, olefins, and their mixtures," *Ind. Eng. Chem. Fundam.*, vol. 17, no. 3, pp. 174–182, 1978.
29. O. Olsvik and F. Billaud, "Modelling of the decomposition of methane at 1273 K in a plug flow reactor at low conversion," *J. Anal. Appl. Pyrolysis*, vol. 25, pp. 395–405, 1993.
30. K. M. Sundaram and J. M. Fernandez-Baujain, "Effect of methane and hydrogen during thermal cracking of light hydrocarbons," *AIChE J.*, vol. 34, no. 2, pp. 321–325, 1988, doi: 10.1002/aic.690340216.
31. D. H. B. Branko Ruscic, "Active Thermochemical Tables," *Active Thermochemical Tables - Home*. [Online]. Available: <https://atct.anl.gov>

32. Z. Chen, Y. Li, W. P. Oladipupo, E. A. R. Gil, G. Sawyer, and R. Agrawal, "Alternative ordering of process hierarchy for more efficient and cost-effective valorization of shale resources," *Cell Reports Phys. Sci.*, vol. 2, no. 10, 2021.
33. B. Crynes, L. Albright, and L.-F. T., "Thermal Cracking," *Encycl. Phys. Sci. Technol.*, 2003.
34. R. Agrawal and P. Oladipupo, "METHANE ETHANE CRACKERS," US 11,267,768 B2 (2022)
35. EA. Rodriguez-Gil, and R. Agrawal, "SYSTEMS AND PROCESSES FOR TURBOQUENCHING," US20230391687A1, 2023

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open-access article under the Creative Commons CC-BY-SA licensing terms. Credit must be given to the creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Mathematical Optimization of Separator Network Design for Sand Management

Pooja Zen Santhamoorthy<sup>a</sup>, Selen Cremaschi<sup>a\*</sup>

<sup>a</sup> Auburn University, Department of Chemical Engineering, Auburn, Alabama, USA

\* Corresponding Author: selen-cremaschi@auburn.edu.

---

## ABSTRACT

Sand produced along with well-production fluids accumulates in the surface facilities over time, taking valuable space, while the sand carried with the fluids damages downstream equipment. Thus, sand is separated from the fluid in the sand traps and separators and removed during periodic clean-ups. But at high sand productions, the probability of unscheduled facilities shutdowns increases. Such extreme production conditions can be handled by strategic planning and optimal design of the separator network to enable maximum sand separation at minimal equipment cost while ensuring the accumulation extent is within tolerable limits. This paper develops a mathematical model to optimize the separator network design to maximize sand separation while the sand accumulation extent and total equipment cost are minimal. The optimization model is formulated using multi-objective mixed-integer nonlinear programming (MINLP). The capabilities of the developed model to assist sand management in the separator network are demonstrated with a case study of optimizing the network for two wells producing sand particles of different sizes. A residence time distribution-based model is used to predict sand settling behavior. The developed Pareto Front shows the trade-off between the increase in total sand accumulation rate and total equipment cost for an increase in the fraction of sand settled.

**Keywords:** Optimization, Planning, Oil and Gas, Separator, Sand

## INTRODUCTION

On the surface, sand produced with oil and gas accumulates in the separators, leading to costly unscheduled shutdowns, while the sand carried over damages the downstream equipment [1]. Sand management practices include sand separation from the fluid in the sand traps and separators and removal during periodic clean-ups. Multiple separators are generally necessary to separate the multiphase production containing oil, gas, water, and sand. The extent of sand separation and accumulation in separators depends on the production conditions, separator design, and separator network. Previous studies have developed algorithms for simultaneous optimization of the well production and separation systems for production management and enhancement [2,3]. However, these studies do not account for the sand separation in the separators.

This paper develops a mathematical programming model to optimize separator network design with three

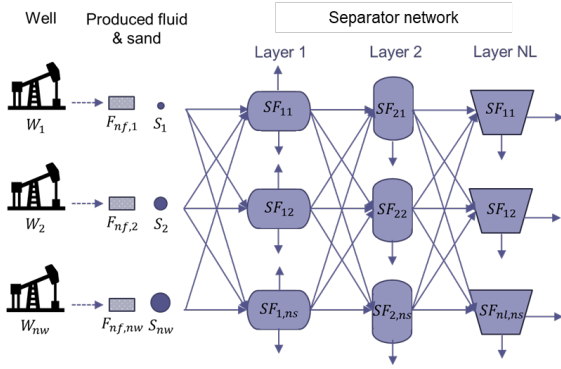
objectives: (i) to maximize the overall sand separation, (ii) to minimize sand accumulation in each separator, and (iii) to minimize the total purchased equipment cost. The model is formulated as a multi-objective mixed-integer nonlinear programming model (MINLP) and solved using epsilon-constraint method.

## PROBLEM STATEMENT

The main goal is to develop an optimization-based framework that assists the design of the separator network by (i) selecting separators from each layer of the separator network for the separation of produced fluid and sand, (ii) determining the design specifications of the selected separators, and (iii) estimating sand settling rate in the separators. The objective is to maximize sand separation from the produced fluid while minimizing sand accumulation in each separator and the total equipment cost of the separators.

The superstructure in Figure 1 illustrates the

separator network. Production fluids ( $F_{nf,nw}$ ) and sand ( $S_{nw}$ ) from wells ( $W_{nw}$ ) enter the network. The sets  $nf \in NF$  and  $nw \in NW$  represent the fluid phases and wells. The layers correspond to different stages of separation, and each layer consists of separators ( $SF_{nl,ns}$ ). The sets  $nl \in NL$  and  $ns \in NS$  represent the layers and separators in each layer. The well fluids are separated into their constituents in each layer.



**Figure 1.** The superstructure of the separator network.

## MULTI-OBJECTIVE OPTIMIZATION MODEL FORMULATION

The three objectives of the separator network design optimization model are: i) to maximize the overall sand separation, ii) to minimize the sand accumulation in each separator, and iii) to minimize the total purchased equipment cost of the separators in the network. The overall sand separation is represented as a fraction of sand produced by the wells that are separated from the produced fluid in the separators ( $FS$ ) (Eq. 1). It is computed using the total sand production rate from the wells and the total sand settling rate in the separators in each layer of the network. The sand accumulation in each separator in the network is minimized by minimizing the summation of the sand accumulation rate in each separator ( $AC$ ). The sand accumulation rate is determined as the rate at which the sand bed height increases in the vessel (Eq. 2). The total purchased equipment cost of the separators network ( $TC$ ) is computed from the equipment cost of each separator in the network (Eq. 3).

$$\begin{aligned} \text{Max } FS &= (\sum_{nl1 \in NL1} ms1'_{nl1} + ms2'_{nl2} \\ &+ \dots + \sum_{nln \in NLN} msn'_{nln}) / \sum_{nw \in NW} Mw_{nw} \quad (1) \end{aligned}$$

$$\text{Min } AC = \sum_{nl1 \in NL1} ac1_{nl1} + \dots + \sum_{nln \in NLN} acn_{nln} \quad (2)$$

$$\text{Min } TC = \sum_{nl1 \in NL1} ce1_{nl1} + \dots + \sum_{nln \in NLN} cen_{nln} \quad (3)$$

The sets  $nw \in NW$  and  $nl1 \in NL1$ ,  $nl2 \in NL2$ , ...,  $nln \in NLN$  represent the oil and gas wells and separators in the different layers of the network, respectively. The parameter  $Mw_{nw}$  represents the sand production rate from each well  $nw$ . The variables  $ms1'_{nl1}$ ,  $ms2'_{nl2}$ , and  $msn'_{nln}$

represent the sand settling rate in each separator in each network layer. The variable  $FS$  represents the fraction of produced sand separated (i.e., settled) in the separators network. The variables  $ac1_{nl1}$  and  $acn_{nln}$  represent the sand accumulation rate in each separator in each network layer and the summation of the accumulation rates is represented by the variable  $AC$ . The variables  $ce1_{nl1}$  and  $cen_{nln}$  represent the purchased equipment cost of the separators in different network layers, and  $TC$  represents the total equipment cost of all the separators in the network.

## Binary Variables

A set of binary variables,  $y1_{nl1}, \dots, yn_{nln}$ , is defined in Eqs. (4) and (5) to represent the selection of separators in each network layer ( $nl1 \in NL1, \dots, nln \in NLN$ ). Another binary variable,  $l1_{nw,nl1}$ , is defined in Eq. (6) to represent the interaction between the wells and the first-layer separators ( $nl1 \in NL1$ ). Similarly, another set of binary variables,  $l2_{nf,nl1,nl2}, \dots, ln_{nf,nln-1,nln}$ , is defined in Eqs. (7) and (8) to represent the interaction between the first- and second-layer separators and the separators in the subsequent  $NLN - 1$  and  $NLN$  layers. If fluid phase separation occurs in the first-layer separators, the individual or mixed phases (the phases that are still not separated) enter the second network along with sand. Thus, the interaction variables  $l2_{nf,nl1,nl2}, \dots, ln_{nf,nln-1,nln}$  are written in terms of the set of fluid phases  $nf \in NF$ . All fluid phases from the  $NLN$  layer separators exit the separator network.

$$y1_{nl1} = \begin{cases} 1, & \text{if separator } nl1 \text{ in } NL1 \text{ layer is selected} \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

$$yn_{nln} = \begin{cases} 1, & \text{if sep. } nln \text{ in } NLN \text{ layer is selected} \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

$$l1_{nw,nl1} = \begin{cases} 1, & \text{if fluid from well } nw \text{ enters} \\ & \text{sep. } nl1 \text{ in } NL1 \text{ layer} \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

$$l2_{nf,nl1,nl2} = \begin{cases} 1, & \text{if phase } nf \text{ from sep. } nl1 \text{ in } NL1 \\ & \text{layer enters sep. } nl2 \text{ in } NL2 \text{ layer} \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

$$ln_{nf,nln-1,nln} = \begin{cases} 1, & \text{if phase } nf \text{ from sep. } nln - 1 \\ & \text{in } NLN - 1 \text{ enters} \\ & \text{sep. } nln \text{ in } NLN \text{ layer} \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

$$\forall nf \in NF, nw \in NW, nl1 \in NL1, \dots, nln \in NLN$$

## Parameters

The parameters are defined for the produced fluid, produced sand, separators design, and separators cost models. (i) Produced fluid: The volumetric flow rate of each fluid phase  $nf$  ( $Q_{nf,nw}$ ) produced from each well  $nw$  are given. Similarly, the physical properties of each fluid phase ( $\varepsilon_{nf,nw}$ ), such as density and viscosity, are given for each well  $nw$ . (ii) Produced sand: The sand density



( $\rho_s$ ), sphericity ( $\psi$ ), and the sand production rate ( $M_{w_{nw}}$ ) and mean particle diameter of sand produced ( $d_{nw}$ ) from each well  $nw$  are given. The fraction of produced sand dispersed in each fluid phase ( $Fd_{nf,nw}$ ) is given for each well  $nw$ . (iii) Separators design: Lower and upper limits are given for each design specification of the separators in the network,  $Dl$  and  $Du$ , respectively. (iv) Separators cost models: The cost model parameters,  $\alpha_{1nl1}, \alpha_{2nl2}, \dots, \alpha_{nln}$  specific to separators in each network layer are given.

## Constraints

If the separator  $nl1$  is selected in the first layer of the network, the binary variable  $y_{1nl1}$  takes a value of one, and if the separator has input from the well  $nw$ , the binary  $l_{1nw,nl1}$  takes a value of one. The fraction of the total produced fluid and sand from well  $nw$  that enters the separator  $nl1$  is represented by the variable  $fr_{1nw,nl1}$ . Thus, it can take a value of zero to one. If the separator  $nl1$  does not have input from the well  $nw$ , the variables  $l_{1nw,nl1}$  and  $fr_{1nw,nl1}$  take a value of zero. This disjunction is given in Eq. (9). Similarly, the interaction between separators in the subsequent layers  $NLN - 1$  and  $NLN$  are represented by the binary variable  $ln_{nf,nln-1,nln}$  and the fraction of each fluid phase  $nf$  entering the separator in  $NLN$  layer from  $NLN - 1$  layer is represented by the variable  $fr_{nf,nln-1,nln}$  (Eq. 10).

$$\left[ \begin{array}{l} l_{1nw,nl1} = 1 \\ 0 \leq fr_{1nw,nl1} \leq 1 \end{array} \right] \vee \left[ \begin{array}{l} l_{1nw,nl1} = 0 \\ fr_{1nw,nl1} = 0 \end{array} \right] \quad (9)$$

$\forall nw \in NW, nl1 \in NL1$

$$\left[ \begin{array}{l} ln_{nf,nln-1,nln} = 1 \\ 0 \leq fr_{nf,nln-1,nln} \leq 1 \end{array} \right] \vee \left[ \begin{array}{l} ln_{nf,nln-1,nln} = 0 \\ fr_{nf,nln-1,nln} = 0 \end{array} \right] \quad (10)$$

$\forall nf \in NF, nln - 1 \in NLN - 1, nln \in NLN$

The design specifications, fluid flow rate and properties, and sand properties in the separator  $nln$  selected in  $NLN$  network layer are given in Eq. (11). The design specifications of the separator ( $Dn_{nln}$ ) should be within the specified lower and upper bounds ( $Dl$  and  $Du$ ). The total volumetric rate of each fluid phase ( $qn_{nf,nln}$ ) entering the separator is computed from the fraction of each fluid phase entering the separator  $nln$  from each separator  $nln - 1$  in the previous layer  $NLN - 1$  ( $fr_{nf,nln-1,nln}$ ) and the volumetric rate of each fluid phase in the previous layer separators ( $qn_{nf,nln-1}$ ). Similarly, the rate at which sand of size  $nw$  dispersed in phase  $nf$  enters the separator  $nln$  ( $men_{nf,nw,nln}$ ) is computed using the amount of sand carried by the fluid phase from the previous layer separators ( $mcn_{nw,nf,nln-1}$ ) and the fraction of that fluid phase entering the separator  $nln$ ,  $fr_{nf,nln-1,nln}$ .

In the previous description, the set  $nw \in NW$  is used to denote the different particle sizes from different wells in the network. The physical properties of each fluid phase are estimated ( $\epsilon n'_{nf,nln}$ ) based on the properties of fluid from each previous layer separator ( $\epsilon n_{nf,nln-1}$ ). The design specifications, fluid flow rate and properties, and sand amount variables take a value of zero if the separator is not selected in the network, i.e., when  $y_{nln} = 0$ . The design specifications ( $Dn_{nln}$ ) depend on the total volumetric flow rate of each fluid phase in the separator ( $qn_{nf,nln}$ ) and other separator type-specific factors ( $\sigma n_{nln}$ ) (Eq. 12).

$$\left[ \begin{array}{l} y_{nln} = 1 \\ Dl \leq Dn_{nln} \leq Du \\ qn_{nf,nln} = \sum_{nln-1} fr_{nf,nln-1,nln} qn_{nf,nln-1} \\ men_{nf,nw,nln} = fr_{nf,nln-1,nln} mcn_{nw,nf,nln-1} \\ \epsilon n'_{nf,nln} = f(fr_{nf,nln-1,nln}, \epsilon n_{nf,nln-1}) \end{array} \right] \vee \left[ \begin{array}{l} y_{nln} = 0 \\ Dn_{nln} = 0 \\ qn_{nf,nln} = 0 \\ men_{nf,nw,nln} = 0 \\ \epsilon n'_{nf,nln} = 0 \end{array} \right] \quad (11)$$

$$Dn_{nln} = f(qn_{nf,nln}, \sigma n_{nln}) \quad (12)$$

$$\forall nf \in NF, nw \in NW, nln - 1 \in NLN - 1, nln \in NLN$$

Sand particles of different sizes ( $d_{nw}$ ) from different wells are settling in the separator  $nln$ . The fraction of sand particles of each size initially dispersed in phase  $nf$  that settle to the vessel bottom ( $Fsn_{nw,nf,nln}$ ) depends on the properties of the fluid ( $\epsilon n'_{nf,nln}$ ) it is dispersed in, separator design specifications ( $Dn_{nln}$ ), and particle diameter ( $d_{nw}$ ) (Eq. 13). Similarly, the fraction of each sand size carried by each phase from the separator  $nln$  ( $Fcn_{nw,nf,nln}$ ) is estimated. Some amount of sand initially dispersed in the fluid phase  $nf = 1$  can move to the phase  $nf = 2$ , and vice versa. This amount is accounted for by adding the variables  $\delta sn_{nw,nf,nl1}$  and  $\delta cn_{nw,nf,nl1}$  to the fraction settled and fraction carried, respectively. This variable can take a value in the range  $[-1, 1]$  based on the sand dispersion and settling model selected for the separator. The rate of sand settling at the bottom ( $msn_{nw,nf,nln}$ ) and sand being carried ( $mcn_{nw,nf,nln}$ ) for different sand sizes are computed from the amount of each sand size entering the separator ( $men_{nf,nw,nln}$ ). If the separator  $nln$  is not selected in the network, all these variables take a value of zero.

$$\left[ \begin{array}{l} y_{nln} = 1 \\ Fsn_{nw,nf,nln} = f(\epsilon n'_{nf,nln}, Dn_{nln}, d_{nw}) + \delta sn_{nw,nf,nln} \\ Fcn_{nw,nf,nln} = f(\epsilon n'_{nf,nln}, Dn_{nln}, d_{nw}) + \delta cn_{nw,nf,nln} \\ msn_{nw,nf,nln} = men_{nf,nw,nln} Fsn_{nw,nf,nln} \\ mcn_{nw,nf,nln} = men_{nf,nw,nln} Fcn_{nw,nf,nln} \end{array} \right]$$

$$\forall \begin{cases} \mathbf{yn}_{nl_n} = \mathbf{0} \\ FSN_{nw,nf,nl_n} = 0 \\ FCN_{nw,nf,nl_n} = 0 \\ msn_{nw,nf,nl_n} = 0 \\ mcn_{nw,nf,nl_n} = 0 \end{cases} \quad (13)$$

$$\forall nf \in NF, nw \in NW, nl_n \in NLN$$

The fluid and sand produced from each well should be treated in the separator network. Thus, the total fraction of fluid entering from each well into different separators in the first network layer ( $fr1_{nw,nl_1}$ ) is equal to one (Eq. 14). The fluid phases exiting separators in each network layer either enter the separators in the subsequent layer or exit the network based on the interaction variable  $frn_{nf,nl_n-1,nl_n}$  given in the Eq. (10).

$$\sum_{nl_1} fr1_{nw,nl_1} = 1 \quad \forall nw \in NW \quad (14)$$

The total amount of sand settled in each separator  $nl_n$  ( $msn'_{nl_n}$ ) is computed as the total amount of each sand size dispersed in each phase that has settled to the separator bottom ( $msn_{nw,nf,nl_n}$ ) (Eq. 15). The sand accumulation rate (i.e., rate of increase of sand bed height) in each separator is computed based on the amount of sand particles of each size settled in the separator and the void space between the settled particles. The void space is accounted for using void fraction, which depends on the mean particle size, particle size distribution, and the sphericity of the particles ( $\psi$ ) [7]. The mean particle size and particle size distribution are estimated based on the amount of different particle sizes settled in the separator. Thus, the void fraction of the sand particles settled in the separator ( $\phi_{nl_n}$ ) depends on the particle sizes present ( $d_{nw}$ ) and the amount of each particle size settled in the separators ( $msn_{nw,nf,nl_n}$ ) (Eq. 16). The sand accumulation rate in the separator ( $acn_{nl_n}$ ) is computed using the void fraction ( $\phi_{nl_n}$ ), the total amount of sand settled in the separator ( $msn'_{nl_n}$ ), sand density ( $\rho_s$ ), and cross-sectional area of the separator ( $CSn_{nl_n}$ ) (Eq. 17).

$$msn'_{nl_n} = \sum_{nw,nf} msn_{nw,nf,nl_n} \quad \forall nl_n \in NLN \quad (15)$$

$$\phi_{nl_n} = f(d_{nw}, msn_{nw,nf,nl_n}) \quad (16)$$

$$acn_{nl_n} = msn'_{nl_n} / (CSn_{nl_n} \rho_s (1 - \phi_{nl_n})) \quad (17)$$

$$\forall nf \in NF, nw \in NW, nl_n \in NLN$$

The purchased equipment cost of each separator depends on its design specifications. The cost of the separator  $nl_n$  ( $cen_{nl_n}$ ) is estimated based on its design specifications ( $Dn_{nl_n}$ ) using the separator-specific cost model ( $\alpha n_{nl_n}$ ) (Eq. 18).

$$cen_{nl_n} = f(Dn_{nl_n}, \alpha n_{nl_n}) \quad \forall nl_n \in NLN \quad (18)$$

## CASE STUDY

The equations (1) to (18) yield a deterministic multi-

objective mixed-integer nonlinear programming model (MINLP) that maximizes sand separation, minimizes sand accumulation, and minimizes the total purchased equipment cost of the separator network. The disjunctive constraints in the model are reformulated using Big-M Reformulation [4]. In the developed multi-objective model, a set of decision variables improving one objective's value could negatively affect another objective and vice versa. Thus, a single optimal solution is not possible; rather, a set of optimal solutions is obtained, which forms the Pareto front [5]. This work uses the epsilon-constraint method [6] to obtain the Pareto front for the developed model. The model is solved by selecting one of the objective functions as the only objective and the remaining objective functions as constraints. Then, the constraint bounds are systematically varied to obtain the Pareto front.

The capabilities of the developed model to optimize a separator network design for sand management are demonstrated with a test problem. The well production and the separation of the produced fluid and sand in the separator network are shown in Figure 2. Two wells producing oil, gas, water, and sand are considered. The amount of fluid produced, the fluid's physical properties, and the sand's size differ for the two wells. It is assumed that the production from the two wells is completely mixed before it enters the separator network. The produced fluid now contains two different sizes of sand particles, which are assumed to be evenly dispersed in the liquid phases (oil and water). The produced fluid and sand particles enter the separator network containing two layers of separators, and each layer contains two separators. The separators in the first layer separate the produced fluid into its constituents: oil, water, and gas. A fraction of the sand particles is separated from the produced fluid by settling to the bottom of the separators. The remaining sand particles are carried by oil and water. These liquid phases carrying sand enter the separators in the second layer, where more sand particles are separated from the liquid. It is assumed that liquid phases separated in the first-layer separators cannot enter the same second-layer separator, i.e., the second-layer separators can handle sand/oil or sand/water only. Thus, a two-phase vertical gravity separator is used for separation in the second layer.

The design specifications of the four-phase gravity separator are given in Eqs. (19-21). The vessel diameter ( $D_s$ ) is selected based on the volumetric flow rate of liquid (oil and water) separated in the vessel ( $Q_o$  and  $Q_w$ ) and the mean residence time of the liquid phases in the vessel ( $\tau$ ). The separator height ( $H_s$ ) is selected based on the ratio of the height to the diameter for the separator ( $r$ ). The vessel's capacity should be large enough that the liquid level is within the maximum allowable level ( $H_{s,lmax}$ ) (Eq. 19). A minimum vessel diameter ( $D_{min}$ ) is required to

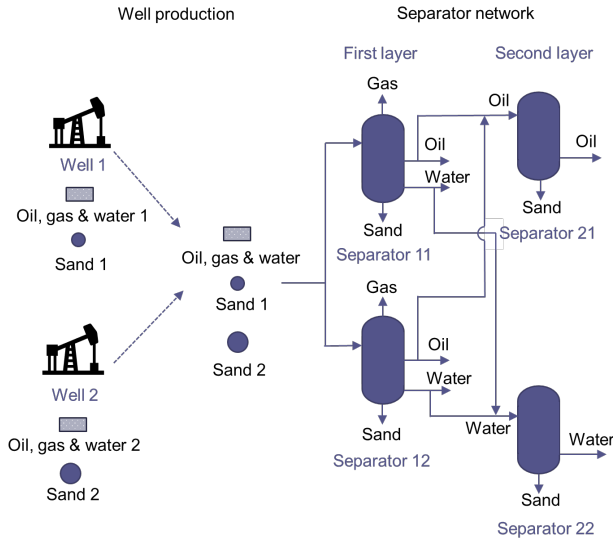
slow the gas to a velocity at which the liquid droplets settle and separate from the gas phase [8]. This minimum diameter depends on the volumetric flow rate of the gas ( $Q_g$ ) and the settling velocity of liquid droplets in the gas phase in the separator ( $v_l$ ) (Eq. 20). Thus, the vessel diameter should be greater than the minimum diameter (Eq. 21). The two-phase gravity separator has only sand/oil or sand/water phases, and it is also designed using Eqs. (19-21).

$$(Q_o + Q_w)\tau/(\pi/4D_s^2H_s) = H_{max}/H_s \quad (19)$$

$$D_{min} = [4Q_g/(\pi v_l)]^{0.5} \quad (20)$$

$$D_s \geq D_{min} \quad (21)$$

Sand settles at the bottom of the gravity separator when its settling time is less than the residence time of the fluid it is dispersed. The residence time of the fluid is expressed using Residence Time Distribution (RTD) models [9]. In our previous work [10], an RTD-based model was developed to compute the fraction of sand settled in a vertical gravity separator. This model is used to estimate the extent of sand separation in each separator in the network.



**Figure 2.** Well production (oil, gas, water, and sand) from two wells enters a separator network with two layers. The separators in the first layer separate the produced fluids. Sand is separated from the produced fluids in both layers.

The model assumptions for the sand settling behavior in a four-phase gravity separator are: (i) As oil and water separate inside the vessel, the sand particles initially dispersed in the oil move to the oil column, while the particles dispersed in the water move to the water column. (ii) Sand particles in the oil column gradually settle with a constant terminal settling velocity, during which a

fraction of the sand particles get carried away by the flowing oil. Similarly, in the water column, the sand particles settle down at a constant velocity specific to the water phase, and the water carries a fraction of these particles. (iii) The remaining sand particles in the water column eventually settle down at the bottom of the separator. The remaining particles in the oil column enter the water phase and settle at a constant velocity. The water carries away some particles, and the rest settle at the separator's bottom.

The terminal sand settling velocity in the oil and water phases is estimated using Stokes' law [11]. The settling time of a sand particle in a liquid column is estimated as the time required to settle through the entire liquid column height at terminal velocity. The distribution of residence time ( $E(t)$ ) is expressed by the RTD model for a single CSTR [9] (Eq. 22). The mean residence time is represented as  $\tau$ .

$$E(t) = \frac{e^{-t/\tau}}{\tau} \quad (22)$$

The fraction of sand particles initially dispersed in the oil phase is  $Fd_o$ , and the rest in the water phase ( $Fd_w$ ). The fraction of sand dispersed in oil fraction with residence time less than sand settling time in oil ( $ts_o$ ) is carried away by oil ( $F_{o,co}$ ), and it is estimated by Eq. (23). The remaining fraction of sand in oil ( $F_{o,ew}$ ) settles through the oil column and enters the water phase (Eq. 24). Thus, two sets of sand particles are dispersed in water, particles initially dispersed in water ( $Fd_w$ ) and particles that entered the water from oil ( $F_{o,ew}$ ). Sand dispersed in the water fraction with residence time less than sand settling time ( $ts_w$ ) is carried away by water. Thus, a fraction of sand initially dispersed in water, and another fraction that entered the water from oil gets carried by water,  $F_{w,cw}$  and  $F_{o,cw}$ , respectively (Eqs. 25 and 26). The remaining fraction of sand in the water settles to the separator bottom ( $F_s$ ) (Eq. 27).

$$F_{o,co} = Fd_o \int_0^{ts_o} E(t_o) dt_o \quad (23)$$

$$F_{o,ew} = Fd_o - F_{o,co} \quad (24)$$

$$F_{w,cw} = Fd_w \int_0^{ts_w} E(t_w) dt_w \quad (25)$$

$$F_{o,cw} = F_{o,ew} \int_0^{ts_w} E(t_w) dt_w \quad (26)$$

$$F_s = (Fd_w - F_{w,cw}) + (F_{o,ew} - F_{o,cw}) \quad (27)$$

The void fraction of the sand particles settled at the separator bottom is estimated using a semi-empirical relation [7] developed for a loosely packed bed.

The purchased equipment cost of the separators ( $C_e$ ) is approximated using the cost model for vertical pressure vessel [8] (Eq. 28). The equipment cost depends on the shell mass of the separator ( $SM$ ), which is a function of the separator diameter, height, and shell

thickness. The equipment cost is in the United States Gulf Coast basis, January 2006 (Chemical Engineering (CE) index = 478.6 and Nelson-Farrar Cost Index (NFCI) = 1961.6). The cost is converted to 2023.

$$C_e = -10000 + 600 SM^{0.6} \quad (28)$$

## RESULTS AND DISCUSSIONS

In the test problem, the total fluid production rate from the two wells is 50 m<sup>3</sup>/h of oil, 50 m<sup>3</sup>/h of water, and 900 m<sup>3</sup>/h of gas. Sand is produced along with wellbore fluid at a rate of 26.5 kg/h. The sand concentration in the liquid is 0.01 vol%. The particles are spherical and have diameters of 150 microns and 50 microns. The sand density is 2650 kg/m<sup>3</sup>. The produced oil, water, and gas densities are 700 kg/m<sup>3</sup>, 990 kg/m<sup>3</sup>, and 23.6 kg/m<sup>3</sup>, respectively. The viscosities of the oil and water are 0.77 x 10<sup>-3</sup> kg/m/s and 0.55 x 10<sup>-3</sup> kg/m/s. The lower and upper bounds for the separator diameter ( $D_s$ ) are 1 m and 3.5 m, respectively, and for the mean residence time of production fluids in the separator ( $\tau$ ), they are 10 mins and 20 mins. The height-to-diameter ratio ( $r$ ) for the separator is 3. The maximum liquid level in the separator ( $H_{s,max}$ ) is 50% of the separator height. The shell thickness is assumed to be 5 mm.

The optimization models are formulated in Python V3.8.6 using PYOMO V6.4.1. The MINLP models are solved using BARON V23.11.13 to 5% optimality gap, all on an Intel Xeon Gold 6248R 3 GHz processor with 48 cores and utilizing a maximum of 100 GB RAM.

Firstly, the separator network optimization model is solved for the three objectives ((i) maximizing sand separation, (ii) minimizing equipment cost, and (iii) minimizing accumulation rate) separately to obtain the bounds for the Pareto front. The results for separator design specifications, separation extent, accumulation rate, and solution time for solving for each objective are given in Table 1. (i) For the given production conditions, a maximum of 84% of the produced sand can be separated from the production fluid with the separator network under study. One large separator (D=2.98 m) and another small separator (D=1.18 m) are selected for the first-layer separation, and the residence time of production fluids is 20 mins in both separators. For second-layer separation, two more equally sized separators are selected to separate sand from the oil and water phases further, with a residence time of 20 mins in the separator. (ii) To minimize the total cost of the network (\$43,000), a single separator (D=2.4 m) with a residence time of 10 mins is selected to separate the production fluids to their constituents. (iii) A single large separator (D=3.04 m) is selected to ensure that the bed height increase rate is minimal (0.93 m/week). Only 56% of the sand is separated with this network design. The highest sand separation (84%) is achieved with a total network cost of \$107,000

and a weekly total sand accumulation rate of 1.77 m in the separators in the network. The solution time is the highest while solving for the equipment minimization objective (6958 CPU s).

**Table 1:** Sand separation extent, accumulation rate, network cost, network design specifications, and solution times for three cases: (i) maximum sand separation, (ii) minimum equipment cost, and (iii) minimum total accumulation rate.

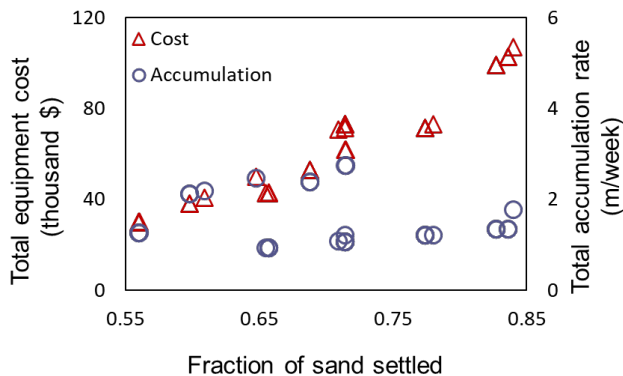
Design spec. / Solution time	Results
<b>(i) Maximum sand separation</b>	
Fraction of sand settled*	0.84
Total cost (thousand \$)	107
Total accumulation rate	1.77
Layer 1	
Diameter (m)	1.18 2.98
Residence time (min)	20 20
Layer 2	
Diameter (m)	2.37 (O) 2.37 (W)
Residence time (min)	20 (O) 20 (W)
Solution time (CPU s)	355
<b>(ii) Minimum equipment cost</b>	
Total cost* (thousand \$)	30
Fraction of sand settled	0.56
Total accumulation rate	1.26
Layer 1	
Diameter (m)	2.4
Residence time (min)	10
Layer 2	
Solution time (CPU s)	Not selected
<b>(iii) Minimum accumulation rate</b>	
Total accumulation rate* (m/week)	0.93
Fraction of sand settled	0.65
Total cost (thousand \$)	43
Layer 1	
Diameter (m)	3.04
Residence time (min)	20
Layer 2	
	Not selected

\* denotes the model objective for each case  
O and W denote the two-phase separators  
for oil/sand and water/sand separation

A Pareto front plot has been developed by plotting



the total equipment cost and total accumulation rate in the network against the fraction of total sand settled in Figure 3. For low sand separation extent, a single separator is selected in the network. A larger separator is used to decrease the accumulation rate, while a smaller residence time is used to decrease the equipment cost. To achieve higher sand separation, a second separator is added in the second layer to separate the sand from the oil phase further while minimizing accumulation. Cost is reduced for a higher separation extent by using two smaller separators in the first layer and a second-layer separator for the water phase, in which the sand settling rate is higher than that in oil. But this arrangement significantly increases the sand accumulation rate, as seen for 71% separation extent. Maximum sand separation is achieved by using all the separators in the network. A high residence time and smaller diameter are preferred for these separators.



**Figure 3.** Pareto front for separator network design.

## CONCLUSION AND FUTURE WORK

A multi-objective separator network optimization model was developed to maximize the sand separation while minimizing the total equipment cost and total sand accumulation rate in the separators. The developed model was used to obtain a Pareto front for a separator network with two separation stages containing two separators each. The Pareto front can be used for the selection and design of the separators to achieve a desired sand separation extent with the separator network, without high accumulation in the separators and at less equipment cost.

This work will be extended to use the developed model for solving larger separator networks with more design and decision flexibility, such as optimizing the height-by-diameter ratio and determining the conditions under which the production fluids from two wells should/should not be separated in the same separator. Reformulation approaches will be investigated for reducing the solution time of the developed MINLP model.

## REFERENCES

- Woodroof, N. The importance of sand monitoring in separator tanks - Part One. *Oilfield Technology* (2019).
- Giorgio, V., Danilo, A., Marco, D., & Almatasem, S. Integrated production optimization and surface facilities management through advanced optimization techniques. *OnePetro* (2012).
- Zainal, S. A., Daud, W. R., Putra, Z. A., & Nor, N. Integrated constraints optimization for surface and sub-surface towards CAPEX free maximizing production. *Materials Science and Engineering* (Vol. 458, No. 1, p. 012045) (2018).
- Raman, R., & Grossmann, I. E. Modelling and computational techniques for logic based integer programming. *Computers & Chemical Engineering*, 18(7), 563-578 (1994).
- Gunantara, N. A review of multi-objective optimization: Methods and its applications. *Cogent Engineering*, 5(1), 1502242 (2018).
- Haimes, Y. On a bicriterion formulation of the problems of integrated system identification and system optimization. *IEEE transactions on systems, man, and cybernetics*, (3), 296-297 (1971).
- Hoffmann, A. C., & Finkers, H. J. A relation for the void fraction of randomly packed particle beds. *Powder Technology*, 82(2), 197-203 (1995).
- Towler, G., & Sinnott, R. Chemical Engineering Design: Principles, Practice and Economics of Plant and Process Design. Butterworth-Heinemann (2021).
- Fogler, H. S. Essentials of Chemical Reaction Engineering: Essenti Chemica Reactio Engi. Pearson Education (2010).
- Santhamoorthy, P., Willams, B., Sambath, K., Subramani, H. J., and Cremaschi, S. Sand management in wellbore and surface facilities using HYSYS. *2021 AIChE Annual Meeting* (2021).
- Stokes, G. G. On the effect of the internal friction of fluids on the motion of pendulums (1851).

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Sustainable Process Systems Engineering – You’re Doing It Wrong!

Raymond L. Smith<sup>a\*</sup>

<sup>a</sup> U.S. Environmental Protection Agency, Office of Research and Development, Center for Environmental Solutions and Emergency Response, 26 W. Martin Luther King Dr., Cincinnati, OH 45268 USA

\* Corresponding Author: [smith.raymond@epa.gov](mailto:smith.raymond@epa.gov)

---

## ABSTRACT

Most studies in process systems engineering are applying incomplete methods when incorporating sustainability. Including sustainability is a laudable goal, and practitioners are encouraged to develop systems that promote economic, environmental, and social aspects. Ten methods that are often overlooked in performing sustainable process systems engineering are listed in this effort and discussed in detail. Practitioners are encouraged to create designs that are inherently safer, to be more complete in their identification of process chemicals used and released, to be complete in their definitions of supply chains, and to apply additional environmental impact categories. Other methods point to items that are factors in process systems engineering such as disruptive recycling, robust superstructures for optimizations, and employing complete sets of objectives. Finally, users should be aware that sustainability tools are available, which might have been outside of their awareness.

---

**Keywords:** Optimization, Process Design, Supply Chain, Life Cycle Analysis, Environment, Sustainability

## INTRODUCTION AND BACKGROUND

Practitioners of process systems engineering stand at the cusp of chemical process design / analysis and life cycle assessment. For those interested in sustainability the methods available to users are growing in popularity but perhaps remain still largely unknown.

In an effort to promote sustainability, studies that developed into the modern methods of life cycle assessment (LCA) were first reported in the 1960s [1]. These studies considered cumulative energy use and comparisons of different beverage containers to quantify the use of natural resources and releases to the environment. Today, rules are being set to formally compare products within categories [2], and LCA studies can model regions [3] and the whole economy [4].

Unlike the above evaluations of supply chain (i.e., cradle-to-grave) sustainability, efforts to optimize chemical processes have smaller single-system domains. The potential addition of new chemicals or technologies into a process makes them open ended, and demonstrated methods for the conceptual design of processes are available [5]. A subset of the open-ended problems can

be optimized, and this is where various process system engineering methods realize their power.

These fields cross pollinate as described in a review of design methods for the environment by Cano-Ruiz and McRae [6]. Alternative generation for designs is combined with optimization methods. One can consider the sustainability of processes within a system by continually expanding the system boundaries to the enterprise, life cycle, economy, and ecosystem [7]. Most studies which make an attempt towards sustainability are incomplete, and so a listing of commonly overlooked methods is offered here.

## TEN UNDERVALUED METHODS OF SUSTAINABLE PROCESS SYSTEMS ENGINEERING

Awareness of methods to improve sustainable process and system design are available as listed in Table 1. Following the table, a series of descriptions will be provided to describe each item in more detail.

**Table 1:** Ten methods for sustainable process systems engineering that are undervalued.

Number	Name
1	Use Inherently Safer Designs
2	Incorporate Inputs / Outputs of Complete Supply Chains
3	Account for All Chemicals and Releases
4	Define Final Fate / Destination for Every Chemical
5	Include All Emissions, Discharges, and Solid Wastes
6	Apply Additional Impact Categories
7	Recognize the Disruptive Nature of Recycling to Design / Optimization
8	Create a Robust Superstructure of Alternatives
9	Use Complete Sets of Objectives
10	Realize Sustainability Tools are Available

### Use Inherently Safer Designs

Inherently safer designs avoid circumstances that could cause accidents as well as chemicals that are unnecessarily toxic. This is one of the original principles of green chemistry, which are listed along with engineering principles in one reference [8]. While not repeated here, the intent is to include all of these principles in this listing of methods for process systems engineers to consider. Only designs that advance these principles will be aligned with sustainability.

### Incorporate Inputs / Outputs of Complete Supply Chains

Designing chemical processes is an open-ended problem with the potential to introduce various chemicals to be part of a product formulation, reacted, or to ease processing. Each chemical introduced requires its own supply chain of processes to manufacture and transport the used chemical. This complete inventory of reactants, solvents, catalysts, processing aids, and cleaning agents is not simple to assemble [9], but the reaction products may be a much longer and less well-known list. Each of these byproducts should be considered for where it ends up and the output system of processes required to handle non-products. Some of these non-product processes may involve flaring, wastewater treatment, hazardous waste treatment, or recycling, each with its own supply chain and releases to the environment. Many of these processes are never seen in studies that show process system diagrams. Therefore, in many cases of process design and analysis, it can generally be concluded that little attention is paid to the fact that every input / output to / from a system requires an upstream / downstream supply chain.

### Account for All Chemicals and Releases

The reactions most process systems engineering studies and life cycle inventory databases consider are overly simplified. As an example, acetic acid production can be modeled according to the stoichiometry of carbon monoxide and methanol reacting to form acetic acid. The actual components found in a real process are often much higher [10]. Beyond the process inputs, chemicals might be introduced to a process for a number of reasons: absorption, boiler feedwater circulation, catalysis, cleaning, cooling tower circulation, input water treatment, wastewater treatment, etc. Some examples of these are presented in the context of early-stage process development [11]. All of these inputs might be released to the environment. In addition, boilers, cooling towers, and fugitive emissions can dramatically increase the number of chemicals released to the environment.

### Define Final Fate / Destination for Every Chemical

When considered holistically, one can envision that every chemical must eventually be reacted, recycled, released, or treated. Certainly, some chemicals go into products, but holistically they will have to realize one of these fates. The challenge for the practitioner of process systems engineering is to logically identify the chemicals present and then track their fates. First, one can apply a version of the methodology described by Douglas [12] where each reactor effluent component is given a destination code. Products and by-products exit as product streams, but impurities of these products will likely be in other streams. In addition, consider that there will be other impurities in the product streams. Unused inputs will mostly be recycled, where possible, but input impurities and unrecycled quantities will exit in various streams. Some streams may have a high enough energy content to legitimately be used as fuel. The other streams will exit as process wastes, where their phase will dictate their form and whether they are vented, process liquid wastes, or solid wastes. Where captured, each of these can be treated with absorption, flaring, land disposal, thermal oxidation, wastewater treatment, etc., and each of these processes has its own resource use and releases to the environment [13]. In addition to exiting in the paths described above, the uncaptured chemicals (some of which will be valuable products, reactants, etc.) may be released or exit as fugitive emissions.

### Include All Emissions, Discharges, and Solid Wastes

In designing and analyzing processes, the releases to the environment include air emissions, liquid discharges, and solid waste. Air emissions are the best represented and analyzed of the releases. The most common of the reported air emissions are greenhouse gases

(GHGs), which are easily determined through calculations of energy type and amount used.

One can speculate that the search for data that is relevant to environmental impact categories suffers from the “streetlight effect”. This effect is told as a story of someone looking for their keys under a streetlight, and a passerby who stopped to help look for them finally asks where the keys were dropped. The answer is two blocks away, but the light is better here, with the analogy for emission data being the easy-to-see GHG data. Other forms of bias, like availability, can play a role as well, as that bias selects easily recalled items as being important. In general, analyses would benefit from not prescribing what important emissions, discharges, and solid waste flows are in advance of evaluations.

To improve process models that have been ignoring storage, transfer, vent, and fugitive emissions as well as liquid discharges and solid wastes, a number of methods are reviewed in the literature [10]. In addition, specific methods for estimating emissions for unit operations such as boilers, loaders / forklifts, and cooling towers are available [11].

### Apply Additional Impact Categories

The number of human health and environmental impact categories included in most studies are very limited. Global warming potential is often incorporated, but many other categories are dismissed through omission. I.e., no one has made a conscious decision to give a zero weighting to other categories; they have simply been omitted. Even in examples where more categories are used, the number is still relatively small [14]. A more extensive taxonomy of environmental impacts is available consisting of many tables of detailed midpoint effects (e.g., global warming potential (GWP) is a midpoint effect determined from emissions), endpoint effects (e.g., skin cancer, reduced lung function), and damage groupings (e.g., disability adjusted life years) [15]. Seldom, if ever, will a study use all of the available impact categories, but a review of those that are available can inform a more complete analysis, i.e., one in which more categories are included, or the text better describes the intent and caveats of the system analysis.

### Recognize the Disruptive Nature of Recycling to Design / Optimization

In chemical engineering process design, an early lesson is the effect a recycle loop can have on a design. Whereas in a straight-line process, or single-pass process, the highest yield is the best use of raw materials, when a recycle loop is added the optimum conversion can move towards the highest selectivity for use of raw materials. Recycle loops can return raw materials to the reactor system to react them more efficiently at higher selectivity. The optimum is often balanced by larger

equipment and recycle streams that can increase energy use and costs.

The advantages of recycling along a supply chain are somewhat different. For post-consumer use materials, a meta-analysis using 366 datasets for 14 materials was performed, and the mean GWPs for secondary production were better than the mean GWPs for primary production (i.e., with virgin materials) [16]. However, experience has shown that the quality of recycled materials is not equivalent to virgin for many reasons (e.g., for PET plastic bales other materials are present, and for PET items other materials are part of their composition) [17]. Thus, the secondary production process is different from the primary one. This represents a radical divergence from internal recycle loops in chemical processes, where higher selectivity accomplished with more recycling leads to less impurities.

For both process systems and supply chain systems, recycling must be evaluated. A combined analysis was done for producing waste-recycled feeds using styrene “tar” from the bottoms of a styrene-refining column [18]. There is no generalization to make regarding the desirability of recycling such streams, as the system depends both on the technology and the materials recycled and produced. A practitioner of process systems engineering would do well to design and analyze each system.

### Create a Robust Superstructure of Alternatives

In process systems engineering a common problem studied is the superstructure-based reactor synthesis. An early paper on the subject used a recycle reactor with heat exchange as the basis [19]. More recently, superstructure methods were applied to complex sets of reactions modeled with uncertain inputs and limiting reagents [20]. The process synthesis problem is aimed at achieving a conceptual design that identifies the operations to do. A similar larger-scale problem is challenging at the process level when new chemicals can be introduced into a process [5]. In each case the idea is to have a robust superstructure of equipment and interconnections that is flexible enough to capture designs when applying optimization. Further, at a larger supply chain scale the analysis of designs is challenging when new chemicals are introduced, as up- and down-stream processes and their resource use and environmental releases will be affected.

### Use Complete Sets of Objectives

In the real world, decisions are not simple because there is seldom a real-world decision that only has one objective. Our models of processes and systems can diverge strongly from this generalization, as assumptions are made to only consider a single dimension. In process design, one might optimize economics, flexibility,

controllability, safety, and environmental impacts. The GREENSCOPE tool of Ruiz-Mercado et al. has approximately 140 indicators in the four E's of Environment, Economics, Energy, and mass Efficiency, thus providing many possible objectives to optimize chemical processes [21].

There are various ways of handling multiple objectives. One method is to rearrange an objective as a constraint on acceptable solutions. Solving over a range of different values for the constraint will create a series of solutions. Another method to address multiple objectives is to normalize each objective (i.e., dividing by a maximum possible value is one normalization), weighting the multiple objectives with respect to each other, and finally adding the objectives together on a single scale. A method for applying these steps using marginal rates of substitution and total utility is described by Smith and Ruiz-Mercado [22]. Others may approach the multiple objectives through the simultaneous development of many Pareto solutions, although dimensionality issues require a method for spacing solutions among the various objectives [23].

## Realize Sustainability Tools are Available

In designing or analyzing process systems it may be that people are unaware of tools that are available. Examples of tools one can use to further sustainability include GREENSCOPE [21], release estimation [11], and LCA [4] tools. In addition, solvent replacement and toxicity prediction methods are available from the U.S. EPA's Office of Research and Development (ORD). The Program for Assisting the Replacement of Industrial Solvents (PARIS) allows one to quickly find solvent replacements (either individual solvents or mixtures) that are similar in physical and chemical properties to the original. The program also provides a calculation of potential environmental impacts in eight categories, from global warming potential, ozone depletion potential, acidification potential, and smog potential, to four categories of potential toxicity [24]. Additional toxicities and many other physical properties can be estimated with the Toxicity Estimation Software Tool (TEST), also available from EPA's ORD [25].

## DISCUSSION

Practitioners of process systems engineering are encouraged to employ the above list to improve their chemical processes and associated supply chains. To summarize the above, one should first create designs that are inherently safer, including other principles of green chemistry and engineering. Sustainable research and development breakthroughs advanced through sustainable chemistry and engineering improve system performance while reducing environmental burdens and

economic and social costs. In the absence of breakthroughs, one can only optimize systems to make them more sustainable.

Practitioners should also be as complete as possible in their identification of process chemicals used and released, be more complete in their definitions of supply chain processes, resources used, and releases, and apply additional environmental impact categories as appropriate.

Users are likely already familiar with recycling, superstructure optimizations, and creating sets of objectives. This effort simply advises to employ these in a manner that positively disrupts systems, defines the widest possible variety of system structures, and includes complete sets of objectives. This setting of objectives, as choices of objectives, constraints, and boundaries of what is included and excluded, is critically important in defining the scope of studies.

Finally, users should be aware that sustainability tools are available, which might have been outside of their awareness. In the end practitioners of process systems engineering will likely still limit their studies with incomplete methods for sustainability, but perhaps the listing here can be used as a checklist to address what could be considered and help refine some future work.

## DISCLAIMER

The views expressed in this article are those of the author and do not necessarily represent the policies or views of the U.S. Environmental Protection Agency.

## ACKNOWLEDGEMENTS

The authors have not reported funding for this work.

## REFERENCES

1. U.S. EPA (2006). *Life Cycle Assessment: Principles and Practice*, Scientific Applications International Corporation, Reston, VA. EPA/600/R-06/660. Office of Research and Development.
2. Ingwersen WW, Subramanian V. Guidance for product category rule development: process, outcome, and next steps. *Int. J. LCA* 19:532-537 (2014).
3. Faturay F, Vunnava VSG, Lenzen M, Singh S. Using a new USA multi-region input output (MRIO) model for assessing economic and energy impacts of wind energy expansion in USA. *Appl. Energy* 261: 114141 (2020).
4. Ingwersen WW, Li M, Young B, Vendries J, Birney C. USEEIO v2.0, The US Environmentally-Extended Input-Output model v2.0. *Sci. Data* 9:194 (2022).
5. Smith, RL. Conceptual Chemical Process Design for

- Sustainability, in *Sustainability in the Design, Synthesis and Analysis of Chemical Engineering Processes*, Gerardo Ruiz-Mercado and Heriberto Cabezas, eds. Butterworth-Heinemann (2016).
6. Cano-Ruiz JA, McRae GJ. Environmentally conscious chemical process design. *Annu. Rev. Energy* 23:499-536 (1998).
  7. Bakshi BR, Fiksel J. The quest for sustainability: Challenges for process systems engineering. *AIChE J.* 49(6):1350-1358 (2003).
  8. Gonzalez MA, Smith RL. A methodology for the evaluation of process sustainability. *Environmental Progress* 22(4):269-276 (2003).
  9. Parvatker AG, Eckelman, MJ. Comparative evaluation of chemical life cycle inventory generation methods and implications for life cycle assessment results. *ACS Sust Chem Eng.* 7:350-367 (2019).
  10. Smith RL, Ruiz-Mercado GJ, Meyer DE, Gonzalez MA, Abraham JP, Barrett WM, Randall PM. Coupling computer-aided process simulation and estimations of emissions and land use for rapid life cycle inventory modeling. *ACS Sust. Chem. Eng.* 5:3786-3794 (2017).
  11. Smith RL, Tan ECD, Ruiz-Mercado GJ. Applying environmental release inventories and indicators to the evaluation of chemical manufacturing processes in early stage development. *ACS Sust. Chem. Eng.* 7:10937-10950 (2019).
  12. Douglas JM. *Conceptual Design of Chemical Processes*. McGraw-Hill (1988); pp.121-
  13. Li S, Feliachi Y, Agbleze S, Ruiz-Mercado GJ, Smith RL, Meyer DE, Gonzalez MA, Lima FV. A process systems framework for rapid generation of life cycle inventories for pollution control and sustainability evaluation. *Clean Technol. Environ. Policy* 20:1543-1561 (2018).
  14. Smith RL, Ruiz-Mercado GJ, Gonzalez MA. Using GREENSCOPE indicators for sustainable computer-aided process evaluation and design. *Comput. Chem. Eng.* 81:272-277 (2015).
  15. Bare JC, Gloria TP. Environmental impact assessment taxonomy providing comprehensive coverage of midpoints, endpoints, damages, and areas of protection. *J. Clean. Prod.* 16:1021-1035 (2008).
  16. Brogaard LK, Damgaard A, Jensen MB, Barlaz M, Christensen TH. Evaluation of life cycle inventory data for recycling systems. *Resources Conservation and Recycling* 87:30-45 (2014).
  17. Smith RL, Takkellapati S, Riegerix RC. Recycling of plastics in the United States: Plastic material flows and polyethylene terephthalate (PET) recycling processes. *ACS Sust. Chem. Eng.* 10:2084-2096 (2022).
  18. Smith RL. Hierarchical design and evaluation of processes to generate waste-recycled feeds. *Ind. Eng. Chem. Res.* 43:2508-2515 (2004).
  19. Achenie LKE, Biegler LT. A superstructure based approach to chemical reactor network synthesis. *Comput. Chem. Eng.* 14(1):23-40 (1990).
  20. Ramapriya GM, Won W, Maravelias CT. A superstructure optimization approach for process synthesis under complex reaction networks. *Chem. Eng. Res. Des.* 137:589-608 (2018).
  21. Ruiz-Mercado GJ, Smith RL, Gonzalez MA. Sustainability indicators for chemical processes: I. Taxonomy. *Ind. Eng. Chem. Res.* 51:2309-2328 (2012).
  22. Smith RL, Ruiz-Mercado GJ. A method for decision making using sustainability indicators. *Clean Techn. Environ. Policy* 16:749-755 (2014).
  23. Kim K-J, Smith RL. Parallel multiobjective evolutionary algorithms for waste solvent recycling. *Ind. Eng. Chem. Res.* 43:2669-2679 (2004).
  24. Harten P, Martin T, Gonzalez M, Young D. The software tool to find greener replacement solvents, PARIS III. *Environ. Prog. & Sust. Energy* 39:e13331 (2020).
  25. Martin, TM. User's Guide for T.E.S.T. (Toxicity Estimation Software Tool). (2020)  
<https://www.epa.gov/sites/production/files/2016-05/documents/600r16058.pdf>

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Designing Reverse Electrodialysis Process for Salinity Gradient Power Generation via Disjunctive Programming

Carolina Tristán<sup>a</sup>, Marcos Fallanza<sup>b</sup>, Raquel Ibáñez<sup>b</sup>, Ignacio E. Grossmann<sup>c</sup> and David Bernal Neira<sup>a,d,e\*</sup>

<sup>a</sup> Purdue University, Davidson School of Chemical Engineering, West Lafayette, IN, USA

<sup>b</sup> University of Cantabria, Department of Chemical and Biomolecular Engineering, Santander, Spain

<sup>c</sup> Carnegie Mellon University, Department of Chemical Engineering, Pittsburgh, PA, USA

<sup>d</sup> Universities Space Research Association, Research Institute of Advanced Computer Science, Mountain View, CA, USA

<sup>e</sup> Quantum Artificial Intelligence Laboratory, NASA Ames Research Center, Moffett Field, CA, USA

\* Corresponding Author: [dbernaln@purdue.edu](mailto:dbernaln@purdue.edu).

## ABSTRACT

Reverse electrodialysis (RED) is a nascent renewable technology that generates clean, baseload electricity from salinity differences between two water streams, a renewable source known as salinity gradient energy (SGE). Full-scale RED progress calls for robust techno-economic and environmental assessments. Using generalized disjunctive programming (GDP) and life cycle assessment (LCA) principles, this work proposes cost-optimal and sustainable RED process designs involving different RED stack sizes and width-over-length ratios to guide the design and operation from the demonstration to full-scale phases. Results indicate that RED units will benefit from larger aspect ratios with a relative increase in net power of over 30% with 6 m<sup>2</sup> membrane size. Commercial RED unit sizes (0.25–3 m<sup>2</sup>) require larger aspect ratios to reach an equal relative increase in net power but exhibit higher power densities. The GDP model devises profitable RED process designs for all the assessed aspect ratios in a foreseeable scenario for full-scale deployment, that is, the energy recovery from desalination concentrates mixed with reclaimed wastewater effluents. A RED system with 3 m<sup>2</sup> RED units nine times wider than its length could earn a net present value of \$2M at a competitive levelized cost of electricity of \$111/MWh in the Spanish electricity market. On-site, RED-based electricity could abate roughly 7% of the greenhouse gas emissions from the desalination plant's energy supply, given the low emissions contribution of RED supply share. These findings demonstrate that optimization-based eco-technoeconomic assessments are a vital ally in making RED a full-scale reality.

**Keywords:** Process Design, Renewable and Sustainable Energy, Optimization, Pyomo, Modelling and Simulations, Life Cycle Analysis

## INTRODUCTION

Demonstrating and deploying clean renewable energy technologies must be a global priority in pursuing a net-zero emissions economy by mid-century [1]. Salinity gradient energy (SGE) technologies offer deep and sustained reductions in greenhouse gas (GHG) emissions to keep the 2050 goal within reach. These technologies recover the chemical energy released when high-salinity and low-salinity streams are reversibly mixed. Reverse electrodialysis (RED) is one of the most researched and advanced SGE technologies.

RED employs ion-exchange membranes (IEMs) to generate electricity from SGE directly. These IEMs allow ions of opposite charge but not water to pass through. A RED device is built by stacking a series of alternating cation (CEMs) and anion exchange membranes (AEMs) that separate salt solutions of different concentrations. Selective transport of ions through the IEMs creates an electric potential across the pairs of AEMs and CEMs that drive redox reactions at electrodes on either side of the membrane pile. The overall electric potential of the set of cell pairs and the electric current then power an external load that closes the circuit [2].

In 2014, REDstack BV hit a significant milestone when their demonstration plant was successfully put into operation on the Afsluitdijk in the Netherlands, meeting Technology Readiness Level (TRL) 7. With a total membrane area per stack of 250 m<sup>2</sup>, this plant is powered by a blend of salt water and freshwater, producing 50 kW at present [3,4]. The REAPower pilot plant in Trapani, Italy is not functioning now and is being used as a demonstration project. Using both natural saturated brine from a saltworks and brackish water from a shoreline well, a trio of RED stacks were able to generate a total power output of 330 W, with a combined membrane area of more than 400 m<sup>2</sup> [5]. The main barriers preventing RED technological readiness are the low power density of large-scale RED systems (0.38–2.7 W/m<sup>2</sup> total membrane area), fouling, and the high cost of commercial membranes (> \$100/m<sup>2</sup>) [6–8]. The development of high-performing membranes, electrode segmentation, and multi-staging are some of the approaches to enhance the power density and energy efficiency of RED.

The water sector opens new avenues to prove and advance full-scale RED. Desalination concentrates and treated wastewater effluents are abundant yet largely untapped waste streams from which RED can extract sustainable and clean electricity [9]. On-site RED electricity generation in desalination plants can also lessen the dependence on the water and energy-intensive grid mix and reduce the environmental burden and costs associated with brine treatment and disposal [10]. This, in turn, contributes to more sustainable and self-sufficient water supply systems. Besides, RED operation with desalination brines delivers higher power densities than river/seawater pairs, and the reject brine does not require further energy-intensive treatment as raw seawater.

Even so, the complex process configuration and operational decision space make it technically challenging to estimate the costs and performance of RED with conventional heuristics. In previous work, we developed a Generalized Disjunctive Programming (GDP) optimization model incorporating a RED stack predictive model to define the cost-optimal RED process design in different scenarios [11]. The solution for the GDP model provided the flowsheet design that maximizes the process net present value (NPV) for a given RED stack design.

The quantification of the RED process environmental loads is also a valuable input to devise environmentally sound design alternatives to RED technology. In this regard, we conducted a life cycle assessment (LCA) of the RED stack to define the environmental profile of RED and to estimate greenhouse gas (GHG) emissions reduction in desalination plants partly sourced with SGE [12].

Building on the LCA of the RED unit and the GDP optimization model of the RED process, this follow-up work explores how scaling up and the design of the RED units' compartments may affect the eco-technoeconomic performance of the optimal RED process flowsheets.

## METHODS

The performance metrics in the eco-techno-economic assessment (eTEA) are the net power output (NP), the net power density (NPD, i.e., net power per total membrane area), the NPV, the levelized cost of electricity (LCOE), and the global warming potential (GWP).

### Problem Statement

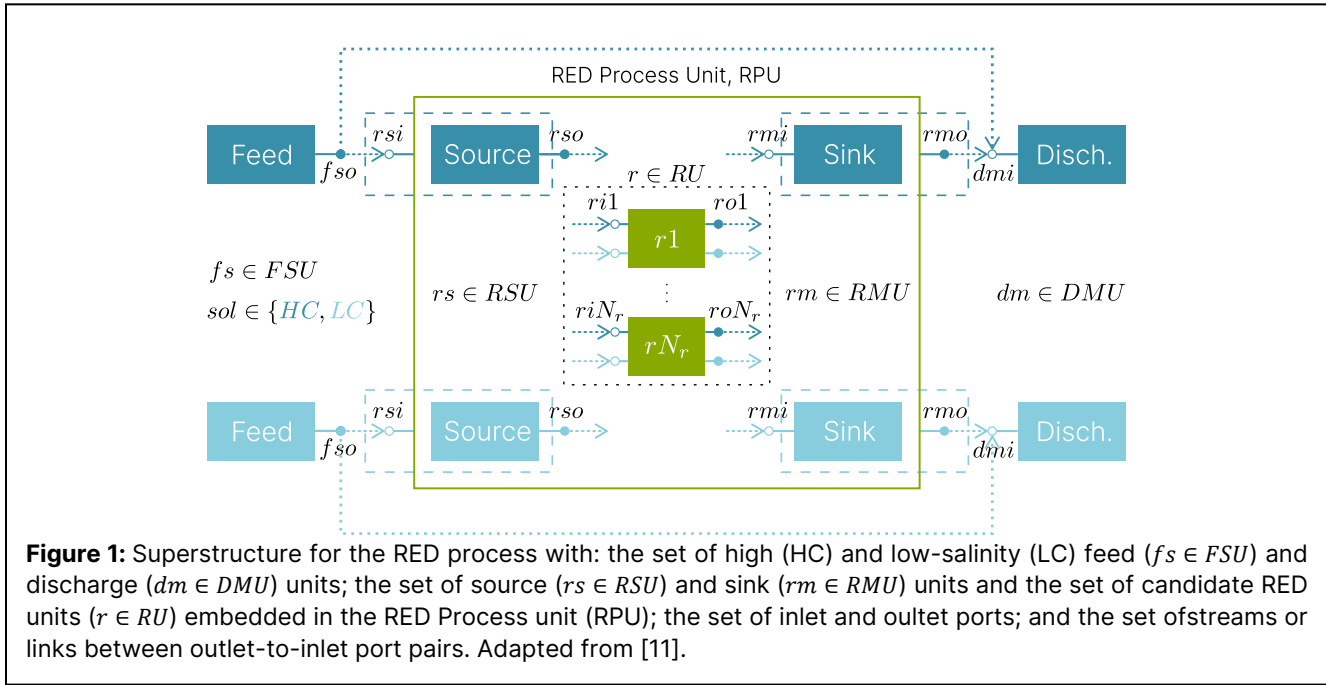
Given a set of identical candidate RED units  $r \in RU = \{r1, \dots, rN_r\}$ , the goal is to determine which ones are active, how they are hydraulically arranged, and their working conditions (*e.g.*, electric current, inlet flow rate, and molar concentration of the HC and LC streams) that yield the cost-optimal flowsheet design of the RED process for a given concentration, volume, and temperature of the high-salinity and low-salinity feed streams, and design parameters of the RED units.

The superstructure in Figure 1 incorporates all the alternative hydraulic topologies for the RED system with  $N_r$  conditional RED units. The superstructure and notation are fully described in previous work [11].

Table 1 summarizes the design parameters of the RED units spanning pilot to commercial scales and different compartment geometries. We set seven distinct sizes (active area in Table 1) and, for each size, nine different width-over-length ratios (aspect ratio, Table 1). Using the predictive model, we estimate the operational conditions that maximize the net power of the stand-alone RED unit for each size and aspect ratio. This sensitivity analysis provides guidelines for RED unit design and operation in all development stages.

We quantify the GWP of the RED units in Table 1 based on the previous LCA of pilot-scale RED units [12].

Later, to explore how the design of the RED stack compartments affects the techno-economic and environmental performance of the optimal RED process design, we solve the GDP model of the RED system with 3 m<sup>2</sup> size RED units varying their aspect ratio.



**Table 1:** Design parameters of the RED stack [11].

Parameter	Value
Number of cell pairs	1000
Aspect ratio, w/L	1–9
Active area	0.25, 1, 2, 3, 4, 5, 6 m <sup>2</sup>
<b>Spacers</b>	
Thickness	270 μm
Porosity	82.5%
<b>IEM: fumasep® CEM (FKS-50) / AEM (FAS-50)</b>	
Areal resistance	1.8 / 0.6 Ω/cm <sup>2</sup>
Permselectivity	0.93
Thickness	50 μm

## Optimization Model

The set of equations (1) defines the general form of the Generalized Disjunctive Programming (GDP) optimization model for the superstructure in Figure 1.

We code the GDP model using the algebraic modeling language Pyomo [13] and Pyomo.GDP [14], a dedicated Pyomo library extension for logic-based modeling and optimization.

$$\begin{aligned}
 \max NPV &= f(x) \\
 \text{s. t. } &g(x) \leq 0 \\
 &\begin{bmatrix} Y_r \\ h_r(x) \leq 0 \end{bmatrix} \vee \begin{bmatrix} \neg Y_r \\ B^r x = 0 \end{bmatrix} \quad \forall r \in RU \quad (1) \\
 &\Omega(Y_r) = True \\
 &x \in X \subseteq R^n \\
 &Y_r = \{True, False\} \quad \forall r \in RU
 \end{aligned}$$

In problem (1), the objective is to maximize the NPV of the RED process. The continuous variables  $x$  are the molar concentration and flow rate of the streams and the internal variables of the active RED units. The decision variables are the electric current, inlet concentration, and flow rate of the RED stacks.

The global constraints,  $g(x) \leq 0$ , describe specifications and physical relationships that must hold for any selection of alternatives in the superstructure, *e.g.*, mass balances of the feed, source, sink, and discharge units, and concentration and flowrate upper and lower bounds.

The  $N_r$  two-term disjunctions denote the discrete activation and deactivation of the  $N_r$  candidate RED units governed by the corresponding Boolean variables  $Y_r$  in each disjunct. When the unit exists ( $Y_r = True$ ), the active constraints  $h_r(x) \leq 0$  impose the RED unit discretized model equations (*e.g.*, mass and energy balances or other physicochemical phenomena within the RED unit), compute the capital and operating costs, and set bounds on the internal variables and the concentration and flow rate of the inlet and outlet streams; otherwise, ( $\neg Y_r$ ) the RED unit equations in the inactive disjunct are ignored, and  $B^r x = 0$  constraints set to zero a subset of the continuous variables and cost terms in the objective function.

The logical relationships ( $\Omega(Y_r) = True$ ) establish the logic conditions for selecting the candidate RED units.

To formulate the GDP problem, we assume:

1. Pure sodium chloride (NaCl) feed solutions, thus presuming ideal aqueous solution (*i.e.*, unity activity coefficients) and the absence of other species.
2. The ionic resistances of solutions and membranes are the unique internal energy loss.

3. Constant membranes' permselectivity and ionic resistance with concentration and temperature.
4. No water transport across membranes due to osmosis, so the streamwise volumetric flow rate in the RED channel is constant.
5. Salt diffusivities in the membrane phase are independent of concentration and temperature.
6. No fluid leakage or ionic shortcut currents in the RED stack's manifolds.
7. Co-current flow.
8. Isothermal and isobaric conditions.

Be aware that simplifying the RED stack model [11] leads to an increased net power output, causing the LCOE to be underestimated and the NPV to be overestimated.

The NPV of the RED process (2) accounts for discounted annual revenues from electricity sales and carbon pricing incentives and discounted operating costs (OPEX in \$/year) and capital expenses (CAPEX in \$). The OPEX and annualized CAPEX define the total annual cost (3), TAC, of the RED system. The CAPEX is annualized over the expected lifetime of the plant  $LT$  in years, using the capital recovery factor,  $CRF$ , given in (4) with a discount rate  $DR$ .

We assume the RED plant electricity is sold to the grid at the Spanish average price of electricity for non-house consumers,  $ep$  [15], and that the abated GHG emissions from the grid mix (Spanish emission factor,  $ef$ ) are subsidized at the average price,  $cp$ , in the European Union Emission Trading System (EU ETS) [16].

$$NPV = \frac{(ep + cp ef) TNP - TAC}{CRF} \quad (2)$$

$$TAC = CRF CAPEX + OPEX \quad (3)$$

$$CRF = \frac{DR}{1 - (1 + DR)^{-LT}} \quad (4)$$

$$TNP = \sum_{r \in RU} NP_r \quad (5)$$

We adapt our RED stack model [10] for a tractable yet rigorous solution. When the RED unit is active ( $Y_r = True$ ), the discretized model computes the net power output,  $NP_r$ , that is added to the RED system net power capacity, *i.e.*, total net power,  $TNP$  in kW (5). The net power output equals zero when the RED unit is absent ( $\neg Y_r$ ).

We apply a load factor,  $LF$ , to the annual full-capacity energy yield (kWh/year) of the RED plant to account for plant downtime due to membrane cleaning and system maintenance.

The capital investment involves the cost of RED stacks,  $\sum_{r \in RU} CC_{stack,r}$ , pumps,  $CC_{pump}$ , and civil and electrical infrastructure costs,  $CC_{civil}$ .

$$CAPEX = \sum_{r \in RU} CC_{stack,r} + CC_{pump} + CC_{civil} \quad (6)$$

The annual operating cost comprises the electricity cost from pumps,  $\sum_{r \in RU} OC_{pump,r}$ , the replacement cost of membranes,  $\sum_{r \in RU} OC_{IEMsrep,r}$ , and maintenance and labor costs (2% of CAPEX).

$$OPEX = \sum_{r \in RU} OC_{pump,r} + \sum_{r \in RU} OC_{IEMsrep,r} + 0.02 CAPEX \quad (7)$$

When the RED unit is active,  $CC_{stack,r}$  is added to CAPEX, and  $OC_{pump,r}$  and  $OC_{IEMsrep}$  to the OPEX; if not, these terms take zero values.

The objective function in (2) is maximized subject to constraints in the GDP detailed in [11]. The main financial parameters are reported in Table 2.

## Solution Strategy

We solve the GDP problem with the Global Logic-based Outer Approximation (GLOA) algorithm [17,18] implemented in the logic-based solver GDPopt version 20.2.28 built on Pyomo.GDP. The GLOA algorithm decomposes the solution to the GDP into a sequence of mixed-integer linear programming (MILP) problems and reduced nonlinear programming (NLP) subproblems.

We solve the MILP master problems with CPLEX and the NLP subproblems with BARON setting the time limit at 1 hour and 1% optimality gap on a machine running Windows 10 (x64) with 6 cores processor (Intel® Core™ i7-8700 CPU @3.2 GHz) and 16 GB of RAM. We use the MINLP and NLP solver versions from GAMS 34.1.0.

**Table 2:** Financial parameters of the RED process.

Parameter	Value
Plant lifetime, $LT$ [19]	30 years
Membrane lifetime [19]	10 years
Membrane price [20]	\$10/m <sup>2</sup>
Load factor, $LF$ [19]	90%
Discount rate, $DR$ [19]	5%
Spanish GWP, $ef$	0.374 kg CO <sub>2</sub> -eq/kWh
Carbon price, $cp$	\$27.8/t CO <sub>2</sub> -eq
Electricity price, $ep$	\$197/MWh

Spanish 2019-average price of electricity for non-house consumers. Band IB: annual consumption between 20 MWh and 500 MWh, excluding taxes and levies.

## RESULTS AND DISCUSSION

In all the assessments the RED units retrieve energy from the concentrate effluent of the Maspalomas II sea-water reverse osmosis desalination plant in Gran Canaria (Canary Islands, Spain) [21–23]. Maspalomas II plant rejects 17,602 m<sup>3</sup>/day (733 m<sup>3</sup>/h) of brine (1.67 M NaCl, 20°C) and consumes 3.77 kWh per cubic meter of desalted water. The low-salinity feedwater (20mM NaCl) is obtained from nearby wastewater treatment plants (*e.g.*, el Tablero, las Burras) [24], so the same LC and HC feed volume is available for SGE conversion.

### Stand-alone RED unit

The discretized RED unit NLP model involves 107–1187 variables and 107–1232 constraints is solved in 237 s up to an hour CPU time with BARON depending on the number of finite elements (from 3 up to 48 finite elements) that is set to keep the same axial discretization accuracy between the different sizes and aspect ratios.

Pilot-scale research often employs stack designs based on their counterparts in desalination, *i.e.*, electro-dialysis [6,25]. These modules feature greater length and smaller width, as the objective is to dilute the feed to comply with a given quality standard [26]. Alternatively, square geometries are usually adopted [5,6].

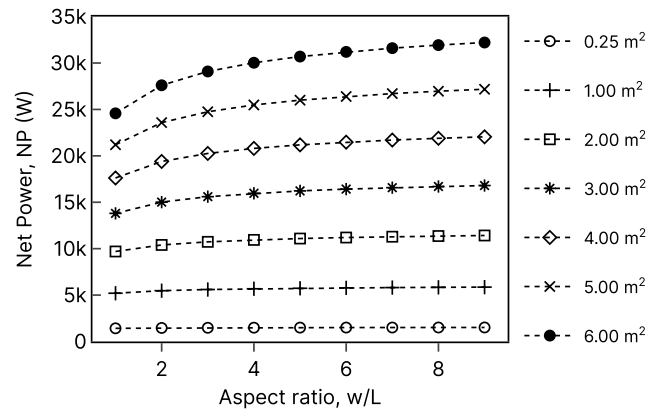
The premise of this study is that using modules that are wider than longer (*i.e.*, aspect ratio greater than one) while housing the same membrane area would allow more powerful systems and the treatment of larger feed volumes with fewer units. This would result in more compact and cost-effective systems.

The RED unit power generation sensitivity to aspect ratio increases with size (Figure 2). The smaller units require larger aspect ratios to reach an equal relative increase in NP. For instance, 1 m<sup>2</sup> units reach a 13% relative increase in NP when the aspect ratio moves from 1:1 to 9:1, while a RED unit twice its size requires a 4:1 ratio to reach the same increase. Increasing the width of the largest 6 m<sup>2</sup> stack nine times would lead to a 31% boost in net power generation, from 24.6 kW to 32.2 kW. In comparison, the smaller unit with an area of 0.25 m<sup>2</sup> and the same shape only generates 6% more net power than the square one.

Findings also indicate that while the largest RED unit delivers more net power than its pilot counterparts (Figure 2), it exhibits a lower power density (Figure 3). This may raise the cost per kWh of the RED unit despite the improvement in net power output.

Depending on the stack's geometry and size, and ultimately its ability to sustain the salinity gradient along the flow path, the optimal solution tunes the linear flow velocity in the HC and LC channels, the inlet concentration of the LC feedstream, and the electric current of the RED unit to maximize the net power output.

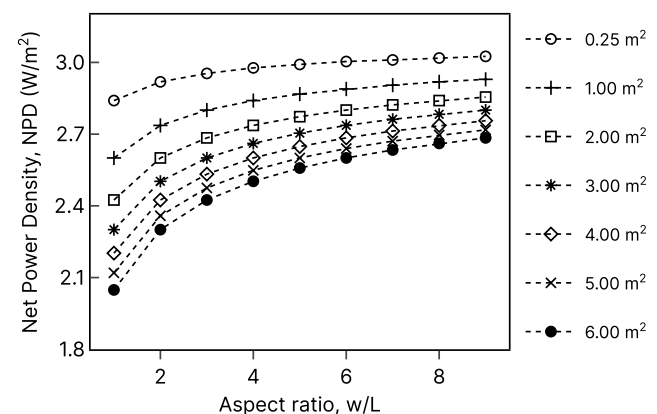
Shorter flow paths—realized by reducing the size for a given aspect ratio or increasing the aspect ratio for a given size—keep the inlet concentration gradient, *i.e.*, the highest driving force, along the channels. If the length is enlarged, there is enough time for ions to flow from the HC to the LC side, fading the concentration gradient.



**Figure 2:** Optimal net power output of the stand-alone RED unit with different sizes (active area) and width-over-length ratios.

Recovering energy from high salinity gradients gives rise to two opposing effects on power generation. On one hand, the electric potential of the cell pairs increases, resulting in a higher gross power generation. On the other hand, the low conductivity of the LC channel increases the internal losses, leading to a decrease in gross power.

Given that the electric current drives the migration of ions across membranes from high-salinity to low-salinity compartments, the optimal electric current should decrease with longer RED units to extend the concentration gradient along the flow path. The opposite is true for shorter RED units, where the optimal solution sets a higher electric current, balancing the increase in the electric potential and internal resistance loss that arises from a higher concentration gradient, as shown in Table 3.



**Figure 3:** Net power density of the stand-alone RED unit with different sizes (active area) and width-over-length ratios under optimal net power conditions.



**Table 3:** Optimal operation variables of the 3 m<sup>2</sup> stand-alone RED unit.

		Aspect Ratio, w/L			
		1:1	3:1	6:1	9:1
Current density		7.64 mA/cm <sup>2</sup>	8.08 mA/cm <sup>2</sup>	8.26 mA/cm <sup>2</sup>	8.35 mA/cm <sup>2</sup>
Potential per cp		68.76 mV	69.94 mV	70.43 mV	70.65 mV
Linear velocity	HC	2.9 cm/s	2.5 cm/s	2.2 cm/s	2.1 cm/s
	LC	4.8 cm/s	3.9 cm/s	3.4 cm/s	3.1 cm/s
Flow rate	HC	40.66 m <sup>3</sup> /h	59.88 m <sup>3</sup> /h	76.08 m <sup>3</sup> /h	87.40 m <sup>3</sup> /h
	LC	66.53 m <sup>3</sup> /h	94.43 m <sup>3</sup> /h	116.15 m <sup>3</sup> /h	130.72 m <sup>3</sup> /h
LC Concentration		29 mM	37 mM	42 mM	44 mM
Net power, NP		13.8 kW	15.6 kW	16.4 kW	16.8 kW
Net Power Density, NPD		2.30 W/m <sup>2</sup>	2.60 W/m <sup>2</sup>	2.74 W/m <sup>2</sup>	2.80 W/m <sup>2</sup>

**Table 4:** NPV-optimal solution of the RED process with 3 m<sup>2</sup> RED units and different width-over-length ratios.

		Aspect Ratio, w/L			
		1:1	3:1	6:1	9:1
Active RED units		22	21	22	21
Net Present Value, NPV		\$1.8M	\$2.3M	\$2.5M	\$2.6M
Net Power Capacity, TNP		267 kW	274 kW	286 kW	282 kW
Net Power Density, NPD		2.02 W/m <sup>2</sup>	2.17 W/m <sup>2</sup>	2.16 W/m <sup>2</sup>	2.24 W/m <sup>2</sup>
Levelized Cost of Electricity, LCOE		\$132/MWh	\$116/MWh	\$114/MWh	\$111/MWh
Global Warming Potential, GWP		2.93	2.73	2.74	2.65
		kg CO <sub>2</sub> /MWh	kg CO <sub>2</sub> /MWh	kg CO <sub>2</sub> /MWh	kg CO <sub>2</sub> /MWh
# variables		24,249	15,849	12,249	10,449
# constraints		25,093	16,343	12,593	10,718

As the optimization model predicts, shifting to a wider-than-long stack design allows preserving the salinity gradient along the flow path with a lower linear velocity in the compartments. This, in turn, reduces hydraulic losses and pumping power to overcome head losses. Table 3 highlights the impact of this effect on the module with a membrane area of 3 m<sup>2</sup>. As the module length decreases, a downward trend in the optimal linear velocity is observed in both the HC and the LC. The flow rate rises despite the lower linear velocity due to the larger cross-sectional area. Concurrently, the concentration of the LC feed increases to offset the rise in the internal resistance.

### Optimal RED process design

We assume the superstructure has 25 identical candidate RED units with an active membrane area of 3 m<sup>2</sup>. For each aspect ratio (*i.e.*, 1, 3, 6, and 9) and the given financial parameters, the optimal solution provides the topology and decision variables that balance electricity production and capital and operating expenses. Discrete decisions involve the working RED units and the active water streams. Continuous variables are the inlet streams flow rate and concentration and active RED stacks electric current.

The GDP model finds profitable RED process designs for all the assessed RED units' aspect ratios (Table 4). But wider-than-long RED units earn more profits than the square peers with almost the same number of active RED units. As anticipated in the former section, the shorter RED stacks exhibit a higher power density. If the total membrane area and HC and LC feed volumes are the same, the RED process using shorter units can produce more power. In the NPV-optimal solution, to accommodate more wider-than-long RED units, the optimal solution makes them operate with flow rates below the optimal ones in Table 3. Such reduced inlet flowrate declines the RED units' net power density. Nevertheless, the increase in electricity production revenues considerably outstrips the increase in capital and operating cost of the wider-than-long RED units with lower power density.

By incorporating RED-based electricity, the grid mix share of the desalination plant supply could be decreased by as much as 7%, thereby reducing GHG emissions. This results from RED's relative emissions contribution to the energy supply being perceptibly slighter, at 2.6–2.9 kg CO<sub>2</sub>-eq/MWh compared to the Spanish grid mix, which emits 374 kg CO<sub>2</sub>-eq/MWh.

Overall, the GDP model defines cost-effective and sustainable RED process designs that improve the environmental profile and resource circularity of energy-intensive desalination and wastewater treatment plants; however, the nonconvexities leads to GDP problems that takes hours to solve with conventional global solvers (Table 4). This may be particularly true in full-scale RED systems with large-scale RED units.

## CONCLUSIONS

This work provides environmentally sustainable and cost-effective RED process designs exploring the RED units' different sizes and aspect ratios based on mathematical programming and the LCA framework. As a case study, we define energy recovery from mixing a real desalination plant's brine with reclaimed wastewater treatment plant effluents, a promising scenario for full-scale RED implementation.

The technical assessment of the size and different width-over-length ratios gives design and operation guidelines to derive compact systems that treat larger feed volumes with fewer yet powerful RED units. The assessment can assist in identifying the best aspect ratio for each module size.

Regarding the NPV-optimal RED process design with 3 m<sup>2</sup> RED units, the 9:1 width-over-length ratio yields the highest profit, \$2M, with an LCOE of \$111/MWh below the Spanish electricity market price (\$197/MWh) and a net power capacity of 282 kW from 22 RED units and virtually no added emissions to desalination plant's energy supply. As a result, RED-based electricity can abate around 7% of desalination plant's GHG from the grid mix supply at a competitive cost.

Overall, these results indicate that fine-tuning the aspect ratio is an effective way to advance in the development and commercial deployment of RED technology and prove that optimization-based eTEA is a robust tool to assist all development stages of emerging technologies such as RED electricity production.

Nonconvexities in the mixers and the RED unit model led to multiple optimal local solutions, therefore requiring computationally demanding global optimization techniques to solve to global optimality. This is particularly true, in large-scale RED systems where the model size significantly grows. A natural progression of this work is to reformulate the nonlinear equations into quadratic or linear approximations to exploit the bilinear nature of the GDP problem that solvers like Gurobi may effectively solve.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial support from projects TED2021-129874B-I00 and PDC2021-120786-I00 through European Union NextGenerationEU/PRTR and MCIN/AEI/10.13039/501100011033.

Carolina Tristán acknowledges the financial support from the research fellowship PRE2018-086454 funded by the Spanish Ministry of Science and Innovation (MCIN/AEI/ 10.13039/501100011033) and "ESF Investing in your future".

David Bernal was supported by the NASA Academic Mission Services, Contract No. NNA16BD14C.

David Bernal and Carolina Tristán acknowledge the support of the startup grant of the Davidson School of Chemical Engineering at Purdue University.

## REFERENCES

1. International Energy Agency, Net Zero by 2050: A Roadmap for the Global Energy Sector, 2021.
2. Pattle R.E., Production of Electric Power by mixing Fresh and Salt Water in the Hydroelectric Pile, *Nature*. 174:660–660 (1954)
3. Post J.W., Goeting C.H., Valk J., Goinga S., Veerman J., Hamelers H.V.M., Hack P.J.F.M., Towards implementation of reverse electrodialysis for power generation from salinity gradients, *Desalin. Water Treat.* 16: 182–193 (2010)
4. IEA-OES, Annual Report: An Overview of Ocean Energy Activities in 2022, (2023).
5. Tedesco M., Cipollina A., Tamburini A., Micale G., Towards 1 kW power production in a reverse electrodialysis pilot plant with saline waters and concentrated brines, *J. Memb. Sci.* 522:226–236 (2017)
6. Tufa R.A., Pawlowski S., Veerman J., Bouzek K., Fontananova E., di Profio G., Velizarov S., Goulão Crespo J., Nijmeijer K., Curcio E., Progress and prospects in reverse electrodialysis for salinity gradient energy conversion and storage, *Appl. Energy* 225:290–331 (2018)
7. Chae S., Kim H., Gi Hong J., Jang J., Higa M., Pishnamazi M., Choi J.Y., Chandula Walgama R., Bae C., Kim I.S., Park J.S., Clean power generation from salinity gradient using reverse electrodialysis technologies: Recent advances, bottlenecks, and future direction, *Chem. Eng. J.* 452:139482 (2023)
8. Nazif A., Karkhanechi H., Saljoughi E., Mousavi S.M., Matsuyama H., Recent progress in membrane development, affecting parameters, and applications of reverse electrodialysis: A review, *J. Water Process Eng.* 47:102706 (2022)

9. Rani A., Snyder S.W., Kim H., Lei Z., Pan S.Y., Pathways to a net-zero-carbon water sector through energy-extracting wastewater technologies, *Npj Clean Water* 5:49 (2022)
10. Tristán C., Fallanza M., Ibáñez R., Ortiz I., Recovery of salinity gradient energy in desalination plants by reverse electrodialysis, *Desalination* 496:114699 (2020)
11. Tristán C., Fallanza M., Ibáñez R., Ortiz I., Grossmann I.E., A generalized disjunctive programming model for the optimal design of reverse electrodialysis process for salinity gradient-based power generation, *Comput. Chem. Eng.* 174:108196 (2023)
12. Tristán C., Rumayor M., Dominguez-Ramos A., Fallanza M., Ibáñez R., Ortiz I., Life cycle assessment of salinity gradient energy recovery by reverse electrodialysis in a seawater reverse osmosis desalination plant, *Sustain. Energy Fuels*. 4:4273–4284 (2020)
13. Hart W.E., Laird C.D., Watson J.-P., Woodruff D.L., Hackebeil G.A., Nicholson B.L., Siirola J.D., Pyomo — Optimization Modeling in Python, Second Edition, Springer International Publishing, Cham, (2017)
14. Chen Q., Johnson E.S., Bernal D.E., Valentin R., Kale S., Bates J., Siirola J.D., Grossmann I.E., Pyomo.GDP: an ecosystem for logic-based modeling and optimization development, *Optim. Eng.* 23:607–642 (2022)
15. EUROSTAT, Electricity prices for non-household consumers - bi-annual data (from 2007 onwards) [NRG\_PC\_205], (n.d.).  
<https://ec.europa.eu/eurostat/databrowser/bookmark/65f83096-1534-4ddc-a561-44764c07601c?lang=en>
16. ICAP, Allowance Price Explorer, (2022).  
<https://icapcarbonaction.com/en/ets-prices>
17. Lee S., Grossmann I.E., A global optimization algorithm for nonconvex generalized disjunctive programming and applications to process systems, *Comput. Chem. Eng.* 25:1675–1697 (2001)
18. Chen Q., Johnson E.S., Siirola J.D., Grossmann I.E., Pyomo.GDP: Disjunctive Models in Python, *Comput. Aided Chem. Eng.* 44: 889–894 (2018)
19. Giacalone F., Papapetrou M., Kosmadakis G., Tamburini A., Micale G., Cipollina A., Application of reverse electrodialysis to site-specific types of saline solutions: A techno-economic assessment, *Energy*. 181:532–547 (2019)
20. Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR) programs, Low-Cost Manufacturing of High-Performance Ion Exchange Membranes for Electrodialysis using Initiated Chemical Vapor Deposition, (n.d.).  
<https://www.sbir.gov/node/2282063>
21. Meyer-Steele S., von Gottberg A., Talavera J.L., New Sea Water Reverse Osmosis Plant for the Caribbean "Energy Recovery, Brine Recovery & Cost Reduction" Ionics Technical Paper (2001)
22. Portillo E., de la Rosa M.R., Louzara G., Quesada J., Ruiz J.M., Mendoza H., Dispersion of desalination plant brine discharge under varied hydrodynamic conditions in the south of Gran Canaria, *Desalin. Water Treat.* 52:164–177 (2014)
23. Sadhwani Alonso J.J., Melián-Martel N., Environmental Regulations—Inland and Coastal Desalination Case Studies. In: Sustainable Desalination Handbook: Plant Selection, Design and Implementation, Ed: V.G. Gude, Butterworth-Heinemann (2018)
24. Pérez Talavera J., Quesada Ruiz J., Identification of the mixing processes in brine discharges carried out in Barranco del Toro Beach, south of Gran Canaria (Canary Islands), *Desalination*. 139:277–286 (2001)
25. Yasukawa M., Mehdizadeh S., Sakurada T., Abo T., Kuno M., Higa M., Power generation performance of a bench-scale reverse electrodialysis stack using wastewater discharged from sewage treatment and seawater reverse osmosis, *Desalination*. 491:114449 (2020)
26. Gurreri L., Tamburini A., Cipollina A., Micale G., Electrodialysis applications in wastewater treatment for environmental protection and resources recovery: A systematic review on progress and perspectives, *Membranes (Basel)*. 10:1–93 (2020)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Optimal Design of a Biogas-based Renewable Power Production System

Vikram Uday<sup>a</sup> and Sujit Jogwar<sup>a\*</sup>

<sup>a</sup> Indian Institute of Technology Bombay, Department of Chemical Engineering, Mumbai, Maharashtra, India

\* Corresponding Author: [jogwar@iitb.ac.in](mailto:jogwar@iitb.ac.in).

---

## ABSTRACT

This paper presents optimal design for an energy-integrated biogas-fuel cell system for renewable electricity generation. The integrated process consists of two steps. The first step generates hydrogen from biogas via methane steam reforming (SMR), whereas the second step electrochemically converts this hydrogen into electricity using a solid oxide fuel cell (SOFC). These two steps are coupled via material and energy integration. Specifically, various design alternatives like anode and/or cathode gas recycling, biogas upgradation by CO<sub>2</sub> removal, external versus direct internal reforming, and auxiliary power production through steam and/or micro gas turbine are explored to improve the overall efficiency and total annualized cost of the system. Specifically, a flowsheet superstructure is developed by incorporating all the available design alternatives. An optimal flowsheet with minimum total annualized cost is extracted from this superstructure using formal optimization techniques to meet the desired power target. Heat exchanger network superstructure is used to incorporate energy integration effectively. The proposed flowsheet and the corresponding optimal operating conditions are explained by analyzing the trade-offs associated with the corresponding design variables in terms of power production, capital expenditure, and utility consumption. For a power target of 300 kW, the proposed optimal energy-integrated process has a total annualized cost of \$608,955/y with a net electrical efficiency of 67.1% and corresponds to electricity cost of \$0.23/kWh.

**Keywords:** Process design, Optimization, Heat integration, Renewable electricity

## INTRODUCTION

Growing push towards sustainable practices has led to an increase in research activity in the area of power production from clean and renewable energy sources. Biogas, generated from the anaerobic digestion of organic matter, holds significant potential as a promising source for renewable power generation. Traditionally, biogas has been utilized to generate power through the use of reciprocating engines, microturbines, gas turbines, and steam turbines, resulting in low electrical efficiency. An efficient way to enhance the overall electrical efficiency of biogas is to integrate it with hydrogen fuel cells [1].

Integration of biogas with a fuel cell is, in principle, a two-step process. The first step generates hydrogen from biogas via a reforming reaction, like steam methane reforming (SMR), autothermal reforming, partial oxidation or dry reforming [2]. The second step electrochemically

converts this hydrogen into electricity using a fuel cell like solid oxide fuel cell (SOFC), proton exchange membrane fuel cell or molten carbonate fuel cell. These two steps can be coupled via material and energy integration. Specifically, various processing options like anode and/or cathode gas recycling, biogas upgradation by CO<sub>2</sub> removal, external versus direct internal reforming, use of steam turbine and/or micro gas turbine can be incorporated to improve the overall efficiency of the system.

Several studies have reported flowsheets for such integrated process. Piroonlerkgul et al. [3] proposed four designs differing on the basis of reforming agent (steam, air, or both) for external reformer and auxiliary power generation via steam turbine. Farhad et al. [4] proposed three designs involving external reformers with steam methane reforming or partial oxidation and explored the possibility of anode gas recycling. Trendewiz and Braun [5] presented designs for small, medium, and large-scale

integrated biogas-SOFC systems with anode gas recycling, external pre-reforming and direct internal reforming. They subsequently performed techno-economic analysis to estimate minimum electricity cost. Siefert and Lister [6] proposed a pressurized system with anode gas recycling and a micro gas turbine to generate auxiliary power from combustion of unspent fuel. Lastly, Baldineili et al. [7] proposed a design with biogas upgradation using membrane separator and proposed the use of direct internal reforming. It can be noted that most of these studies focus on only few select options for performance comparison, use parametric analysis to obtain final design instead of formal optimization and incorporate energy integration in an ad hoc manner. Thus, there is a need for a systematic approach for design and optimization of such highly integrated systems.

Motivated by this, the objective of this work is to develop an optimal flowsheet for power production via integrated SMR-SOFC system with biogas feed. A flowsheet superstructure is proposed based on the available design alternatives. Using formal optimization techniques, the optimal flowsheet is extracted from this superstructure by selecting the best combination of the processing alternatives to meet the set power target. The proposed flowsheet as well as the corresponding optimal operating conditions are explained by analyzing the trade-offs associated with the corresponding design variables in terms of power production, capital expenditure and utility consumption.

The rest of the paper is organized as follows. The next section describes the proposed flowsheet superstructure. The following section presents the optimal design framework with special emphasis on the trade-offs associated with key decision variables. The solution of the optimization problem and the corresponding results are discussed in the subsequent section.

## PROCESS SUPERSTRUCTURE

Superstructure optimization is a systematic technique to evaluate the different process alternatives. The first step involves generating a process flowsheet superstructure incorporating all the potential design alternatives. In the second step, this superstructure is converted into an optimization problem by incorporating material and energy balance equations along with operating and capital cost correlations. In the final step, this optimization problem is solved to obtain the optimal flowsheet along with the corresponding performance indices.

In this work, the various design alternatives proposed in literature (such as anode gas recycle, biogas upgradation, pre-reformer and direct internal reforming, steam and gas turbine) are combined together to generate the flowsheet superstructure. Previously unexplored options like cathode gas recycle or micro gas turbine on

membrane separator retentate are also incorporated. Lastly, a superstructure of heat exchanger networks is included to ensure efficient energy integration. The resulting superstructure is depicted in Figure 1.

The feed biogas at ambient conditions is supplied to the process via a blower or a compressor (C1) based on the system's operating pressure. Biogas production/cleanup is not explicitly considered in this work. Instead, cost of cleaned biogas is considered as feed cost as shown in Eq. (3). There are two potential routes for this biogas. In the anode gas recycle route, the feed biogas is mixed with a part of the spent anode gas which is rich in steam and  $H_2$  and proposes advantages in operating cost due to reduced consumption of fresh feed and steam. However, there is a penalty in terms of capital cost of reformer and fuel cell due to large flow rates and dilute conditions. The biogas upgradation route considers removal of  $CO_2$  via membrane separation. In this route, the biogas is pressurized using compressors C2 and C3.  $CO_2$  is removed as a low pressure permeate. The high-pressure retentate is expanded over a micro gas turbine to produce auxiliary power. The output of both these routes is heated using internal or external heating sources and subsequently fed to the external reformer. The  $H_2$ -rich exhaust of this reformer is fed to the anode section of the SOFC. If required, heating is provided by internal sources or hot utility. The SOFC anode is also capable of performing internal reforming to generate  $H_2$ . The relative extent of external and internal reforming can be manipulated by adjusting reformer operating conditions and the bypass provided over the reformer. The air required by the SOFC is sent through a blower/compressor (C4). Similar to the anode gas, it can be heated to the required temperature through internal or external sources. A part of the hot cathode effluent gas, which is rich in oxygen, is recycled back to reduce air demand as well as utility consumption. Similarly, in the case of anode gas recycle route, a part of the hot anode effluent gas is recycled back to the reformer. The rest of the anode and cathode spent gas (after taking out the anode and cathode recycle) is sent to the combustor. The combustor converts the chemical energy available with the spent fuel (unconverted hydrogen) into thermal energy which can be used for heating or auxiliary power generation. A part of the hot exhaust gas from the combustor is used to generate high-pressure steam via PPX10. This steam subsequently drives a steam turbine (T1) to generate auxiliary power. The rest of the hot gas is used for internal energy integration. It can be used to supply heat to the external reformer (PPX2) or meet high-temperature heating demand of the anode (PPX1) and cathode (PPX3) or reformer (PPX4 and PPX5) feed. The steam turbine exhaust can also be used to meet low temperature demands of reformer feed (PPX7), cathode (PPX6), reformer steam (PPX8) and membrane separator retentate



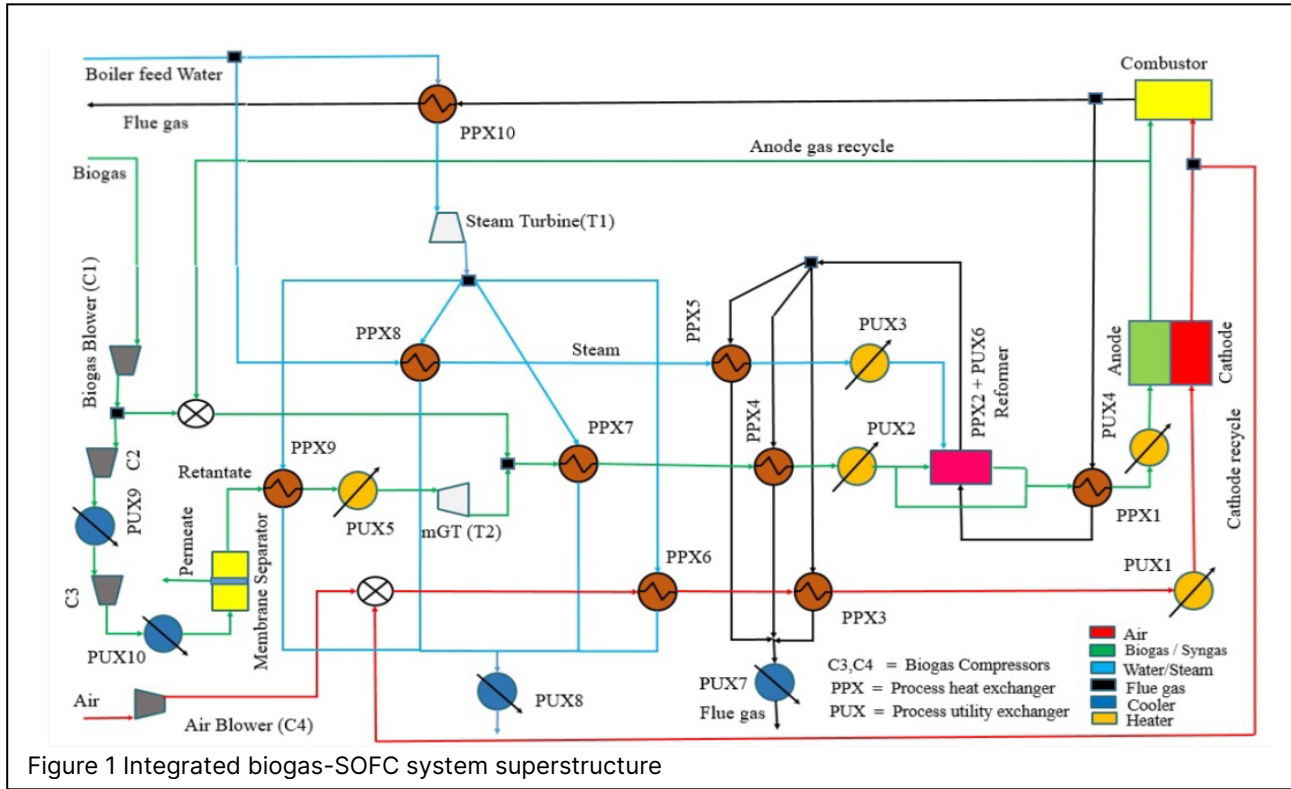


Figure 1 Integrated biogas-SOFC system superstructure

(PPX9). Process utility heat exchangers (PUX1 through PUX10) are placed to meet the rest of the heating and cooling demands.

It can be noted that the SOFC, along with steam (T1) and micro gas (T2) turbine produce electrical power. A part of this power is consumed internally to drive compressors C1, C2 and C3. The performance of the optimal integrated process is quantified by net electrical efficiency ( $\eta_{net-elect}$ ). It represents the conversion efficiency of the process and is defined as the ratio of the net power output to the total energy fed to the system. It can be computed using Eq. (1).

$$\eta_{net} = \frac{P_{SOFC} + \sum_j W_{turb,j} - \sum_j W_{comp,j}}{LHV_{BG} + \sum_j Q_{H,j}} \quad (1)$$

where  $P_{SOFC}$  represents power generated by the SOFC,  $W_{turb,j}$  and  $W_{comp,j}$  represent power generation and consumption by turbine and compressor, respectively.  $LHV_{BG}$  represents the heating value of the biogas feed and  $Q_{H,j}$  represents the hot utility consumption. Along similar lines, the performance of the SOFC is quantified by its conversion efficiency and is defined as the ratio of its power output to thermal energy input. The corresponding SOFC efficiency ( $\eta_{SOFC}$ ) is computed using Eq. (2).

$$\eta_{SOFC} = \frac{P_{SOFC}}{LHV_{BG}} \quad (2)$$

## OPTIMIZATION PROBLEM FORMULATION

The superstructure optimization problem consists of three components; the objective function, the decision variables and the process constraints. This section provides details about each of these components.

### Objective function

In order to obtain an economically efficient design, it is important to consider capital as well as operating cost. To this end, minimization of total annualized cost (TAC) is considered as an objective function. The capital cost for each equipment is computed using cost correlations given in Table 1. In the case of SOFC, the capital cost consists of cost of electrode active surface area ( $A_{SOFC}$ ), enclosure cost, converter cost and stack replacement cost. The cost of membrane separator is estimated based on required membrane area ( $A_{membrane}$ ). This area is computed through material balance and considering membrane permeance of 100 GPU.

The installed capital cost is computed by multiplying the above capital cost by the corresponding installation factors. It is subsequently scaled to 2023 cost using cost index (CECPI). Finally, the annualized capital cost is obtained by dividing this cost by equipment life. For this study, expected life of 10y is considered for all equipment, except membrane separator whose expected life is taken as 5y.

Operating cost consists of biogas feed cost along

with hot and cold utility as all the electrical power requirements are met internally. Thus the total annualized cost is computed using Eq. (3).

$$TAC = C_{BG}n_{BG} + C_{HU/CU} \sum_k Q_{H,k/C,k} + \sum_k \frac{E_k}{L_k} \quad (3)$$

where  $C_{BG}$  is the cost of feed biogas,  $C_{HU/CU}$  is the cost of hot or cold utility,  $E_k$  is the installed equipment cost and  $L_k$  is the expected life of that equipment.

**Table 1.** Cost correlations and parameters

Equipment/Feed/Utility	Cost correlation/value
SOFC Active surface area	\$ 1700 × $A_{SOFC}$ (m <sup>2</sup> ) [6]
SOFC Enclosure	\$ 80 × ( $P_{rSOFC}$ (atm)) <sup>0.33</sup> [6]
SOFC stack replacement	\$ 350 × $P_{SOFC}$ (kW) [6]
DC/AC convertor	\$ 70 × $P_{SOFC}$ (kW) [6]
Air/Biogas Blower	\$ 16,216 × ( $Q$ (m <sup>3</sup> )) <sup>0.56</sup> [8]
Steam turbine	\$ 29,410 × ( $P$ (kW)) <sup>0.49</sup> [8]
Micro gas turbine	\$ 200 × ( $P$ (kW)) [6]
Combustor	\$ 266.58 × ( $Q$ (kW)) <sup>0.871</sup> [8]
Utility heater	\$ 173.6 × ( $Q$ (kW)) <sup>0.765</sup> [8]
Process heat exchanger	\$ 147.5 × $A$ (m <sup>2</sup> ) [6]
Membrane Separator	\$ 150 × $A_{membrane}$ (m <sup>2</sup> ) [6]
Reformer	\$204,000 ( $m_{H_2}/1125$ ) <sup>0.7</sup> [5]
Biogas feed ( $C_{BG}$ )	\$18/kmol [9]
Hot utility ( $C_{HU}$ )	\$ 3.45 × 10 <sup>-5</sup> /kJ [9]
Cold utility ( $C_{CU}$ )	\$ 2.00 × 10 <sup>-6</sup> /kJ [9]

## Decision variables

The key decision variables and the corresponding trade-offs are given below.

1. SOFC operating pressure ( $P_{rSOFC}$ ): High SOFC pressure favors high open circuit voltage and SOFC efficiency. However, it increases capital cost, air compressor power consumption and lowers power production via micro gas turbine.
2. SOFC operating temperature ( $T_{SOFC}$ ): High SOFC temperature reduces voltage losses in SOFC, requires less airflow to carry exothermic heat and exit gases have high energy integration potential. However, it reduces open circuit voltage and increases hot utility consumption.
3. SOFC Fuel utilization factor ( $U_f$ ): It represents the fractional conversion of hydrogen in the SOFC. A high value of  $U_f$  increases power generation from the SOFC. However, it leads to increased current density and voltage losses as well as an increase in airflow to maintain SOFC temperature. To accommodate voltage losses, a higher SOFC area is required, which leads to increased capital cost. Furthermore, it reduces opportunity for anode gas recycle, auxiliary power production via steam turbine and internal energy integration.
4. SOFC Air utilization factor ( $U_a$ ): It represents the fractional conversion of oxygen in the SOFC. A low value of  $U_a$  allows for better heat management in SOFC, but results in increased air compression power and dilution at the cathode reduce SOFC efficiency.
5. Reformer temperature ( $T_{ref}$ ): High reformer temperature results in increased CH<sub>4</sub> conversion and H<sub>2</sub> yield, subsequently leading to high cell voltage and SOFC power generation. However, it also increases the heat load of the reformer and reduces net electrical efficiency. Moreover, with more CH<sub>4</sub> conversion, reformer capital cost and hot utility cost increases at the expense of improved SOFC efficiency.
6. Reformer S/C ratio ( $SCR$ ): A high value of SCR increases H<sub>2</sub> yield and potential of high SOFC efficiency. However, it also dilutes the anode gas, leading to lower open circuit voltage and increased hot utility for steam generation. Furthermore, it increases reformer and steam generator capital cost.
7. Reformer bypass fraction ( $f_{byp}$ ): This fraction controls the extent of external and internal reforming. External reforming results in high hydrogen partial pressure in the SOFC; hence, higher SOFC efficiency, thus lowering the feed demand and operating cost. However, it also suffers from increased hot utility and reformer capital cost. On the other hand, direct internal reforming provides better synergy between endothermic reforming and exothermic electrochemical reactions, thereby reducing the utility cost but results in lower partial pressure of hydrogen, cell voltage and thus increases SOFC capital cost.
8. Anode gas recycle fraction ( $f_{AGR}$ ): A high value of anode gas recycle fraction represents better material and energy integration and thus reduces feed and hot utility cost. However, it reduces auxiliary power production through steam turbine and thus increases SOFC capital cost.
9. Cathode gas recycle fraction ( $f_{CGR}$ ): A high value of cathode gas recycle fraction reduces the fresh air flow and the corresponding heating and compression loads, reducing hot utility cost and compression cost. However, it causes dilution of cathode gas and reduces cell voltage.

10. Flue gas split fraction ( $f_{split}$ ): This fraction controls the extent of internal energy integration and auxiliary power production via steam turbine. A high fraction of flue gas going to steam generator increases turbine power contribution and unloads the SOFC, resulting in lower SOFC capital cost. However, it limits internal energy integration and increases hot utility consumption. Consequently, the hot utility cost increases.
11. SOFC active surface area ( $A_{SOFC}$ ): A high value of active surface area reduces the current density, resulting in lower voltage loss and high operating voltage, and high SOFC efficiency, leading to a reduction in feed demand and corresponding operating cost. However, increased SOFC area results in increased capital cost of the SOFC.
12. Hot effluent split fractions ( $f_s$ ): There are total 5 such split fractions, two on exhaust gas and three on steam turbine exhaust. These split fractions control the relative extents of the heat transfer among various process heat exchangers.
13. Route selector ( $y$ ): This binary variable provides selection between anode gas recycle and biogas upgradation route. While anode gas recycle reduces feed cost through material integration, it also results in increased utility cost as well as increased SOFC capital cost. On the other hand, biogas upgradation route reduces SOFC capital cost but increases capital cost contribution from membrane separator and compressors.

### Process constraints

The key equality constraints are the material and energy balance equations for each process unit. These equations relate the various process inputs like pressure, temperature, flow rate, area, etc. to the key outputs like power generation, reaction conversion, etc. The details of these equations can be found in our earlier publication [9]. The key inequality constraint corresponds meeting power production target and is represented by Eq. (4).

$$P_{target} \leq P_{SOFC} + \sum_j W_{turb,j} - \sum_j W_{comp,j} \quad (4)$$

For each heat exchanger, there is a need to ensure positive thermal driving force. Correspondingly, the following constraints are incorporated for each of the process heat exchangers (PPX1 through PPX10).

$$T_{H,in,j} - T_{C,out,j} \geq \Delta T_{min}, \quad T_{H,out,j} - T_{C,in,j} \geq \Delta T_{min} \quad (5)$$

Lastly, upper and lower bounds for each decision variable are included.

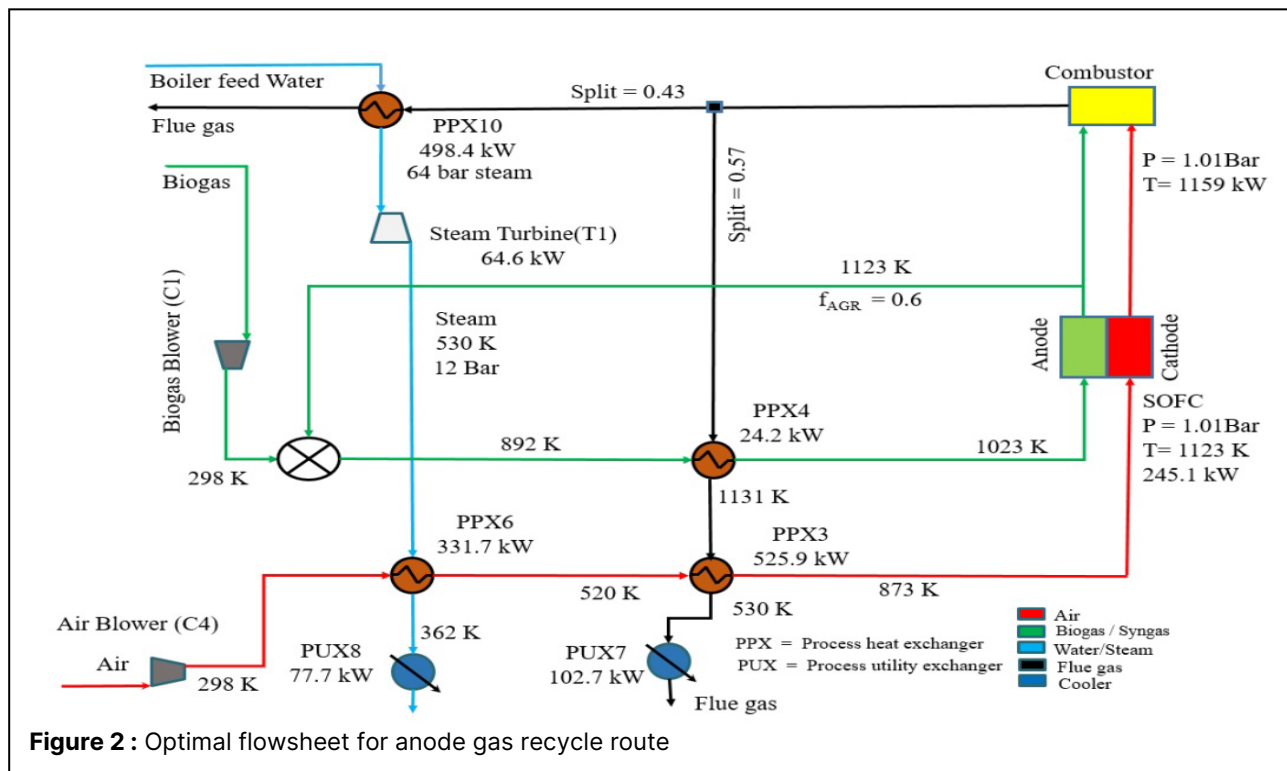
## RESULTS AND DISCUSSION

For this study, biogas composition of 65% CH<sub>4</sub> and 35% CO<sub>2</sub> is considered. Optimal flowsheet is designed for a power target of 300 kW. This capacity represents power potential of a small-scale wastewater treatment facility [5]. The optimization problem presented in the previous section is a mixed integer nonlinear programming (MINLP) problem. However, there is only one binary variable. It can therefore be easily split into two nonlinear programming problems (NLP), one for each processing route. As a NLP problem is easier to solve, we have obtained separate optimal flowsheets for each processing route. These NLP problems are solved using GAMS with Baron as a solver and the corresponding solutions are presented through Figure 2, 3 and Table 2.

**Table 2:** Optimal solution for each processing route (LB: lower bound, UB: Upper bound)

Variables	Anode gas recycle	Biogas upgradation
TAC (\$/y)	608,955	687,475
$\eta_{net}$ (%)	67.14	65.1
$\eta_{SOFC}$ (%)	54.86	55.48
$P_{SOFC}$ (kW)	245.119	255.712
$W_{turb,steam}$ (kW)	64.685	67.63
$W_{turb,gas}$ (kW)	-	0
$W_{comp,C1}$ (kW)	0	0
$W_{comp,C2}$ (kW)	-	6.702
$W_{comp,C3}$ (kW)	-	6.416
$W_{comp,C4}$ (kW)	-	0
$n_{BG}$ (mol/s)	0.857	0.884
$n_{Air}$ (mol/s)	46.143	46.094
$n_{BFW}$ (mol/s)	8.134	8.505
$Pr_{SOFC}$ (bar)	1.01 (LB)	1.01 (LB)
$T_{SOFC}$ (K)	1123 (UB)	1123 (UB)
$U_f$	0.78	0.85 (UB)
$U_a$	0.10 (LB)	0.10 (LB)
$T_{ref}$ (K)	-	-
SCR	-	-
$f_{byp}$	1.0 (UB)	1.0 (UB)
$f_{AGR}$	0.604	-
$f_{CGR}$	0.0 (LB)	0.0 (LB)
$f_{split}$	0.433	0.445
$A_{SOFC}$ (m <sup>2</sup> )	128.76	159.7

It can be seen that the anode gas recycle route results in lower TAC as compared to the biogas upgradation route. The optimal process recommends operating



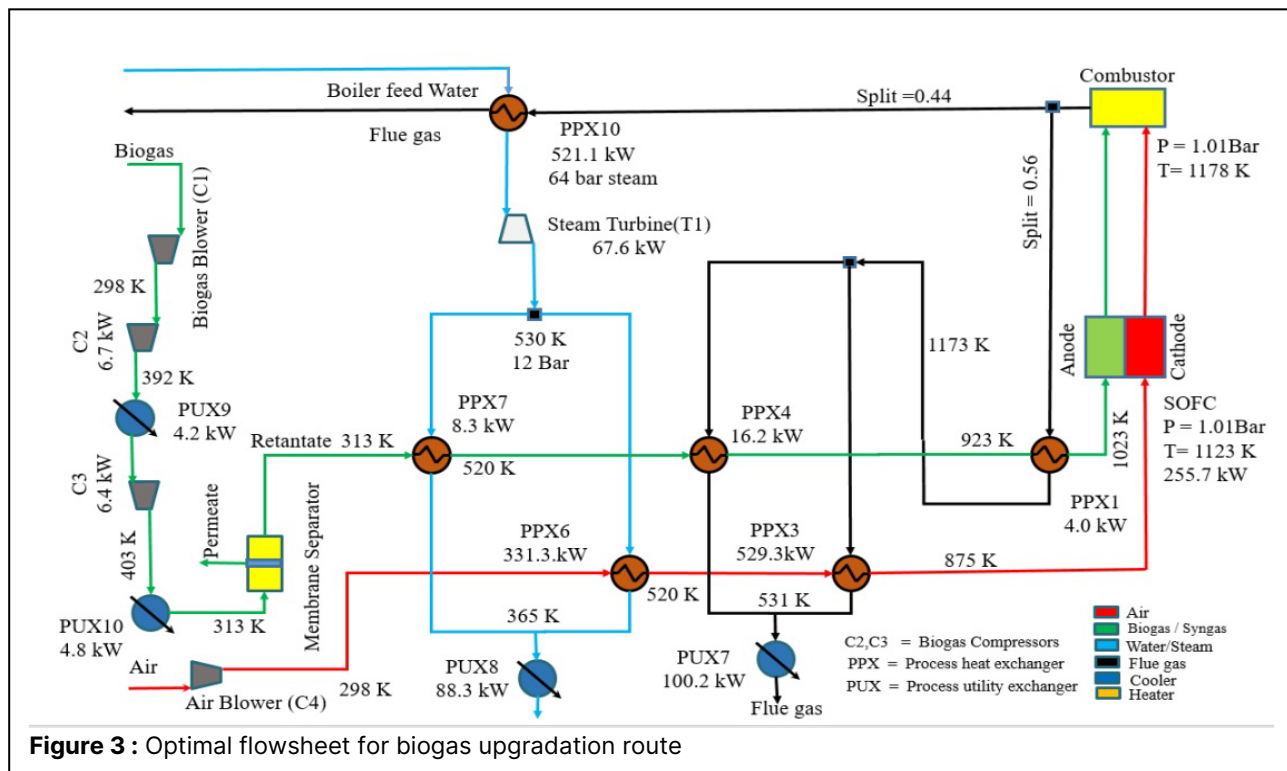
the SOFC at atmospheric conditions and high temperature (1123K). Total internal reforming is chosen over external reforming to save the capital cost of reformer and hot utility consumption for reforming reaction. Cathode gas recycling is not selected as dilution effects outweigh material and energy integration benefits. It can be noted that considerable fraction (60%) of anode gas is recycled back to utilize the unconverted hydrogen and generate power with higher efficiency (SOFC versus combustor + steam turbine). 43% of the high temperature flue gas is sent to the high-pressure steam generator for the auxiliary power generation and the rest of the fraction is sent to the process for heat integration to meet the internal heat demands of the process. By using heat from this flue gas as well as turbine exhaust, all of the heat demands are met. The SOFC generates 78% of the total power, and the steam turbine contributes the rest. Removal of the external reformer, efficient energy integration and synergy between the SOFC and steam turbine resulted in high net electrical efficiency of 67.1%. The total annualized cost of the plant is \$608955/y which corresponds to electricity production cost of \$0.23/kWh. This is slightly higher than typical electricity cost from conventional resources. The feed cost dominates the overall economics and contributes to 72% of the total annualized cost. This is due to capital intensive biogas production and cleanup process. (%). The cost of electricity can be further reduced by optimizing biogas production and cleanup process leading to reduced feed cost. The second most significant cost contribution comes from the SOFC (13%). Another way to reduce the cost of electricity

is to improve SOFC electrode design to reduce its cost per unit active surface area. This will allow for increasing the SOFC active surface area to further improve net electrical efficiency and lower the contribution of feed cost.

A similar trend is observed for the biogas upgradation route as well. Specifically, SOFC temperature and fuel utilization are set at upper bound, while pressure and air utilization are constrained at lower bound. Internal reforming is chosen over external reforming and cathode gas recycling is excluded. Even though more power is generated through the steam turbine, SOFC generates 82% of the total power, which is slightly higher than the previous case. This is due to increased power consumption via biogas compressors C2 and C3. It should be noted that the option of gas turbine is not selected as the power generated through it is sufficient enough to justify capital expenditure in turbine and the entry heat exchanger. 44.5% of the flue gas goes for auxiliary power generation, and the rest is utilized for heat integration. Similar to the previous case, this process also does not require any hot utility, resulting in net electrical efficiency of 65.1%. The total annualized cost for this process is \$687475/y and the corresponding cost of electricity is \$0.26/kWh. Similar to the previous process, the feed cost contributes significantly (66%) towards the total cost.

Let us now compare the two routes. The process with anode gas recycle has lower feed demand for the same power target and thus results in higher SOFC efficiency. Overall, the net electrical efficiency of anode gas recycle route is slightly better than biogas upgradation route. The other operating conditions of the two routes





**Figure 3 :** Optimal flowsheet for biogas upgradation route

are similar.

Let us now assess the impact of capital cost on the optimal design of the integrated process. Previously, optimal flowsheet was designed to minimize total operating cost [9]. Anode gas recycle was the most optimal option in that case as well. Most of the optimization variables had similar values. The total annualized cost of that process was \$609688/y, marginally higher than the proposed solution. This corroborates well considering that the feed cost dominates the overall economics.

The optimization problem for anode gas recycle involved 753 equations, 692 variables, and took 2583 s (CPU time) to solve. Similarly, the optimization problem for biogas upgradation involved 863 equations, 765 variables and took 4489 s to solve. The problems were solved on an Intel i7-7700 CPU @ 3.60 GHz processor.

## CONCLUSION

This paper presents optimal design for a biogas-based fuel cell system for electricity generation. A biogas superstructure consisting of various design alternatives is constructed. An optimal flowsheet is synthesized by minimizing the total annualized cost. The optimal flowsheet results in electrical cost of \$0.23/kWh. The following key conclusions can be drawn from this study.

1. Anode gas recycling route is more optimal as compared to biogas upgradation. A high value of recycle fraction is recommended to balance the trade-offs between material and energy integration and

auxiliary power generation.

2. Internal reforming provides better heat management, high net electrical efficiency, low capital investment, and low utility costs as compared to external reforming.
3. Feed biogas cost is the major contributor towards the total annualized cost.
4. Cathode gas recycle option does not provide net benefits towards efficiency improvement.

## REFERENCES

1. Galvagno A, Chiodo V, Urbani F, Freni F. Biogas as hydrogen source for fuel cell applications. *Int J Hydrogen Energy* 38(10):3913 – 3920 (2013)
2. Alves HJ, Junior CB, Niklevicz RR, Frigo EP, Frigo MS, Coimbra-Araújo CH. Overview of hydrogen production technologies from biogas and the applications in fuel cells. *Int J Hydrogen Energy* 38(13):5215 – 5225 (2013)
3. Piroonlerkgul P, Assabumrungrat S, Laosiripojana N, Adesina AA. Selection of appropriate fuel processor for biogas-fuelled SOFC system. *Chem Eng J* 140(1-3):341 – 351 (2008)
4. Farhad S, Hamdullahpur F, Yoo Y. Performance evaluation of different configurations of biogas-fuelled SOFC micro-CHP systems for residential applications. *Int J Hydrogen Energy* 35(8):3758 – 3768 (2010)



5. Trendewicz AA, Braun RJ. Techno-economic analysis of solid oxide fuel cell-based combined heat and power systems for biogas utilization at wastewater treatment facilities. *J Power Sources* 233:380 – 393 (2013)
6. Siefert NS, Litster S. Exergy and economic analysis of biogas fueled solid oxide fuel cell systems. *J Power Sources* 272:386 – 397 (2014)
7. Baldinelli A, Barelli L, Bidini G. Upgrading versus reforming: An energy and exergy analysis of two Solid Oxide Fuel Cell-based systems for a convenient biogas-to-electricity conversion. *Energy Conv Manag* 138:360 – 374 (2017)
8. Peters MS, Timmerhaus KD. Plant design and economics for chemical engineers. McGraw-Hill International (2018)
9. Uday V, Jogwar SS. Optimal design of an integrated biogas-based fuel cell system, *Ind. Eng. Chem. Res.* 63 (10):4496 – 4508 (2024)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Dimensionality Reduction in Optimal Process Design with Many Uncertain Sustainability Objectives

Hongxuan Wang<sup>a</sup>, Andrew Allman<sup>a,\*</sup>

<sup>a</sup> University of Michigan, Department of Chemical Engineering, Ann Arbor, MI, USA

\* Corresponding Author: [allmanaa@umich.edu](mailto:allmanaa@umich.edu)

---

## ABSTRACT

The study of sustainable design has gained prominence in response to the growing emphasis on environmental and social impacts of critical infrastructure. Addressing the different dimensions inherent in sustainability issues necessitates the application of many-objective optimization techniques. In this work, an illustrative four-objective design system is formulated, wherein uncertainties lie within two different socially-oriented objectives. A stochastic community detection approach is proposed to identify robust groupings of objectives. The findings reveal that the modularity of the optimal solution surpasses that of the average graph, thus demonstrating the efficacy of the proposed approach. Furthermore, a comprehensive exploration of the Pareto frontiers for both the robust and single-scenario best groupings is undertaken, demonstrating that using the robust grouping results in little to no information loss about tradeoffs.

**Keywords:** Sustainability, Multi-Objective Optimization, Network Theory

---

## INTRODUCTION

Various global events over the past several years have made clear the importance of designing new infrastructure not only at low cost but also that does not negatively impact global climate, that is resilient against disruptive events such as pandemics and wars, and that provides positive social outcomes for all relevant stakeholders [1]. For many large industries, shareholders are increasingly concerned about the significance of environmental, social, and governance (ESG) considerations within the framework of sustainability. The chemical industry is no exception to this trend. Consequently, there exists a compelling impetus to investigate sustainable process designs which perform well both in traditional economic metrics, such as net present value, annualized cost, or payback period, while also achieving positive social and environmental outcomes [2]. Moreover, when such environmental and social outcomes are considered, it is important to consider tradeoffs between different aspects of these main pillars. For instance, a socially positive outcome will consist of outcomes such as added jobs, safe operation, equitable outcomes among relevant stakeholders, and community acceptance, some of which may be in conflict with one another when considering

design alternatives. To evaluate the potential tradeoffs between these various goals, we employ many-objective optimization, an approach with widespread applicability in chemical process systems research. [3] The result of many-objective optimization is not a singular solution but a Pareto frontier, illustrating the tradeoffs between different objectives. All points along the Pareto frontier indicate the best one objective can do without hurting another one.

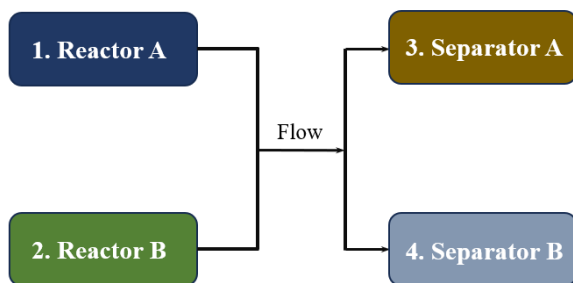
Unfortunately, for problems with large number of objectives, the time that is required to solve the sustainable design problem escalates significantly when the number of objectives increases, with problems of more than three objectives impractical to solve and interpret using rigorous solution methods such as the weighted sum or epsilon constraint approaches. Additionally, when Pareto frontiers can be obtained in these high dimensional spaces, they are challenging if not impossible for relevant stakeholders to interpret. To address these issues, our previous work developed a method for systematically reducing the dimensionality of a many-objective optimization problem on the basis of their competing or correlating nature using a network theoretic approach [4]. We then extended this framework to process operations problems where the underlying economic or

environmental parameters inherently varied in time (due to, for example, varying costs and emissions associated with grid-purchased electricity), demonstrating that the time-varying natures of these signals can alter the appropriate grouping of objectives when repeatedly solving the operation problem over time [5]. This work builds upon our previous efforts in the many-objective optimization space by analyzing how to group objectives which inherently contain a large degree of uncertainty. This phenomenon is most commonly seen when considering social objectives, which are typically ill-defined and difficult to quantify, and thus can be subject to large ranges of uncertainty.

In this work, we focus on an illustrative four objective design problem formulated as a linear program (LP). The objectives to minimize include net present costs, carbon emissions, safety risk and social inequity. We assume the presence of uncertain parameters in the two social objectives, risk and equity. The design problem is straightforward and illustrative in nature, with all four objectives being linear combinations of design variables.

The dimensionality reduction approach is systematically applied in all different scenarios generated from uncertain parameters. We want to find out the most robust grouping of objectives among the scenarios identifying the grouping of objectives with the highest expected modularity over all scenarios. To achieve this, a novel approach is developed which treats the community detection of an uncertain graph as a stochastic optimization problem and employs a column generation approach to decompose the problem into several interconnected deterministic community detection problems with modified modularity objectives. The remainder of this paper is structured as follows: in the next section, we describe the many-objective optimal design problem considered in this work. Then, a novel approach for stochastic community detection of uncertain graphs is described. Next, we present results of solving applying our dimensionality reduction approach and solving the many-objective optimization design problem. Finally, concluding remarks and avenues for future work are presented.

## PROBLEM FORMULATION



**Figure 1.** Two reactor, two separator superstructure considered in this work.

## System description

In this study, we demonstrate a straightforward design problem with two reaction and two separators as shown in Fig. 1, which depicts the superstructure of the reactor/separator system. The chemical reaction design model considered has two reactants and one product, necessitating separation in the designated separators as part of the design problem. Flow with both reactants allows for reactions in either Reactor A, Reactor B or both. Additionally, there are two choices for separators, and all flow resulting from the chemical reaction is directed towards these separators. Each reactor and separator is characterized by a unique conversion for both the reaction and separation processes.

Moreover, specific carbon emission and operating cost parameters are assigned to each unit in the system. There are uncertain parameters in the system that are associated with safety risk and social equity objectives; in this problem, we assume that there are two possible realizations of this objective parameter for each unit considered.

## Objective functions

The first objective that we consider is to minimize the annualized net present cost  $N$  that combines the capital cost and operating cost of the design system:

$$N = \sum_{i=1}^4 cap_i * \frac{1}{\theta} + op_i \quad (1)$$

$$cap_i = c_i^{ref} \left( \frac{F_i}{F_i^{ref}} \right)^\gamma \quad (2)$$

$$op_i = \alpha_i F_i \quad (3)$$

where  $\theta$  is the NPC-scaled lifetime used to annualize the capital cost and  $F$  is the flow through the reactors or separators, which acts as a proxy for unit size. The capital cost as we can see in equation (2) is a nonlinear function, where  $c_i^{ref}$  and  $F_i^{ref}$  are the sizes and capital costs, respectively, of the reference for unit  $i$ , and  $\gamma$  is the scaling exponent. Unit subscripts  $i$  correspond to the four units shown in Fig. 1. In order to formulate the problem as an LP such that our objective reduction framework can be applied, we choose  $\gamma$  as 1; extension of our dimensionality reduction framework to nonlinear problems will be the scope of future work. The operating cost is given in equation (3) and it is proportional to the flow through each unit, where  $\alpha_i$  are cost parameters and  $F_i$  are flow in each units.

Besides the traditional economic factors, we consider minimization of carbon emissions  $H$ , related to the usage of all the facilities in the system, where  $\beta_i$  represent emission parameter of different units:

$$H = \sum_{i=1}^4 \beta_i F_i \quad (4)$$

The emission objective has similar structure with operating cost: it is linear combination of the flow through

the different units. Moreover, two social objectives, risk and inequity, are analyzed in the optimization:

$$R = \sum_i \delta_{ri} F_i \quad (5)$$

$$E = \sum_i \delta_{ei} F_i \quad (6)$$

In this work, the coefficients for the two social objectives,  $\delta_{ri}$  and  $\delta_{ei}$ , are uncertain parameters. Note that for this illustrative example, the emissions, risk, and inequity objectives are a linear combination of all design variables, which for this problem are simply the flows. Thus, while it's uncommon to find that social objectives are determined by the same variables as cost and emission objectives in practical applications, the approach shown here is generalizable for linear objectives as variables which do not impact a given objective can have the corresponding coefficient set to zero. The extension of this approach to more realistic nonlinear problems is still ongoing work which is beyond the scope of this paper.

### Model Constraints

The chemical design system should follow physical and practical limitations. First, we set a total amount of initial feed in the system ( $F_t$ ) that equals to the sum of the flow in the reactors ( $F_1, F_2$ ):

$$F_1 + F_2 = F_t \quad (7)$$

$$F_t \geq F_i \geq 0$$

Since an  $A+B \rightarrow C$  reaction is assumed, total amount of flow (on a molar basis) will reduce after the reaction based on different conversions ( $E_1, E_2$ ) of the reactors:

$$\left(1 - \frac{E_1}{2}\right) F_1 + \left(1 - \frac{E_2}{2}\right) F_2 = F_3 + F_4 \quad (8)$$

We assume that the separators can distinguish the specific product that we want but can't totally remove the product from the flow. Thus, here we use the flow from the reactors to the respective separators ( $F_3, F_4$ ) and parameters for efficiency of the separators ( $E_3, E_4$ ), which are all based on the mixture of reactants and products.

$$E_3 F_3 + E_4 F_4 \geq F_{min} \quad (9)$$

Equation (9) indicates that a minimum requirement of the product ( $F_{min}$ ) needs to be satisfied in the reaction system.

### Case study data

As this problem acts to serve as an illustrative example, problem parameters are chosen by the researchers and not necessarily meant to be representative of a real process. We aim to apply our method to analyze objective tradeoffs for different practical process designs from the literature as future work.

The conversions of the two reactors are 80% and 95% which means that 80% or 95% of the reactants are transformed into product. The two separators can obtain 25%

and 40% of the product from the flow relatively. The total feed flow into the reactors is constrained as 100mol/h and the total flow goes into the separators is calculated from the conversions of the reactors.

For annualized net present cost, cost references of 2000, 5000, 700, 1200 dollars are set for the capital cost of the facilities. Operating cost and carbon emission parameters of all the facilities are set as 2, 0.9, 0.2, 0.15 dollar per mol/h and 80, 50, 80, 35 kg CO<sub>2</sub> per mol/h relatively.

**Table 1:** uncertain parameters in Risk and Equity objectives.

Facility	Risk parameter	Equity parameter
1	100 or 80	800 or 40
2	50 or 30	500 or 30
3	20 or 15	50 or 70
4	10 or 5	300 or 20

Risk and equity objectives are linear combination of the flow in each facility with uncertain parameters. Possible values of the parameters are shown in the Table 1. As social objectives such as equity can be difficult to quantify, we consider that these values can differ by a large amount. Each facility has two possible values of both risk and equity parameters, all 256 scenarios from their combinations are considered in the optimization. Despite having only two uncertain parameters, we note that the wide range of uncertainty and large number potential scenarios resulting from their combination showcases the ability of our proposed approach to deal with highly uncertain objectives. We also note that our proposed approach also allows for the quantification of uncertain parameters with a larger number of values.

### STOCHASTIC OBJECTIVE REDUCTION ALGORITHM

In this section, a process for systematically reducing the dimensionality of a MILP with many uncertain objectives into a problem with three or fewer objective functions is presented. The core of the algorithm, which considers deterministic objectives, was originally developed in our group's previous work and is summarized as follows [4]. First, cost vectors (the gradient vector of the objective function) are projected onto the constraint surfaces. Strength of interaction is defined as the inner product of the projected vectors and a weighted sum of the constraint interaction strengths is used to determine the total objective correlation strength. Objective correlation strength is scaled to be between 0 and 1. This information is embedded into an objective correlation graph, which consists of nodes corresponding to the

different objectives and edges weighted by the objective correlation strength. Community detection is applied to identify groups of objectives that are strongly correlated within the group and competing with objectives in other groups. For more details about the objective reduction algorithm, please reference our previous works [4,5].

For problems with uncertain objective functions, this uncertainty will manifest itself as uncertainty in the edge weights of the objective correlation graph. For each realization of uncertainty, a different objective correlation graph can be obtained. Since the true realization of uncertainty is not known *a priori*, it is essential to identify a grouping of objectives that performs robustly well for all realizations of uncertainty to be able to obtain a full understanding of the tradeoffs between the many objectives.

### Stochastic community detection

The community structure within a network manifests as a statistically significant configuration of edges which can be evaluated through the value of the modularity. [6] The modularity is described by the number of edges falling within groups minus the expected number in an equivalent network with edges placed at random, as depicted in formulation (10):

$$\max_z \sum_{i \in I, i' \in I} \left( \frac{A_{ii'}}{m} - \frac{a_i a_{i'}}{m^2} \right) z_{ii'} \quad (10a)$$

$$\text{s.t. } z_{ii} = 1 \quad \forall i \in I \quad (10b)$$

$$z_{ii'} + z_{ii''} - z_{i'i''} \leq 1 \quad \forall i \in I, i' \in I \setminus \{i\}, i'' \in I \setminus \{i, i'\} \quad (10c)$$

$$z_{ii'} \in \{0, 1\} \quad \forall i \in I, i' \in I \quad (10d)$$

where  $I$  is the set of nodes in the graph,  $A$  is the graph adjacency matrix,  $a_i = \sum_{i' \in I} A_{ii'}$  is the degree of node  $i$ ,  $m = \sum_{i \in I} a_i$  is twice of the total number of edges in the network, and  $z_{ii'}$  is a binary partitioning variable that is one if nodes  $i$  and  $i'$  are assigned to the same community, and zero otherwise. The community detection is accomplished by maximizing modularity.

To apply this approach to uncertain graphs, such as those obtained from our objective reduction approach with uncertain objectives, the traditional community detection problem must be recast as a stochastic optimization problem. In this case, a reasonable goal is to pursue the most robust partition that has consistently high modularity across all realizations of uncertainty. Here, we use expected value of modularity to evaluate partitions:

$$\max_z E_{\Sigma} \left[ \sum_{i \in I, i' \in I} \left( \frac{A_{ii'}}{m} - \frac{a_i a_{i'}}{m^2} \right) z_{ii'} \right] \quad (11)$$

$$\text{s.t. } (10b - d)$$

Note that other stochastic metrics, such as conditional value at risk, may also be used in place of expected value; we will assess how this proposed approach

extends to other metrics in future work. To simplify the equation, we assume all scenarios are equally likely and replace the expectation operator with an average over all scenarios:

$$\max_z \frac{1}{|K|} \sum_{k \in K} \sum_{i \in I, i' \in I} \left( \frac{A_{ii'k}}{m_k} - \frac{a_{ik} a_{i'k}}{m_k^2} \right) z_{ii'k} \quad (12)$$

$$\text{s.t. } (10b - d)$$

Set  $K$  indicates all the scenarios from the uncertain parameters. Unfortunately, fast algorithms for community detection such as spectral partitioning [7], fast unfolding [8], or the Leiden algorithm [9] are not directly applicable to the problem. However, this problem can be decomposed into a set of single-scenario community detection problems with a slightly modified objective. First, copy variables corresponding to each scenario are introduced to generate exploitable structure:

$$\max_z \frac{1}{|K|} \sum_{k \in K} \sum_{i \in I, i' \in I} \left( \frac{A_{ii'k}}{m_k} - \frac{a_{ik} a_{i'k}}{m_k^2} \right) z_{ii'k} \quad (13)$$

$$\text{s.t. } z_{iik} = 1 \quad \forall i \in I, k \in K$$

$$z_{ii'k} + z_{ii''k} - z_{i'i''k} \leq 1 \quad \forall i \in I, i' \in I \setminus \{i\},$$

$$i'' \in I \setminus \{i, i'\}, k \in K$$

$$z_{ii'k} \in \{0, 1\} \quad \forall i \in I, i' \in I, k \in K$$

$$z_{ii'k} = z_{ii'k'} \quad \forall i \in I, i' \in I, k \in K, k' \in K$$

In this problem, the non-anticipativity constraints  $z_{ii'k} = z_{ii'k'}$  are complicating, such that if they were removed, the stochastic problem could be treated as  $|K|$  independent community detection problems. Using column generation [10] can help to solve formulation (13). It can be transformed into the following master problem:

$$\max_{z, \lambda} \frac{1}{|K|} \sum_{k \in K} \sum_{c \in C} \lambda_{ck} f_{ck}^* \quad (14)$$

$$\text{s.t. } \sum_{c \in C} \lambda_{ck} z_{ii'ck}^* = z_{ii'} \quad \forall i \in I, i' \in I, k \in K \quad (\pi)$$

$$\sum_{c \in C} \lambda_{ck} = 1 \quad \forall k \in K \quad (\mu)$$

$$\lambda_{ck} \in \{0, 1\} \quad \forall k \in K, c \in C$$

where  $C$  is the set of columns generated which correspond to a specific partitioning of the graph into communities,  $f_{ck}^*$  is the modularity of column  $c$  in scenario  $k$ ,  $z_{ii'ck}^*$  is the partitioning variable in the corresponding modularity.  $\pi$  and  $\mu$  are the dual variables corresponding to the constraints in the same line, which are used to generate new columns (potential partitions) via the following set of  $|K|$  subproblems, one per scenario  $k$ :

$$\max_z \sum_{i \in I, i' \in I} \left( \frac{A_{ii'k}}{m_k} - \frac{a_{ik} a_{i'k}}{m_k^2} - \pi_{ii'k} \right) z_{ii'k} + \mu_k \quad (15)$$

$$z_{iik} = 1 \quad \forall i \in I, k \in K$$

$$z_{ii'k} + z_{ii''k} - z_{i'i''k} \leq 1 \quad \forall i \in I, i' \in I \setminus \{i\},$$

$$i'' \in I \setminus \{i, i'\}, k \in K$$



$$z_{i'k} \in \{0, 1\} \quad \forall i \in I, i' \in I, k \in K$$

Formulation (15) is the modularity in scenario  $k$  with two Lagrangian terms resulting from the column generation decomposition. Since both are constants obtained from the dual solution of the master problem, only the  $\pi z$  term needs to be considered when implementing the objective change into a community detection algorithm. Doing so is straightforward: using Newman's spectral partitioning algorithm [7], which makes use of the eigenvalues and eigenvectors of the modularity matrix  $M$ , we modify the modularity matrix by subtracting the symmetric part of the  $\pi$  matrix:

$$\widehat{M}_{ii'} = \frac{A_{ii'k}}{m_k} - \frac{a_{ik}a_{i'k}}{m_k^2} - \frac{\pi_{iik} + \pi_{i'ik}}{2} \quad (16)$$

The eigenvalues and eigenvectors of the new matrix  $\widehat{M}$  are then used to partition the graph in the same way as before. Alternatively, when using the fast unfolding (Louvain) [8] and Leiden [9] algorithms, we take grouping steps which give the largest increase in the modularity augmented with the Lagrangian  $\pi z$  terms, rather than just the modularity. As in any column generation approach, the algorithm proceeds by iteratively solving the master problem and subproblems until no subproblem returns a partition with positive objective value, indicating that no partition not already in the set of columns has the potential to improve the objective value.

## RESULTS AND DISCUSSION

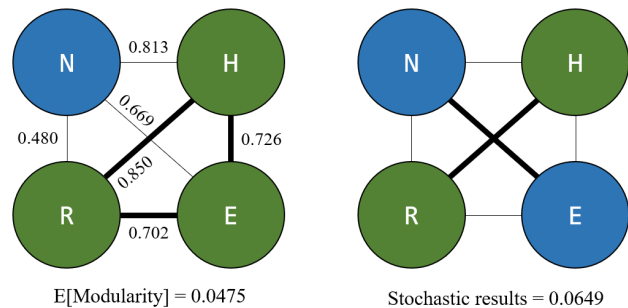
To study the stochastic design problem, objective correlation graphs are generated for all 256 possible scenarios of the social objective parameters. We then compare the results of community detection applied to each individual graph and show the resulting grouping frequencies are in Table 2. Notably,  $[[N, H], [R, E]]$  and  $[[N, E], [H, R]]$  emerge as the most common partitions, where the symbols N, E, H, and R refer to the four different objectives as introduced in the problem formulation section and defined again in Table 2. However, we observe that four different "best" objective groupings occur depending on which scenario is actually realized, and it is unclear which performs best, on average, over all possible scenarios with this approach. It's essential to highlight that there aren't any uncertain values in the annualized net present cost and emission objectives in our system, signifying a constant correlation strength between them. The established correlation strength, determined through our previous algorithm, is 0.813, a relatively but not overwhelmingly high value which explains the prevalence of 111 instances where annualized net present cost and emission fall within the same group.

In contrast, the introduction of uncertain parameters in risk and equity objectives brings variability to the correlation strengths between different objectives. For instance, with risk parameters set at [100, 50, 20, 10] and

equity parameters at [800, 500, 70, 20], the correlation strength between risk and equity is notably high at 0.991, resulting in the grouping  $[[N, H], [R, E]]$ . Conversely, when using [80, 30, 20, 10] and [40, 500, 50, 20] as risk and equity parameters, the correlation strength significantly decreases to 0.094, leading to the grouping  $[[N, E], [H, R]]$ . Thus, the frequencies that risk and equity are in the same group or not depends on the realization of uncertain parameter set of the two objectives. In this case, these two objectives are in the same group over half of the scenarios.

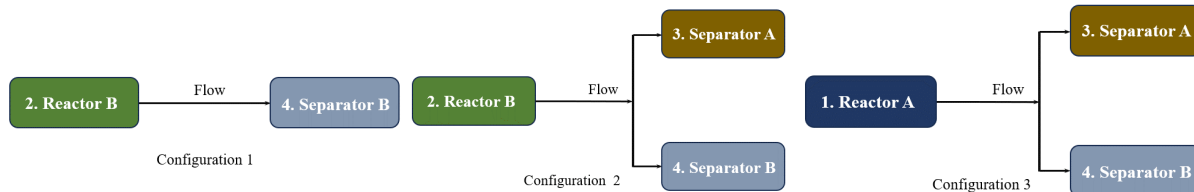
**Table 2:** Community detection results of the uncertain system. N: annualized net present cost, H: carbon emissions, R: risk, E: Equity.

Grouping	Frequency
$[[N, H], [R, E]]$	94
$[[N, E], [H, R]]$	95
$[[N], [H, R, E]]$	50
$[[R], [N, H, E]]$	17

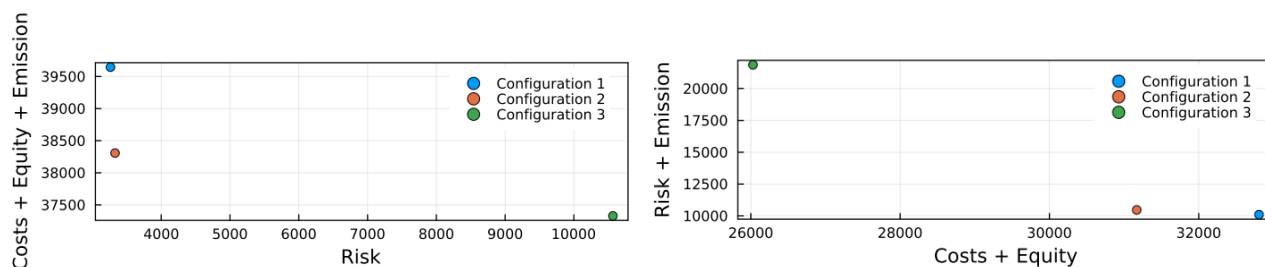


**Figure 2.** Grouping and expectation results of the average graph (left) and the most robust grouping (right)

Despite this variability, to address the tradeoffs within the uncertain design problem effectively, it is rational to identify the most robust partition across all scenarios as defined by the expected modularity over all scenarios. The results, as depicted in Fig. 2 through the application of the stochastic algorithm, reveal that the optimal grouping is,  $[[N, E], [H, R]]$ . It is important to note that this grouping differs from what one would find by just averaging the edge weights from the objective correlation graphs generated from all 256 scenarios,  $[[N], [H, R, E]]$ . The expectation of the modularity in the optimal solution is 0.0649, and the expectation of the second-best partition,  $[[N, H], [R, E]]$  is 0.0615. Both are higher than the partition of the average graph, which gives an expected modularity of 0.0475, suggesting that a



**Figure 3:** Illustration of different design configurations



**Figure 4:** Weighted sum Pareto frontier between (left) risk and cost + equity + emission as well as (right) risk+emission and cost + equity in the scenario of  $[[R], [H, N, E]]$  grouping.

deterministic approach of averaging edge weights is not an effective mechanism for identifying objective groupings that preserve tradeoff information in as many scenarios as possible.

For further study of the system, we compare the weighted sum Pareto frontiers for one scenario using both robust grouping found by community detection and the grouping that is “best” for the specific scenario considered,  $[R], [H, N, E]$ . Fig. 3 illustrates all optimal design configurations occurring along the Pareto frontier. Pareto frontiers are depicted on the left side of Fig. 4 for the scenario which uses risk and equity parameters set at  $[100, 30, 15, 5]$  and  $[40, 30, 70, 300]$ . Points sharing the same color on the Pareto frontier represent the same optimal design configuration within the Pareto frontier. In configuration 1, the emphasis is on risk reduction, leading to the utilization of Reactor B and Separator B, both possessing the lowest risk parameters. Conversely, Configuration 3 aims to minimize the sum of costs, equity, and emissions. Only Reactor A is constructed due to its lower cost, emissions, and equity parameters. Separator A and B are both utilized with specific flow values to optimize the three objectives.

In the middle of the Pareto frontier, the tradeoff point results from balancing the sum of costs, equity, and emissions against risk. In this case, only Reactor B is employed compared to Configuration 3, and the flow values in the separators are also different. The Pareto frontier for the optimal result's grouping,  $[[N, E], [H, R]]$ , is displayed on the right side of Fig. 4. The trade-off point using the  $[[R], [H, N, E]]$  grouping also appears in the Pareto frontier of  $[[N, E], [H, R]]$ , suggesting that employing the most robust grouping can provide valuable insights in a certain extent, and demonstrating that tradeoff

information can be preserved in the robust grouping even when it is not the same as the “best” grouping for a particular scenario.

However, in numerous other scenarios, the Pareto frontier of  $[[N, E], [H, R]]$  does lose a bit of information about design tradeoffs in comparison to the individual scenario's best grouping. This is not unexpected, as choosing a grouping of objectives with suboptimal modularity for a particular scenario inherently means that objectives with some degree of competition are being grouped together, which will result in loss of information regarding tradeoffs between the grouped objectives. The  $[[N, E], [H, R]]$  grouping with 95 times and other scenarios with no information loss in  $[[N, E], [H, R]]$  reveal that opting for the most robust grouping is a practical choice when attempting to understand tradeoffs in many-objective optimization problems with uncertain parameters.

## CONCLUSIONS AND FUTURE WORK

Environmental and social considerations for sustainability are essential aspects of optimal design problems for modern chemical production infrastructure. However, they can be subject to a great deal of uncertainty making critical evaluation of tradeoffs between objectives challenging. In this work, we developed a method for systematically identifying groups of objectives that, on average, tend to be more correlating than competing, forming the basis of a dimensionality reduction in many-objective optimization problems that is robust to uncertainty. Through the use of an illustrative case study with a two reactors, two separators superstructure, the efficacy of our approach towards identifying an objective grouping with low loss of tradeoff information was demonstrated.

As future work, we aim to apply this framework to a set of more practically relevant design problems from the literature such as a hydrogen production process and green ammonia production system, in order to better understand the tradeoffs inherent in designing future chemical production infrastructure. This analysis will entail further development of our objective reduction algorithm, such that it can be applied to nonlinear, as well as linear, many-objective optimization problems.

and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



## ACKNOWLEDGEMENTS

Financial support for this work was provided by the National Science Foundation (NSF) CAREER Award ([CBET-2237284](#)).

## REFERENCES

1. Pástor, L., Stambaugh, R. F., & Taylor, L. A. Sustainable investing in equilibrium. *Journal of Financial Economics* 142(2): 550-571 (2021)
2. Palys, M. J., Kuznetsov, A., Tallaksen, J., Reese, M., & Daoutidis, P. A novel system for ammonia-based sustainable energy and agriculture: Concept and design optimization. *Chemical Engineering and Processing-Process Intensification* 140: 11-21 (2019)
3. Rangaiah, G. P., Feng, Z., & Hoadley, A. F. Multi-objective optimization applications in chemical process engineering: Tutorial and review. *Processes* 8(5): 508 (2020)
4. Russell, J. M., & Allman, A. Sustainable decision making for chemical process systems via dimensionality reduction of many objective problems. *AIChE Journal* 69(2): e17962 (2023)
5. Russell, JM., Allman, A. Dynamic objective correlation in many-objective optimal operation: An ammonia production case study. *FOCAPO/CPC 2022* (2022)
6. Newman, M. E. Modularity and community structure in networks. *Proceedings of the national academy of sciences* 103(23): 8577-8582 (2006)
7. Newman, M. E., & Girvan, M. Finding and evaluating community structure in networks. *Physical review E*, 69: 026113 (2004)
8. Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10): P10008 (2008)
9. Traag, V. A., Waltman, L., & Van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Scientific reports* 9(1): 5233 (2019)
10. Desrosiers, J., & Lübbecke, M. E. A primer in column generation. In *Column generation* (pp. 1-32). Boston, MA: Springer US. (2005)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator

# Sustainable Aviation Fuels (SAF) from Ethanol: An Integrated Systems Modeling Approach

Madelynn J. Watson<sup>a</sup>, Aline V. da Silva<sup>b</sup>, Pedro G. Machado<sup>c</sup>, Celma O. Ribeiro<sup>c</sup>, Cláudio A.O. Nascimento<sup>d</sup>, Alexander W. Dowling<sup>a\*</sup>

<sup>a</sup> University of Notre Dame, Department of Chemical and Biomolecular Engineering, Notre Dame, IN, US

<sup>b</sup> Universidade Estadual de Campinas, Institute of Economics, Campinas, SP, Brazil

<sup>c</sup> University of São Paulo, Industrial Engineering Department, São Paulo, SP, Brazil

<sup>d</sup> University of São Paulo, Chemical Engineering Department, São Paulo, SP, Brazil

\* Corresponding Author: adowling@nd.edu

## ABSTRACT

This work explores the economic and environmental opportunities for sustainable aviation fuel (SAF) in the Brazilian sugarcane industry. Brazil was one of the first countries to use biomass fuels for transportation and is currently the 2<sup>nd</sup> largest producer of the world's bioethanol. Bioethanol produced from sugarcane can be upgraded to SAF via the American Society for Testing and Materials (ASTM)-certified pathway alcohol-to-jet (ATJ); however, at least two challenges exist for commercial implementation. First, technologies to produce bio-jet fuels cost more than their conventional fossil-based counterparts. Second, there is considerable uncertainty regarding returns on investment as the sugar and ethanol markets have been historically volatile. As such, we propose a new optimization model to inform risk-conscious investment decisions on SAF production capacity in sugarcane mills. Specifically, we propose a linear program (LP) to model an integrated sugarcane mill that can produce sugar, ethanol, or SAF. Then, using historical price data as scenarios, we determine optimal operation at different market scenarios. Based on the relationship between ethanol, sugar, and SAF prices, we show that the integrated sugarcane mill operates in four production regions. Furthermore, through sensitivity studies, we quantify the impact of SAF prices showing a premium SAF price of 2 \$ L<sup>-1</sup> results in 100% of scenarios favoring SAF production. These results allow us to guide SAF buyers or policymakers by showing the price point for SAF to become attractive for sugarcane mill integration.

**Keywords:** Biofuels, Optimization, Environment, Energy, Design Under Uncertainty

## INTRODUCTION

The aviation industry accounted for 2.8% of global GHG emissions in 2019. These emissions are expected to increase due to more global air traffic and high dependence on traditional fossil jet fuel [1]. In response to global sustainability efforts following the 2016 Paris Agreement, the aviation industry pledged to reduce emissions by 50% by 2050 [2]; however, decarbonization options are limited in this industry as electrification and advanced fuels (e.g., hydrogen) require two or three decades of development [3]. Sustainable Aviation Fuels (SAFs) are one of the most promising options for reducing greenhouse gas (GHG) emissions in aviation due to their technology readiness and drop-in compatibility with existing aircraft

and fuel systems; however, high costs compared to fossil-based fuels have led to slow capacity development [4].

SAF demand is expected to increase significantly by the year 2050 to combat high GHG emissions in the aviation industry. To reach the SAF demand track for net-zero emissions by 2050, the current global SAF capacity would need to increase by ~46 billion L by 2030 [4,5]. Thus, there is an urgent need to increase global SAF capacity efficiently and economically to meet 2050 sustainability goals.

Brazil was one of the first countries to use biofuels for transportation in the 1970s. In 2023, 25.3 million m<sup>3</sup> of bioethanol was consumed (mostly to fuel private vehicles in the country, 10.1 million m<sup>3</sup> used as a gasoline

blend, and 15.2 million m<sup>3</sup> used as direct fuel) [6]. Brazil is currently the second-largest producer of the world's bioethanol and the first-largest bioethanol producer from sugarcane [7]. A conventional sugarcane mill starts with juice extraction from crushed sugarcane stalks. The extracted juice is treated to remove impurities, and the clarified juice is then shared between the ethanol distillery and the sugar factory. The juice split fractions for ethanol and sugar production are decided at the beginning of a harvesting season and are important operational decisions for the mill [8]. Typical juice split fractions vary between 40% and 60%, with many sugarcane mills only containing installed capacity for sugar production between 50–75% of the total juice extracted during one season [8].

Several existing sugarcane mills have already expanded their product range beyond sugar and ethanol to produce fertilizer, industrial salts, animal feed, etc. These diversified products boost the overall financial security of the sugarcane mill and provide an opportunity to jump into previously unexplored markets [9–10]. Following their footsteps, bioethanol can be upgraded to jet fuel via the ASTM-certified pathway alcohol-to-jet (ATJ) [11]. ATJ converts biomass to an alcohol intermediate (typically ethanol) to jet fuel via the following three conversion steps: alcohol dehydration, oligomerization, and hydrogenation. Partnerships targeting SAF production exist between companies producing alcohols and converting alcohols to fuels, notably BioChemtex/Gevo and LanzaTech/Swedish Biofuels [12].

Bioethanol from sugarcane is considered a first-generation biofuel since it is derived from a food-based feedstock. Although first-generation biofuels are questioned due to food security, they are currently economically feasible and commercially supplied [13]. Sustainable biofuel production generally considers second-generation (non-edible crops) to fourth-generation (waste) feedstocks [14]. From an environmental perspective, Foteinis et al. [15] illustrate the non-linear nature of the relationship between feedstock generation and environmental impact. Foteinis et al. [15] compare life cycle analysis (LCA) results for biodiesel production from sunflower seed oil (first-generation), used cooking oil (second-generation), and micro-algae (third-generation). Biodiesel from used cooking oil (second-generation) had the lowest environmental impact and third-generation biodiesel had the highest environmental impact [15].

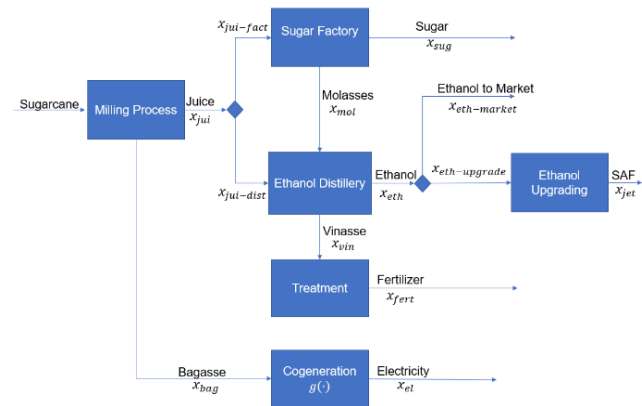
Integration of SAF in the Brazilian sugarcane industry presents complex and interdependent optimization problems. Several factors complicate the decision to invest in new technologies in the sugarcane industry. First, the mill's final product mixes must be decided at the beginning of a harvesting season, with little operational flexibility throughout the remaining season [16]. Second, weather conditions, international sugar prices, and

instability of governmental fuel policies and subsidies create substantial uncertainties in sugar and ethanol market prices [8]. The interdependence of these two factors leads to large uncertainty on the return on investment, which hinders risk-conscious mill owners from investing in new technologies. Furthermore, it is well known that SAF technologies are not cost-competitive with their fossil-based counterparts [4], leading to additional financial risk with SAF integration.

Mutran et al. [8] developed a multi-objective, two-stage stochastic program to de-risk bioenergy investment decisions in the sugarcane industry to address some of these complexities. In this work, we extend the optimization analysis of Mutran et al. [8] to inform risk-conscious investment decisions on SAF capacity in sugarcane mills by considering the uncertainty in sugar, ethanol, and jet fuel markets and the interrelation between production decisions in terms of the final product mix. Furthermore, we quantify the impact of government incentives or environmentally conscious industry contracts on SAF economics.

## METHODS

We propose a linear program (LP), adapted from [8], to study the economic risk of integrated SAF production in a sugarcane mill in Brazil. Specifically, we study the optimal operation of an integrated sugarcane mill that can produce sugar, ethanol, and SAF.



**Figure 1.** Process flow diagram (PFD) of integrated SAF production in a single sugarcane mill.

Fig. 1 presents a process flow diagram of the integrated sugarcane mill process. The units in this process include sugarcane milling (mill), sugar factory (fact), ethanol distillery (dist), treatment of vinasse residues (treat), a cogeneration system using a traditional Rankine cycle with back-pressure turbine (rank), and an ethanol upgrading to SAF (upgrade). Saleable products from the sugarcane plant comprise sugar (sug), ethanol directed to the market (eth-market), SAF (jet), fertilizer from



vinasse residue (fert), as well as electrical power for the regulated market (el). In contrast, intermediates are considered to be juice (jui), juice directed to the sugar factory (jui-fact), juice directed to the ethanol distillery (jui-dist), ethanol directed to the upgrading (eth-upgrade), bagasse (bag), molasses (mol), and vinasse (vin).

**Table 1:** Nominal parameter values from [8] except  $\lambda_{jet}$  and  $\omega_{eth,upgrade,jet}$  which originate from [17].

Symbol	Description	Nominal Value
$\omega_{cane,mill,jui}$	Conversion	0.734 tonne/tonne
$\omega_{cane,mill,bag}$	Conversion	0.266 tonne/tonne
$\omega_{jui,fact,sug}$	Conversion	0.123 tonne/tonne
$\omega_{jui,dist,eth}$	Conversion	0.123 tonne/m <sup>3</sup>
$\omega_{mol,dist,eth}$	Conversion	0.373 tonne/m <sup>3</sup>
$\omega_{eth,upgrade,jet}$	Conversion	0.4284 m <sup>3</sup> /m <sup>3</sup>
$\omega_{vin,treat,fert}$	Conversion	1 m <sup>3</sup> /m <sup>3</sup>
$\rho_{mol,fact}$	Generation	1.05 tonne
$\rho_{vin,dist}$	Generation	10 m <sup>3</sup>
$\Gamma_{mill}$	Maximum Capacity	3,000,000 tonne/year
$\Gamma_{fact}$	Maximum Capacity	2,202,000 tonne
$\Gamma_{upgrade}$	Maximum Capacity	270,000 m <sup>3</sup>
$\Gamma_{dist}$	Maximum Capacity	2,202,000 tonne
$\gamma_{sug}$	Minimum Production	0.4
$\gamma_{eth}$	Minimum Production	0.4
$\gamma_{eth-market}$	Minimum Production	0.2
$\gamma_{jet}$	Minimum Production	0.4
$\lambda_{sug}$	Production Cost	1,150 \$/tonne
$\lambda_{eth}$	Production Cost	1,800 \$/m <sup>3</sup>
$\lambda_{jet}$	Production Cost	4,600 \$/m <sup>3</sup>
$\lambda_{el}$	Production Cost	7 \$/MWh
$\lambda_{fert}$	Production Cost	0.002 \$/m <sup>3</sup>

In the mathematical model formulation that follows, the continuous variables  $x_i \geq 0$  refer to the yearly production of product  $i$ , the parameters  $\omega_{i,u,j}$  refer to the conversion of product  $i$  to product  $j$  in unit  $u$ , the parameters  $\rho_{j,u}$  refers to the yield of intermediate product  $j$  in unit  $u$ , the parameters  $\lambda_i$  refers to the production cost to produce product  $i$ , the parameters  $\pi_i$  refers to the market price for product  $i$ , the parameters  $\Gamma_u$  refers to the maximal annual processing capacity of unit  $u$ , and the parameters  $\gamma_i$  refers to the minimum production fraction of

product  $i$ . Table 1 reports the nominal parameter values.

The LP is formulated below in Eq. (1). The objective is to maximize the sugarcane mill's expected revenue (i.e., profit) Eq. (2) constrained by mass balances Eq. (3)-(13), unit capacity restrictions Eq. (14)-(16), and minimum production constraints Eq. (17)-(20). Minimum production constraints are enforced to maintain typical juice shares for sugar and ethanol production (40 – 60%), limit food supply (sugar) competition, sufficiently utilize installed SAF capacity, and require a contribution to the bioethanol available for private vehicle use.

This optimization formulation is an adaption of the model proposed by Mutran et al. [8] with the addition of the SAF upgrading unit Eq. (9)-(10) and the minimum production constraints Eq. (16)-(19). Electricity generation is determined using a piece-wise linear surrogate model to determine the amount of surplus bagasse based on the shares of ethanol and sugar. We denote this relationship as  $g(\cdot)$ , represented by Eq. (18)-(20) in Mutran et al. [8].

$$\max R(x|\pi, \lambda, \omega, \rho)$$

$$s. t. \text{ Eq. (3) – (20)} \quad (1)$$

$$g(\cdot) \geq 0$$

$$x_i \geq 0$$

$$R(x) = (\pi_{sug} - \lambda_{sug})x_{sug} + \pi_{eth}x_{eth-market} - \lambda_{eth}x_{eth} + (\pi_{jet} + \pi_{premium} - \lambda_{jet})x_{jet} + (\pi_{el} - \lambda_{el})x_{el} + (\pi_{fert} - \lambda_{fert})x_{fert} \quad (2)$$

$$x_{jui} = \Gamma_{mill}\omega_{cane,mill,jui} \quad (3)$$

$$x_{bag} = \Gamma_{mill}\omega_{cane,mill,bag} \quad (4)$$

$$x_{jui} = x_{jui-fact} + x_{jui-dist} \quad (5)$$

$$x_{sug} = x_{jui-fact}\omega_{jui,fact,sug} \quad (6)$$

$$x_{mol} = x_{sug}\rho_{mol,fact} \quad (7)$$

$$x_{eth} = x_{jui-dist}\omega_{jui,dist,eth} + x_{mol}\omega_{mol,dist,eth} \quad (8)$$

$$x_{eth} = x_{eth-market} + x_{eth-upgrade} \quad (9)$$

$$x_{jet} = x_{eth-upgrade}\omega_{eth,upgrade,jet} \quad (10)$$

$$x_{vin} = x_{eth}\rho_{vin,dist} \quad (11)$$

$$x_{fert} = x_{vin}\omega_{vin,treat,fert} \quad (12)$$

$$x_{el} = g(\cdot) \quad (13)$$

$$x_{jui-fact} \leq \Gamma_{fact} \quad (14)$$

$$x_{jui-dist} \leq \Gamma_{dist} \quad (15)$$

$$x_{eth-upgrade} \leq \Gamma_{upgrade} \quad (16)$$

$$x_{jui-fact} \geq \gamma_{sug}x_{jui} \quad (17)$$

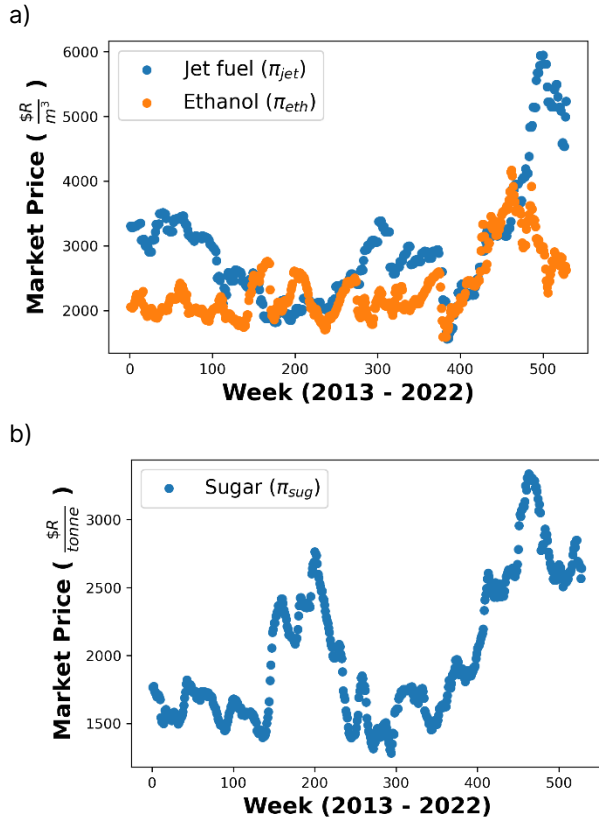
$$x_{eth} \geq \gamma_{eth}x_{jui}\omega_{jui,dist,eth} \quad (18)$$

$$x_{eth-market} \geq \gamma_{eth-market}x_{eth} \quad (19)$$

$$x_{jet} \geq \gamma_{jet}x_{eth} \quad (20)$$

Product prices for sugar, ethanol, and jet fuel are taken from weekly price records in Brazil over the period of 2013–2022 [18-20] a total of 526 observations. Fig. 2 shows the weekly price signals ( $\pi_{jet}$ ,  $\pi_{eth}$ , and  $\pi_{sug}$ ) for the three main products expressed in the 2023 \$R. We

assume that SAF prices are equivalent to market prices of jet fuel plus some premium ( $\pi_{premium}$ ) paid by airline companies or policy incentives.



**Figure 2.** Weekly historical market data for ethanol and jet fuel (a) and sugar (b) in Brazil from 2013 to 2022 expressed in the 2023 \$R [18-20].

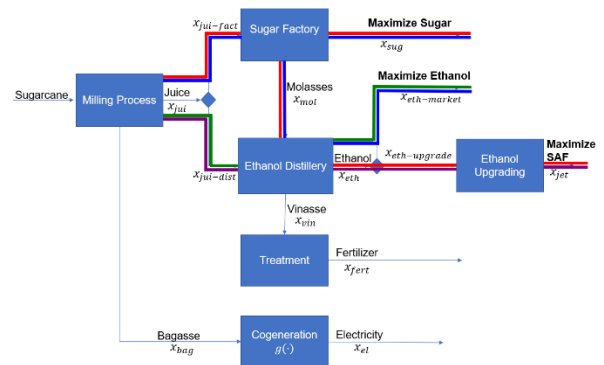
## RESULTS

We start by solving optimization problem Eq (1) – (20) independently for each week from 2013 to 2022, considering historical price data for sugar, ethanol, and jet fuel (Fig. 2). We then map the optimization results for each scenario (week) into four categories based on the core products:

1. Maximize sugar and send the maximum amount of residual ethanol to the market (blue in Fig. 3 and Fig. 4).
2. Maximize sugar and upgrade the maximum amount of residual ethanol to SAF (red in Fig. 3 and Fig. 4).
3. Maximize ethanol and send the maximum amount of residual ethanol to the market (green in Fig. 3 and Fig. 4).
4. Maximize ethanol and upgrade the maximum amount to SAF (purple in Fig. 3 and Fig. 4)

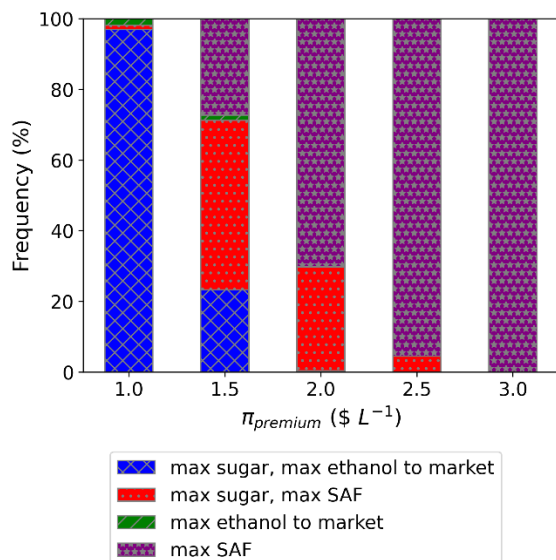
Fig. 3 shows the optimal operation pathways

colored on the process flow diagram. Fig. 4 shows the percentage of market scenarios (weekly) for each of the four operation pathways while varying the SAF premium parameter  $\pi_{premium}$ . At a SAF premium price of 1 \$ L<sup>-1</sup>, the sugarcane mill prefers to operate under standard practice, maximizing sugar production and sending any ethanol produced to the market. On the contrary, at a SAF premium of 3 \$ L<sup>-1</sup>, the sugarcane mill prefers to send all its resources towards maximizing SAF production. Both of these scenarios represent cases where there is one predominant operating mode regardless of market prices. In between these cases ( $\pi_{premium} = 1.5$  to 2.5 \$ L<sup>-1</sup>) we observe variation between optimal operating modes. Furthermore, after applying a SAF premium of 2 \$ L<sup>-1</sup>, 100% of all market scenarios favor SAF production (red and purple) over selling ethanol to the market (green and blue).



**Figure 3.** Optimal operation pathways colored on the process flow diagram (PFD) of integrated SAF production in a single sugarcane mill.

The multiple optimal operating modes in SAF premium cases 1.5 \$ L<sup>-1</sup> to 2.5 \$ L<sup>-1</sup> (Fig. 4) motivate investigating the operational flexibility in the sugarcane mill, i.e., changing the product flow rates in response to weekly market conditions. Standard practice in the sugarcane industry is to choose ethanol and sugar shares at the beginning of the harvesting season with little flexibility throughout the year. Our analysis shows there may be some financial benefit to change operation in response to market prices (different optimal operating modes). Profit savings from flexibility may be significant for SAF integration in the sugarcane industry, especially with uncertainty in monetary incentive timelines for SAF.



**Figure 4.** Percentage of scenarios under each operating region at different levels of SAF premiums

Furthermore, the competition that may arise between ethanol use for ground and air transportation should be carefully considered. If the majority of sugarcane mills favor SAF production over selling ethanol to the market (e.g., SAF premium of 2 \$ L<sup>-1</sup>), private vehicles in the country may be forced to use gasoline (electric vehicles in Brazil are currently underdeveloped). Future work should include LCA to better understand the environmental impact of producing SAF from sugarcane bioethanol and bioethanol produced from other sugarcane mill waste residues (i.e., bagasse).

## CONCLUSION

In summary, we proposed a LP to model SAF production integrated in the Brazilian sugarcane industry. Then, using historical prices, we compute the optimal operation over a variety of market scenarios for ethanol, sugar, and SAF prices. We show that the integrated sugarcane mill operates in four production regions based on the relationship between ethanol, sugar, and SAF prices. Furthermore, through sensitivity studies, we quantify the impact of SAF prices (e.g., price premiums airlines may pay for SAF compared to conventional jet fuel to meet industry/government decarbonization goals) and discuss the potential for operational flexibility in a sugarcane mill.

This research motivates the following open research questions:

1. How to economically distribute ethanol upgrading technologies throughout Brazil to allow for ethanol-use flexibility (integrated vs. stand-alone)?
2. How to develop detailed process designs for integrated SAF-producing systems under uncertainty?
3. How to integrate techno-economic analysis and

policy design for biofuel technologies?

4. What is the difference in carbon intensity in the ground and air transportation sectors?

## ACKNOWLEDGEMENTS

We declare no conflicts of interest. M.J.W. acknowledges the Department of Education [Grant P200A210048] and the Arthur J. Schmitt Presidential Leadership Fellowship administered by the University of Notre Dame. A.W.D. acknowledges the U.S. National Science Foundation [Cooperative Agreement No. EEC-1647722]. A.S, P.M., C.R, and C.N. acknowledge the Project Fapesp/Research Centre for Greenhouse Gas Innovation [Grant 2020/15230-5]; and the CNPq (Brazilian National Research Council), [Grant 308909/2021-6].

## REFERENCES

1. IEA (2020). Tracking Aviation 2020 <https://www.iea.org/reports/tracking-aviation-2020>
2. IATA (2021). Resolution on the Industry's Commitment to Reach Net Zero Carbon Emissions By 2050.
3. Dahal K, Brynolf S, Xisto C, Hansson J, Grahn M, Grönstedt T. Techno-economic review of alternative fuels and propulsion systems for the aviation sector. *Renew. Sustain Energy Rev.* (2021)
4. Watson MJ, da Silva AV, Machado PG, Rivera YS, Dowling AW, Ribeiro CO, Nascimento CAO. Sustainable aviation fuel technologies, costs, emissions, policies, and markets: a critical review. *J. Clean. Prod.* 449 (2024)
5. IEA (2021). Renewables 2021 - Analysis and forecast to 2026.
6. UNICA (2024). Harvest 2023/2024 updates. Retrieved April 1, 2024, from <https://www.novacana.com/noticias/atualizacao-safra-cana-de-acucar-2023-24-quinzena-janeiro-250124>
7. Li L, Lin J, Wu N, Xie S, Meng C, Zheng Y, Wang X, Zhao Y. Review and outlook on the international renewable energy development. *EBE* 3(2):139-157 (2022)
8. Mutran VM, Ribeiro CO, Nascimento CAO, Chachuat B. Risk-conscious optimization model to support bioenergy investments in the Brazilian sugarcane industry. *Appl. Energy* 258 (2020)
9. Klein BC, de Mesquita Sampaio IL, Mantelatto PE, Filho RM, Bonomi A. Beyond ethanol, sugar, and electricity: a critical review of product diversification in Brazilian sugarcane mills. *Biofpr*, 13(3):809-821 (2019)
10. Formann S, Hahn A, Janke L, Stinner W, Sträuber H,

- Logroño W, Nikolausz M. Beyond sugar and ethanol production: value generation opportunities through sugarcane residues. *Front. Energy* 8 (2020)
11. Santos CI, Silva CC, Mussatto SI, Osseweijer P, van der Wielen LAM, Posada JA. Integrated 1st and 2nd generation sugarcane bio-refinery for jet fuel production in Brazil: Techno-economic and greenhouse gas emissions assessment. *Renew. Energ.* 129:733–747 (2019)
  12. Mawhood R, Gazis E, de Jong S, Hoefnagels R, Slade R. Production pathways for renewable jet fuel: a review of commercialization status and future prospects. *Biofpr* 10(4):462-484 (2016)
  13. Liu W, Junming X, Xinfeng X, Yan Y, Xiaolu Z, Changhui P. A new integrated framework to estimate the climate change impacts of biomass utilization for biofuel in life cycle assessment. *J. Clean. Prod.* 267 (2020)
  14. Doliente SS, Narayan A, Tapia JFD, Samsatli NJ, Zhao Y, Samsatli S. Bio-aviation fuel: a comprehensive review and analysis of the supply chain components. *Front. Energy Res.* 8 (2020)
  15. Foteinis S, Chatzisyneon E, Litinas A, Tsoutsos T. Used-cooking-oil biodiesel: life cycle assessment and comparison with first- and third-generation biofuel. *Renew. Energy* 153:588-600 (2020)
  16. Cavalett O, Junqueira TL, Dias MOS, Jesus CDF, Mantelatto EM, Cunha MP, Franco HCJ, Cardozo TF, Filho RM, Rossell CEV, Bonomi A. Environmental and economic assessment of sugarcane first generation biorefineries in Brazil. *Clean Technol. Environ. Policy*, 14(1):399–410 (2012)
  17. Restrepo-Flórez JM & Maravelia CT. Advanced fuels from ethanol-a superstructure optimization approach. *Energy Environ. Sci.*, 14(1):493–506 (2021)
  18. *Sugar*. Center for Advanced Studies on Applied Economics (CEPEA), “Luiz de Queiroz” College of Agriculture (ESALQ), USP. Retrieved February 1, 2023, from <https://www.cepea.esalq.usp.br/br/indicador/acucar.aspx>
  19. *Ethanol*. Center for Advanced Studies on Applied Economics (CEPEA), “Luiz de Queiroz” College of Agriculture (ESALQ), USP. Retrieved February 1, 2023, from <https://www.cepea.esalq.usp.br/br/indicador/etanol.aspx>
  20. *Prices of producers and importers of petroleum and biodiesel derivatives*. National Agency for Petroleum, Natural Gas and Biofuels. Retrieved February 1, 2023, from <https://www.gov.br/anp/pt-br/assuntos/precos-e-defesa-da-concorrenca/precos/precos-de-produtores-e-importadores-de-derivados-de-petroleo>

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Exploring Net-Zero Greenhouse Gas Emission Routes for Bio-Production of Triacetic Acid Lactone: An Evaluation through Techno-Economic Analysis and Life Cycle Assessment

Ching-Mei Wen<sup>a</sup>, Charles Foster<sup>b</sup>, and Marianthi Ierapetritou<sup>a\*</sup>

<sup>a</sup> University of Delaware, Department of Chemical & Bio-molecular Engineering, Newark, Delaware, United States

<sup>b</sup> DSM-Firmenich, 6480 Dobbin Rd, Columbia, Maryland, United States

\* Corresponding Author: [mgi@udel.edu](mailto:mgi@udel.edu).

## ABSTRACT

Triacetic acid lactone (TAL) is a bio-privileged molecule with potential as a chemical precursor, traditionally synthesized from petroleum. Current trends are shifting towards the use of renewable biomass or CO<sub>2</sub>-derived feedstocks to enhance sustainability. However, comprehensive studies on the techno-economic viability and carbon life cycle of such methods are limited. This study assesses TAL production from conventional glucose and a novel approach co-feeding *Yarrowia lipolytica* (YL) with glucose and formic acid (FA), aiming for a more cost-effective and eco-friendly process. We confront the inherent challenges in this process by exploring different technology scenarios using kinetic bioprocess modeling underpinned by techno-economic analysis (TEA) and life cycle assessment (LCA) to identify the most cost-effective and sustainable routes to TAL production. A noteworthy component of our investigation centers around the prospect of recycling and utilizing the CO<sub>2</sub> emitted from the YL bioreactor to eliminate greenhouse gas emissions inherent in aerobic fermentation processes. The study combines TEA and LCA to dissect the proposed TAL bio-production routes, evaluating the sustainability of the process and the implications of net-zero greenhouse gas emission manufacturing. We employed SuperPro Designer and Aspen software for process simulation and energy balance computations. The results underscore the benefits of CO<sub>2</sub> recycling in TAL production, with an estimated minimum selling price (MSP) slightly increasing by 6.21-7.80% compared to traditional methods, but significantly undercutting the market price of \$51000/mt-TAL and achieving net-negative CO<sub>2</sub> emissions. This research illustrates a viable route to bio-production with net-zero emissions, providing a model for future bioprocessing and industrial practices.

**Keywords:** Technoeconomic Analysis, Life Cycle Assessment

## INTRODUCTION

Energy systems across the globe are moving towards more integrated, cleaner, and sustainable processes [1]. However, attaining a carbon-free economy presents a formidable challenge, as it requires a significant reduction in emissions from hard-to-decarbonize sectors such as industrial and chemical processes. This study explores the production of triacetic acid lactone (TAL) through carbon dioxide utilization in aerobic

fermentation cycles. The goal is to achieve Net-Zero emissions by identifying economically viable and environmentally friendly processes to produce chemicals. The study aims to guide the development of sustainable industrial methodologies using TAL as a case-study.

TAL is a promising molecule for sustainable chemical production, obtained through bio-production from renewable resources [2]. It has the potential to serve as a key precursor for the production of valuable chemical intermediates, as well as end products like sorbic acid [3],



which have significant global demand (market size was USD 500 million in 2022, reported by Reports and Data, a market research and consulting company). Traditional synthesis methods rely on petrochemical processes that are not sustainable [4]. With the engineering *Yarrowia lipolytica* (YL), a yeast and model organism and naturally proficient lipid producer, to produce TAL, we can shift towards using biomass as a renewable feedstock [5-6]. This addresses the growing need for carbon-neutral production pathways. However, bioproduction involves metabolic processes that produce significant amounts of CO<sub>2</sub>, which is typically considered waste [7]. In this study, the bioprocess emits CO<sub>2</sub> at concentrations greater than 99 wt.% by assuming high-purity O<sub>2</sub> used in providing aerobic fermentation. [8]. Rather than releasing it into the atmosphere, there is an opportunity to repurpose this emission [8]. The current trend is to convert CO<sub>2</sub> into value-added chemicals, which aligns with the industrial symbiosis paradigm of repurposing waste streams to reduce environmental impact [9].

One innovative approach is the electrocatalytic conversion of CO<sub>2</sub> into formic acid [10-11]. This process recycles carbon and transforms it into a chemical with significant utility in various industrial applications. Formic acid is an auxiliary energy source for numerous microbial species that use formate dehydrogenase enzymes (FDH) to transfer electrons to NAD<sup>+</sup>, generate NADH and release CO<sub>2</sub> [10-11]. This metabolic pathway is central to the co-feeding strategy used in the TAL production process with YL [10-11]. By co-feeding formic acid with glucose, YL can utilize additional reducing power for biosynthesis, contributing to the efficient production of TAL. The integration of CO<sub>2</sub>-to-formic acid utilization in the production process has a dual benefit. It provides a supplementary energy source that enhances microbial growth and product yield and represents a strategic move toward decarbonization. The application of this technology on an industrial scale promises to reduce the carbon footprint of biochemical production and to mitigate the impact of climate change associated with industrial activities that are otherwise difficult to decarbonize.

When developing a bioprocess to supplement or replace an existing chemical product, it is crucial to conduct a techno-economic analysis. Despite the significant economic and environmental implications of developing biorefinery processes for TAL manufacture, only a few studies have examined it through techno-economic analysis (TEA) [12-13]. Based on the current retail price of \$550/kg, sourced from Biosynth (bio-synth.com), and a bulk chemical price of \$51,000/mt from vendors, all the scenarios are economically feasible in the proposed TAL production routes. By integrating carbon dioxide recycling into the current systems, biogenic carbon capture could be achieved, leading to a substantial reduction in greenhouse gas (GHG) emissions.

This study explores a novel route for TAL production by focusing on CO<sub>2</sub>-to-formic acid electrocatalysis and co-feeding formic acid with glucose in the YL process. This approach highlights the transformative potential of recycling CO<sub>2</sub>, which is traditionally viewed as a drawback in economics (options such as storage only add costs) [14], into a value-added input. This enhances the sustainability profile of TAL production, epitomizing the principles of a circular economy. Furthermore, the proposed model has the potential to shift toward more sustainable industrial practices, meet economic and environmental benchmarks, and contribute to the goal of achieving net-zero greenhouse gas emissions.

## PROCESS OVERVIEW

The focus of the present study is to conduct a thorough comparison between two YL production systems, each scaled to produce 100 metric tons of TAL annually, examining their techno-economic and life cycle aspects. The first system utilizes *Yarrowia lipolytica* fed solely with glucose, while the second system involves a simultaneous feed of YL with glucose and formic acid [10-11]. This latter system utilizes electrochemical production of formic acid from carbon dioxide [15-16].

In the present study, process simulations were executed using SuperPro Designer [17], while energy balances and heat integration analyses were computed via Aspen Plus, employing the Non-Random Two-Liquid (NRTL) thermodynamic model to ensure accurate calculations of properties [18]. The scope of the investigation extends to a LCA and TEA across two distinct system configurations as exhibited in Figure 1A and 1B, further delineated into four specific scenarios. Within these scenarios, glucose serves as the primary substrate within the YL bioreactor, facilitating the biosynthesis of bio-TAL. Further explorations, as demonstrated in Figure 1B, illustrate the particular scenario in which a co-feeding strategy of formic acid and glucose is employed, showcasing an integrated method for enhancing the bio-production efficiency of TAL. It considers the sourcing of glucose and the subsequent recycling of off-gas into the three-compartment cell (TCC) reactor [15,19]. This innovative approach integrates the reactor for the electrochemical synthesis of formic acid from carbon dioxide emissions considering three carbon recycle percentage cases: 25%, 50%, and 75%.

## Substrate for the TAL Biosynthesis Process

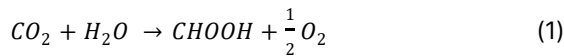
### (A) Production of TAL Using *Yarrowia lipolytica* with Glucose as Substrate

The production of TAL through biotechnological routes is a promising alternative to traditional chemical synthesis. The process starts with the metabolic assimilation of glucose by YL, which undergoes glycolysis and

enters the tricarboxylic acid (TCA) cycle and the associated biosynthetic pathways [10]. Genetic engineering has optimized *YL* to overproduce acetyl-CoA and malonyl-CoA, the precursors necessary for TAL formation [20]. By upregulating the genes involved in these pathways and downregulating those leading to competing pathways, *YL* can produce TAL.

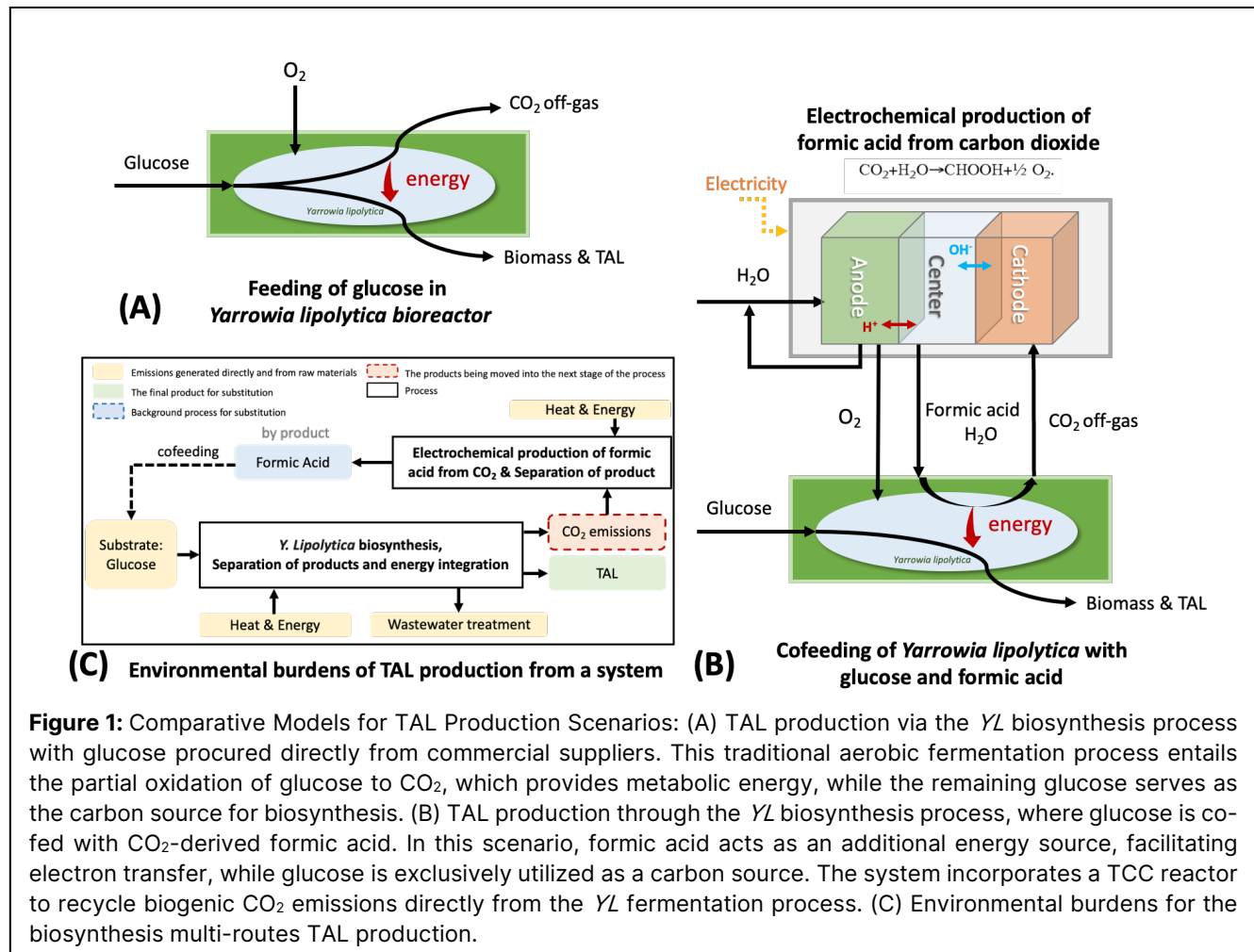
### (B) Cofeeding of *Yarrowia lipolytica* with glucose and formic acid

In the pursuit of sustainable bioproduction, a significant step forward involves capturing and utilizing CO<sub>2</sub> emissions, which is crucial for reducing the carbon footprint of biotechnological applications. According to Noorman (2020) [10-11], an innovative approach involves the electrochemical reduction of biogenic CO<sub>2</sub> using renewable electricity to form valuable organic molecules that can be fed to the fermentation process. This study uses a Three Compartment Cell (TCC) reactor to electrochemically reduce CO<sub>2</sub>, effectively converting it into formic acid with the net reaction [15,21]:



This method creates a closed carbon cycle, capturing emitted CO<sub>2</sub>, reducing it to formic acid, and reintroducing it into the fermentation stage. Theoretically, this cycle could sustain cellular energy requirements (ATP) through respiration alone. Moreover, by utilizing the primary carbon source, such as glucose, predominantly for biomass and product assimilation rather than energy production, this process substantially increases biomass and product yields, thereby enhancing the overall efficiency of the primary carbon source.

The TCC reactor's innovative configuration employs a gas diffusion electrode (GDE) cathode, GDE anode, and central flow compartments [15,19]. The central compartment is flanked by a cation exchange membrane adjacent to the anode, which is coated with an IrO<sub>2</sub> catalyst on Toray paper [15,19], and an anion exchange membrane on the cathode side, loaded with a tin (Sn) catalyst and polytetrafluoroethylene at 5% [15,19]. The key electrochemical reactions are delineated within their respective compartments. Notably, the TCC reactor design does not require the separation of by-products, given their minor concentrations within the effluent stream [21]. The efflux from the anode compartment predominantly comprises



unconverted H<sub>2</sub>O and the by-product O<sub>2</sub>, where O<sub>2</sub> will be sent to the aerobic fermentation reactor and excess water will be recycled back to the TCC reactor [15]. The formic acid flow from the center part of the TCC is sent to a pressurized 4 bar distillation tower to remove the excess water to achieve a 70 wt. % of FA and then co-fed with glucose to the *YL* bioreactor [15].

To produce 1 kg of formic acid, the process requires an input of 0.96 kg CO<sub>2</sub> and 0.57 kg H<sub>2</sub>O within the TCC reactor. The electrochemical reduction itself is powered by 4.79 kWh of electricity [15]. This co-feeding approach, where *Yarrowia lipolytica* is supplemented with both glucose and electrochemically produced formic acid, represents a shift toward more sustainable and partially decarbonized biotechnological processes.

### TAL Purification Process

Purification stands as a critical component in biorefining, with significant implications for costs. The process necessitates the removal of excess biomass, alongside organic and inorganic contaminants from the fermentation broth. Various techniques, including hexanol extraction [22], have been employed to recover TAL. However, an enhanced method utilizing charcoal adsorption has been delineated by Singh R. et al. [13], where TAL is absorbed by activated charcoal and subsequently desorbed using ethanol, which is then recovered for reuse. The selection of charcoal adsorption is predicated on its extractant reusability, providing a method that facilitates continuous operation while aligning with economic and environmental objectives, particularly leveraging the volatility of ethanol and ready availability.

### Bioprocess modeling

A bioprocess model used for this study was trained using fermentation profiles from Markham et al. [33]. A fed-batch process configuration was considered for all conditions explored. Briefly, within the system of ordinary differential equations (ODEs), the amount ATP generated from intracellular carbon catabolism at each timestep was calculated. It was then assumed that 25, 50, or 75% of the CO<sub>2</sub> emitted from the bioreactor could be captured and recycled as formic acid. An equal amount of formic acid was fed as was consumed by the organism (keeping the concentration in the bioreactor at zero). The molecular stoichiometry and energy generation from formic acid consumption defined by van Winden et al. [34] was used in the study. Energy generated from formic acid recycling was considered as a replacement for the same amount of energy generated via intracellular carbon catabolism, and the intracellular carbon that would have been used for catabolism to generate that energy was re-directed towards biomass and TAL in a ratio consistent with the carbon molar ratio of biomass and TAL produced under the glucose only feeding condition at each timepoint in the

ODE integration.

## RESULTS

### TEA results

The TEA conducted for TAL production adhered to the 2019 Chemical Engineering Equipment Cost Index and applied the discounted cash flow methodology for the determination of the Minimum Selling Price (MSP). The derived MSP was subsequently benchmarked against market prices (\$50/kg) to assess economic viability. The Internal Rate of Return (IRR) served as the economic indicator for comparison among different technologies.

Capital and operational expenditures were extrapolated from existing literature and SuperPro Designer [17], with equipment costs being estimated according to size by employing a power scaling factor, of 0.6 [23]. This calculation was based on an equipment price benchmark of a 500 m<sup>3</sup> bioreactor [35], providing a reference for cost estimations. Annual operating costs were assessed, including material, labor, utilities, and laboratory operations, along with facility expenses such as maintenance, depreciation, insurance, taxes, and overheads, with labor costs aligned with project demands. The detailed economic parameters are listed in Table 1. The TEA results, detailed in Table 2 [15,24-26], encompass capital expenditure (CapEx), operating expenditure (OpEx), the MSP of TAL, and the Internal Rate of Return (IRR) across various scenarios. Moreover, there is an additional by-product sale in the recycling scenarios examined in the TEA for TAL production. A key aspect was the conversion of all emitted CO<sub>2</sub> into FA. This process not only contributes to the sustainability of the operation but also adds a revenue stream, as the portion of FA not used for feeding back into the reactor is sold externally.

TEA analysis indicates that the scenario using glucose as the sole substrate for TAL production outperforms the co-feeding scenario with glucose and CO<sub>2</sub>-derived formic acid (FA). Notably, the co-feeding approach, with a 75% carbon recycle rate, demonstrates more than double the conversion efficiency compared to the glucose-only scenario. This enhanced efficiency is attributed to the additional NADH generated by the dissimilation of formic acid, which contributes to ATP production, thus supporting growth [11]. However, the production cost for CO<sub>2</sub>-derived FA, ranging from \$0.78 to \$2.63 per kg, exceeds that of glucose [21], priced at \$0.33 per kg on average [27].

In the co-feeding scenarios, the 25% carbon recycling case achieves the most favorable MSP at \$30.91/kg. The lower MSP in scenarios with less formic acid co-feeding correlates with the profits from by-product sales after accounting for the amounts needed for co-feeding, the surplus FA is sold as a by-product,

**Table 1:** Summary of economic assumptions for capital and operating cost.

TCC process		Capital investment cost		Fixed operating cost	
Cathode	Graphite block with Sn nanoparticles	General & Administrative Overheads	3% DFC <sup>a</sup>	Operating Labor Cost	average annual pre-taxed salary of \$80,000 per employee
Anode	Titanium anode with IrO <sub>2</sub> -based catalyst coating	Contract Fee	19% DFC	Maintenance Cost	2% of installed equipment cost
Membrane	Amberlite® IR120 strong acid ion exchange resin, 620.0 – 830.0 μm and Nafion 212/115/324	General & Administrative Overheads	3% DFC	Operating Charges	25% Operating Labor Cost
Electrolyzer cost	450 \$/kW	Working capital	20% annual operating cost	Operating General & Administrative Cost	20% direct production cost
Pressure, temperature	Ambient pressure, 25 C	Contingencies	37% DFC	<b>Time parameters</b>	
Residence time	4.0 min	Contract Fee	19% DFC	Project life	20 yr
Membrane lifetime	5000 hr	Other Capital Cost	71% DFC	Depreciation period	10 yr
				Annual operating time	8000 hr

offsetting TAL production costs). Based on the SuperPro database, with an FA market price of \$0.91 per kg, approximately 240, 448, and 648 mt/yr of FA can be sold in the 75%, 50%, and 25% carbon recycling scenarios, respectively. In these scenarios MSP is inversely correlated to TAL yield on glucose (0.30, 0.21, and 0.18 g TAL per gram of glucose for 75%, 50%, and 25% carbon recycle cases, respectively). Also worth noting, the opposite trend in TAL MSP is observed if we only produce the required amount of formic acid and release the remainder of the CO<sub>2</sub> to atmosphere (at \$33.6, \$31.74, and \$31.51 per kilogram, for 25, 50, and 75% recycle scenarios, respectively).

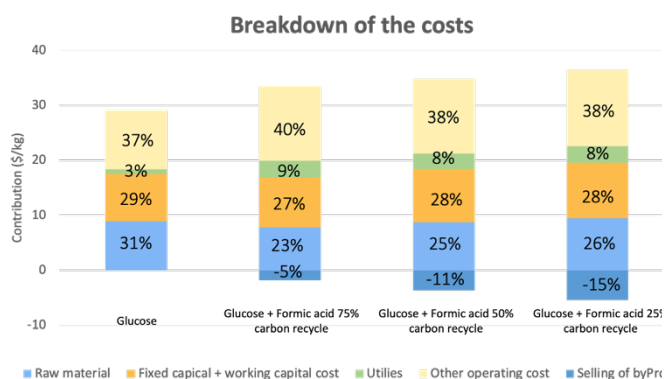
Figure 2 illustrates the detailed cost components contributing to the MSP for each production scenario. The primary utilities impacting the process economics are identified as electricity, steam, process water, and cooling water. The energy-efficient TAL separation process, notably the adsorption and desorption steps, is highlighted for its role in reducing utility costs. However, for the electrochemical production of FA, the requirement of 4.79 kWh of electricity per kilogram of FA [15] results in a higher proportion of utility costs compared to glucose-based TAL production.

**Table 2:** Summary of TEA for various raw materials for a benchmark plant capacity of 100 mt TAL/y.

Stream	CapEx (MM\$)	OpEx (MM\$)	MSP (\$/kg) <sup>a</sup>	IRR (%) <sup>b</sup>
Glucose	4.88	2.11	29.10	51.14
Glucose + Formic acid 75% carbon recycle	5.10	2.47	31.37	45.86
Glucose + Formic acid 50% carbon recycle	5.44	2.58	31.16	45.26
Glucose + Formic acid 25% carbon recycle	5.67	2.69	30.91	44.71

<sup>a</sup> MSP is computed when the sum of the discounted cash flow over 10 years equals zero (minimum selling price without any benefit).

<sup>b</sup> The internal rate of return after twenty years is based on the glucose-based MSP at a scale 0.1 kt/y, 50.00 \$/kg.



**Figure 2.** Cost breakdown of applying different TAL production routes.

The larger FA requirement to feed in the 75%, 50%, and 25% carbon recycle cases, with amounts of 95.95, 58.42, and 28.03 kg/hr, respectively causes an increase in electricity demands for the TCC reactor. However, TAL production using 75% carbon recycling of FA incurs less utility costs in the YL sections, as the cooling water demands of the YL bioreactor decrease, assuming that the reaction expends 115 kcal of heat for each mole of O<sub>2</sub> consumed [32]. To produce an equivalent quantity of TAL, the oxygen mole ratio required for the 75%, 50%, and 25% FA scenarios are 13.67, 15.13, and 16.59, respectively. The trade-off in utility costs between the two sections indicates that the CO<sub>2</sub>-to-FA process is more energy-intensive. When using only glucose as a substrate in the YL bioreactor, a higher capital expenditure is required due to the lower yield and productivity of TAL (larger reactor size and more fermentation broth); they are 0.40, 0.66, 0.52, and 0.45 g-O<sub>2</sub>/kg broth/h for the glucose-based scenario and FA co-feeding with glucose with 75, 50, and 25 % carbon recycle scenarios respectively. However, when FA and glucose are co-fed, yield and productivity increase. For the 75%, 50%, and 25% carbon recycle cases, they are 0.30, 0.21, and 0.18 g TAL/g glucose, that is 0.19, 0.13, and 0.10 g-TAL/kg broth/h in productivity, respectively. In these scenarios, the increase in biomass yield not only enhances product titer but also allows the system to burn less glucose for ATP generation. This efficiency results in smaller reactor sizes, with the 75% carbon recycle case being 44% smaller than the 25% case, and less fermentation broth, further improving the overall productivity (0.19 and 0.10 g-TAL/kg broth/h respectively). Moreover, in the 25% carbon recycling scenario, capital costs for the CO<sub>2</sub>-derived FA segment are lower due to the reduced CO<sub>2</sub> output from the YL section.

## LCA results

The life cycle assessment (LCA) aimed to evaluate the environmental advantages of electrochemically converting CO<sub>2</sub> to FA and the subsequent co-feeding of

glucose and FA for TAL production. The assessment, considering cradle-to-gate scope, tracked the CO<sub>2</sub> from its origin in the YL bioprocess through its transformation into FA and its return to the YL bioprocess in combination with glucose. The functional unit for this analysis was set at 1 kg of TAL.

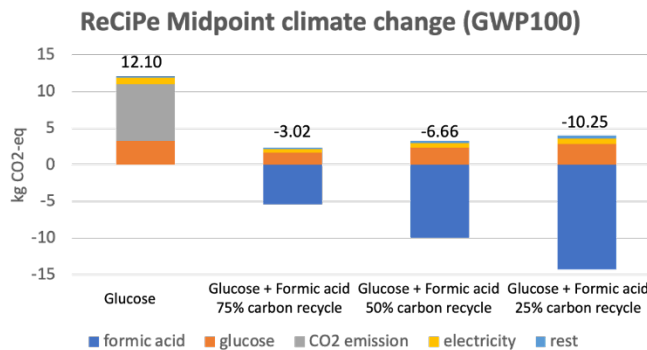
Conforming to the ISO 14040 standard [28], the environmental evaluation of the CO<sub>2</sub>-to-FA reduction was conducted using the Brightway2 framework [29] and the Ecoinvent® v3.9.1 database [30], prioritizing US data but using global data when US data was unavailable. To determine carbon footprint and LCA, emissions factors were created and used for various activities, such as grid electricity usage, measured in g CO<sub>2</sub>/kWh consumed. These factors represent an estimate of emissions from a broad system of power generators, rather than a single point source of emissions. The ReCiPe Midpoint method (version 1.13) was employed to gauge the Global Warming Potential (GWP) of greenhouse gases over a century [31], thus ensuring a comprehensive perspective of both immediate and enduring environmental effects.

Furthermore, the LCA utilized system expansion for product allocation within multi-output processes, a method advantageous for co-products with distinct, quantifiable market values. This approach allows for the apportionment of environmental burdens proportionately to the economic value or utility of each product. In this case, the co-produced FA, derived from recycled CO<sub>2</sub>, is considered an environmental credit, acknowledging its displacement of fossil-derived equivalents, as depicted by the blue squares in Figure 1(C). Oxygen, another by-product of electrolysis, was excluded from further utility considerations in this assessment.

The LCA results demonstrate that co-feeding *Yarrowia lipolytica* with glucose and formic acid significantly improves the greenhouse gas performance of TAL production. Notably, the scenarios with 75%, 50%, and 25% carbon recycling exhibit negative net emissions, thereby enhancing the environmental profile of the process. The 25% carbon recycling case is the most effective, with a net emission reduction to -10.25 kg CO<sub>2</sub>-equivalents per kg of TAL (4.03 CO<sub>2</sub>-eq per kg of TAL, if not considering the CO<sub>2</sub>-to-FA credit), indicating an exemplary case of carbon-negative output. While if not considering the carbon credit from the by-product, FA, the 75% carbon recycling case performs the best in GWP impact. The inclusion of recycled CO<sub>2</sub> as a feedstock for formic acid production contributes to this improvement, showcasing a viable pathway for reducing GHG emissions in chemical manufacturing. On the other hand, the production of TAL using glucose-based methods, without recycling the CO<sub>2</sub> generated during fermentation, results in a higher carbon footprint. In this scenario, the carbon emissions from the fermentation process account for more than 50% of the total net emissions. This assessment solidifies the co-



feeding of glucose with formic acid, particularly with high rates of carbon recycling, as a technologically and environmentally advantageous route for TAL production, aligning with global sustainability goals.



**Figure 3.** Lifecycle greenhouse gas emissions for glucose-based TAL and cofeeding of *Yarrowia lipolytica* with glucose and formic acid-based TAL.

## DISCUSSION

The transition to sustainable energy systems is a critical endeavor in mitigating the environmental impact of hard-to-decarbonize sectors, including industrial and chemical processes. This study's LCA and TEA have shed light on a novel bioprocessing route for TAL production that integrates carbon recycling—a pivotal step towards achieving net-zero emissions.

The LCA results affirm that the co-feeding of *Yarrowia lipolytica* with glucose and formic acid significantly advances the GHG performance of TAL production. The standout 75% carbon recycling scenario not only meets but exceeds net-zero emission goals, venturing into the carbon-negative territory. This demonstrates the profound environmental benefits of incorporating CO<sub>2</sub>-to-formic acid conversion within the TAL production cycle.

From a techno-economic perspective, the TEA findings reveal that the co-feeding approach, particularly with a high rate of carbon recycling, is economically viable. Despite the higher utility costs associated with the electrochemical generation of formic acid, this is counterbalanced by the revenue generated from the sale of excess formic acid, thereby enhancing the process's cost-competitiveness.

The capital cost analyses for the bioreactor configurations further corroborate the economic viability of the co-feeding approach. Although the initial outlay is more substantial for the glucose-only scenario due to the larger reactor requirements (4.88 MM\$ CapEx of YL process), co-feeding with formic acid benefit the reactor sizing and utility consumption, leading to a reduction in both capital and operating expenses (3.21–4.66 MM\$ CapEx of YL process).

Incorporating these findings, future directions will

involve refining the bioprocesses and scaling the production while aligning with policy developments that support economic and environmental sustainability. This research sets a foundation for policy-backed industry practices that promote a circular economy and contribute to the global goal of a net-zero carbon future.

## ACKNOWLEDGEMENTS

This work is financially supported by DSM-Firmenich.

## REFERENCES

- Lyle, John Tillman. Regenerative design for sustainable development. John Wiley & Sons, 1996.
- Obydenov, Dmitrii L., Asmaa I. El-Tantawy, and Vyacheslav Ya Sosnovskikh. "Triacetic acid lactone as a bioprivileged molecule in organic synthesis." *Mendeleev Communications* 29.1 (2019): 1-10.
- Cardenas, Javier, and Nancy A. Da Silva. "Metabolic engineering of *Saccharomyces cerevisiae* for the production of triacetic acid lactone." *Metabolic engineering* 25 (2014): 194-203.
- Weissermel, Klaus, and Hans-Jürgen Arpe. *Industrial organic chemistry*. John Wiley & Sons, 2008.
- Jach, Monika Elżbieta, and Anna Malm. "Yarrowia lipolytica as an alternative and valuable source of nutritional and bioactive compounds for humans." *Molecules* 27.7 (2022): 2300.
- Park, Young-Kyoung, and Jean-Marc Nicaud. "Metabolic engineering for unusual lipid production in *Yarrowia lipolytica*." *Microorganisms* 8.12 (2020): 1937.
- Macedo, Isaias C., Joaquim EA Seabra, and João EAR Silva. "Green house gases emissions in the production and use of ethanol from sugarcane in Brazil: the 2005/2006 averages and a prediction for 2020." *Biomass and bioenergy* 32.7 (2008): 582-595.
- Najafpour, Ghasem. *Biochemical engineering and biotechnology*. Elsevier, 2015.
- Anwar, M. N., et al. "CO<sub>2</sub> utilization: Turning greenhouse gas into fuels and valuable products." *Journal of environmental management* 260 (2020): 110059.
- van Winden, Wouter A., et al. "Towards closed carbon loop fermentations: Cofeeding of *Yarrowia lipolytica* with glucose and formic acid." *Biotechnology and Bioengineering* 119.8 (2022): 2142-2151.
- Noorman, H. J. (2019). Greenhouse gas improved

fermentation (EP3715464B1). DSM IP Assets BV.  
URL (<https://patents.google.com/patent/EP3715464B1/en>)

12. Chia, Mei, et al. "Triacetic acid lactone as a potential biorenewable platform chemical." *Green Chemistry* 14.7 (2012): 1850-1853.
13. Singh, Ramkrishna, et al. "Adsorptive separation and recovery of triacetic acid lactone from fermentation broth." *Biofuels, Bioproducts and Biorefining* 17.1 (2023): 109-120.
14. Quadrelli, Elsje Alessandra, et al. "Carbon dioxide recycling: emerging large-scale technologies with industrial potential." *ChemSusChem* 4.9 (2011): 1194-1215.
15. Kang, Dongseong, Jaewon Byun, and Jeehoon Han. "Electrochemical production of formic acid from carbon dioxide: A life cycle assessment study." *Journal of Environmental Chemical Engineering* 9.5 (2021): 106130.
16. Leitner, Walter. "Carbon dioxide as a raw material: the synthesis of formic acid and its derivatives from CO<sub>2</sub>." *Angewandte Chemie International Edition in English* 34.20 (1995): 2207-2221.
17. Canizales, Licelly, et al. "SuperPro Designer®, user-oriented software used for analyzing the techno-economic feasibility of electrical energy generation from sugarcane vinasse in Colombia." *Processes* 8.9 (2020): 1180.
18. Renon, Henri, and John M. Prausnitz. "Local compositions in thermodynamic excess functions for liquid mixtures." *AIChE Journal* 14.1 (1968): 135-144.
19. Yang, Hongzhou, et al. "Electrochemical conversion of CO<sub>2</sub> to formic acid utilizing Sustainion™ membranes." *Journal of CO<sub>2</sub> Utilization* 20 (2017): 208-217.
20. Liu, Yangming, et al. "Recent Advances and Perspectives on the Biomass-derived Production of the Platform Chemical Triacetic Acid Lactone by Engineered Cell Factories." *Biochemical Engineering Journal* (2023): 108961.
21. Somoza-Tornos, Ana, et al. "Process modeling, techno-economic assessment, and life cycle assessment of the electrochemical reduction of CO<sub>2</sub>: a review." *Iscience* 24.7 (2021).
22. Yu, James, et al. "Bioengineering triacetic acid lactone production in *Yarrowia lipolytica* for pogostone synthesis." *Biotechnology and Bioengineering* 115.9 (2018): 2383-2388.
23. Towler, Gavin, and Ray Sinnott. *Chemical engineering design: principles, practice and economics of plant and process design*. Butterworth-Heinemann, 2021.
24. Davis, Ryan, Andy Aden, and Philip T. Pienkos. "Techno-economic analysis of autotrophic microalgae for fuel production." *Applied Energy* 88.10 (2011): 3524-3531.
25. Eswaran, Sudha, et al. "Techno-economic analysis of catalytic hydrothermolysis pathway for jet fuel production." *Renewable and Sustainable Energy Reviews* 151 (2021): 111516.
26. Mousavi-Avval, Seyed Hashem, and Ajay Shah. "Techno-economic analysis of hydroprocessed renewable jet fuel production from pennycress oilseed." *Renewable and Sustainable Energy Reviews* 149 (2021): 111340.
27. Cheng, Ming-Hsun, et al. "The costs of sugar production from different feedstocks and processing technologies." *Biofuels, Bioproducts and Biorefining* 13.3 (2019): 723-739.
28. Finkbeiner, Matthias, et al. "The new international standards for life cycle assessment: ISO 14040 and ISO 14044." *The international journal of life cycle assessment* 11 (2006): 80-85.
29. Mutel, Chris. "Brightway: an open source framework for life cycle assessment." *Journal of Open Source Software* 2.12 (2017): 236.
30. FitzGerald, D., and T. Sonderegger. "Documentation of changes implemented in the ecoinvent database v3. 9.1." (2022).
31. Pfister, Stephan, and Laura Scherer. "Uncertainty analysis of the environmental sustainability of biofuels." *Energy, Sustainability and Society* 5.1 (2015): 1-12.
32. Bailey, James E., and David F. Ollis. *Biochemical engineering fundamentals*. McGraw-Hill, 2018.
33. Markham, Kelly A., et al. "Rewiring *Yarrowia lipolytica* toward triacetic acid lactone for materials generation." *Proceedings of the National Academy of Sciences* 115.9 (2018): 2096-2101.
34. van Winden, Wouter A., et al. "Towards closed carbon loop fermentations: Cofeeding of *Yarrowia lipolytica* with glucose and formic acid." *Biotechnology and Bioengineering* 119.8 (2022): 2142-2151.
35. Humbird, D., R. Davis, and J. D. McMillan. "Aeration costs in stirred-tank and bubble column bioreactors." *Biochemical engineering journal* 127 (2017): 161-166.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# An Update on Project PARETO - New Capabilities in DOE's Produced Water Optimization Framework

Miguel A. Zamarripa<sup>a,b,\*</sup>, Elmira Shamlou<sup>a,b</sup>, Javal Vyas<sup>a,b</sup>, Travis Arnold<sup>a,b</sup>, Philip Tominac<sup>a,b</sup>, Melody H. Shellman<sup>a,b</sup>, and Markus Drouven<sup>a</sup>

<sup>a</sup> National Energy Technology Laboratory, Pittsburgh, PA, USA

<sup>b</sup> NETL Support Contractor, Pittsburgh, PA, USA

\* Corresponding Author: [Miguel.Zamarripa-Perez@netl.doe.gov](mailto:Miguel.Zamarripa-Perez@netl.doe.gov)

## ABSTRACT

Managing oil and gas produced water, characterized by hypersalinity and large volumes, presents significant challenges. This paper introduces an advanced optimization framework, PARETO, which offers a novel approach to strategic water management, emphasizing produced water (PW) treatment, quality tracking, quantification of emissions, and environmental justice. This work presents a case study showcasing different produced water management challenges. The PARETO framework demonstrated its effectiveness in optimizing water management strategies in line with environmental sustainability and regulatory compliance.

**Keywords:** produced water management, MINLP, process design, network optimization, MILP

## INTRODUCTION

Produced water (PW) management presents significant challenges to oil and gas industry stakeholders due to variable production volumes (unpredictable handling requirements) and water quality, especially due to its high concentrations of total dissolved solids (TDS) and other constituents (necessitating treatment). Given the variability in water produced across different basins along with the difficulty of treating hypersaline brine, oil and gas companies need to identify fit-for-purpose produced water management, treatment, and reuse approaches. Projected increases in PW volumes in coming decades [1], recent injection capacity curtailments [2 - 3], and intensive capital investments (e.g., PW infrastructure) all motivate novel, cost-effective strategies for PW management. Therefore, decision-support tools to assess techno-economic feasibility are critical. However, few such software tools currently exist.

In 2021, the US Department of Energy (DOE) launched a three-year, \$5 million PW optimization initiative to develop, demonstrate and deploy PARETO, a novel optimization framework for PW management and beneficial reuse. PARETO is developed by the National Energy Technology Laboratory (NETL), in cooperation with Lawrence Berkeley National Laboratory (LBNL),

Carnegie Mellon University, Georgia Tech, New Mexico State University, and the Ground Water Protection Council, and is designed to identify cost-effective and environmentally sustainable PW management, treatment and reuse solutions [4]. Specifically, PARETO supports decision-makers with 1) PW management, including infrastructure buildout recommendations and the coordination of PW deliveries; 2) PW treatment, including treatment facility placement recommendations and the selection of effective treatment technologies; and 3) PW beneficial reuse, including the identification of beneficial reuse options and the distribution of treated PW and/or concentrated brine.

In previous versions of PARETO, the tool's core capabilities for optimizing water distribution across the network were demonstrated using the Mixed Integer Linear Programming (MILP) method [4]. These capabilities included user-defined treatment site specifications and post-process water quality tracking. The optimization aimed at various objectives, such as minimizing costs or maximizing water recycling within the network. Building upon this work, the current version of PARETO introduces significant enhancements, including new methods for integrating produced water (PW) treatment and quality tracking. This is achieved through surrogate modeling of rigorous desalination models, which capture the impact

of water quality and flow rate on treatment costs, utilizing both MILP and Mixed Integer Nonlinear Programming (MINLP) methods. Furthermore, this update includes advanced techniques for quantifying emissions and promoting environmental justice, enabling a detailed analysis of the trade-offs between economic factors and environmental objectives. 1) the PARETO treatment module includes a library of PW desalination technologies that can be used to obtain treated water for use outside oil and gas operations; 2) the PARETO environmental module allows a user to track pollutants generated during produced water operations and provides a tool to bring environmental justice to communities located near oil and gas operations.

The results show PARETO’s advanced mathematical modeling capabilities are able to solve large scale water network problems, with an emphasis on complex nonlinearities arising from studying effective formulations for solving water networks with water quality predictions.

## CASE STUDY AND PROBLEM STATEMENT

This work presents an industrial-size produced water network based on the Permian Basin (New Mexico and Texas).

The main characteristics of PW networks in this region are the availability of pipelines and limited use of trucking for transport (the opposite is true in some other basins such as the Appalachian). Figure 1 shows a representative PW network, which consists of 14 production pads, 3 completion pads, 5 disposal wells, 4 treatment sites (one of which is for desalination – R03), and 3 treatment technologies (i.e., membrane vapor compression, membrane distillation, and osmotic assisted reverse osmosis). Figures 2-4 show the time-varying nature of the water production forecast and completions demand. The PW forecasts follow an exponential decay pattern, whereas the completions pads demand large amounts of water during specific time windows. Table 1 summarizes the water volumes, as can be seen, the PW forecast exceeds the completions demand, which is typical in PW networks.

The main challenges associated with PW network management are:

- Handling large volumes of produced water during specific time windows. (see Figure 3 and 4)
- High salinity (TDS > 120,000 mg/L) and other constituents (i.e., oil, grease, barium, lithium, etc.).
- Pipeline hydraulics (i.e., pressure drop, pumping stations, etc.).
- Active development in the area requiring infrastructure buildout (i.e., new pipelines, new

disposal wells, new treatment plants).

- Selecting appropriate treatment and beneficial reuse alternatives for produced water (if available/applicable).

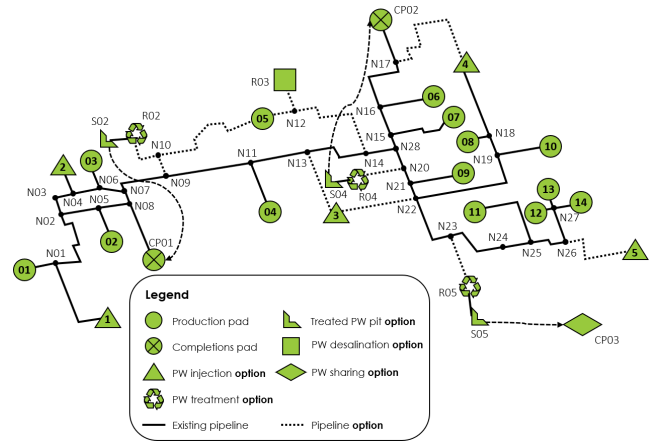


Figure 1: PW network schematic.

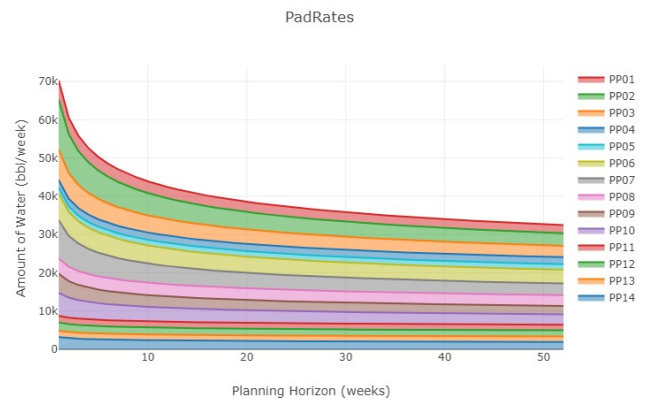


Figure 2: PW forecasts from production pads.

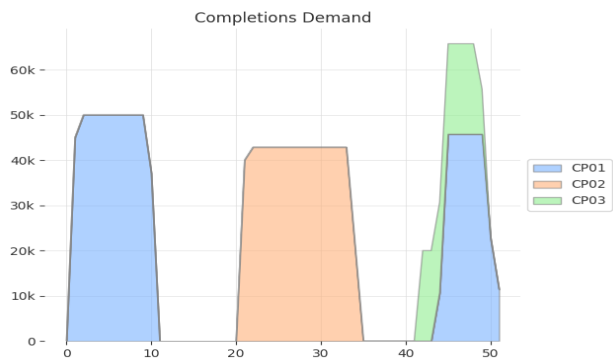
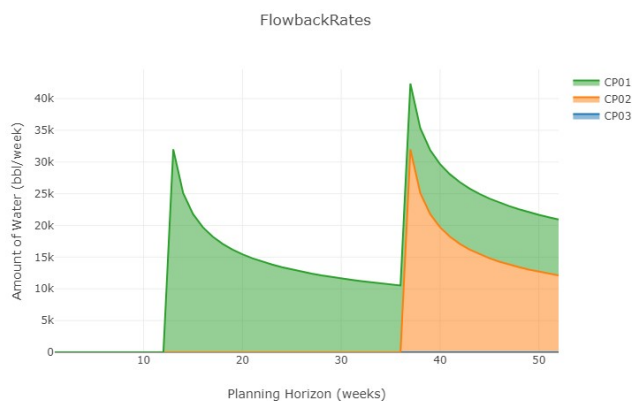


Figure 3: Completions demand volumes.



**Figure 4:** Flowback forecast from completions pads.

**Table 1:** Total PW volume and completion demand.

Item	Volume (bbl/day)
Completions Demand	1,481,429
Produced Water Forecast	2,825,953
Disposal Capacity	2,525,714
Starting Treatment Capacity	0

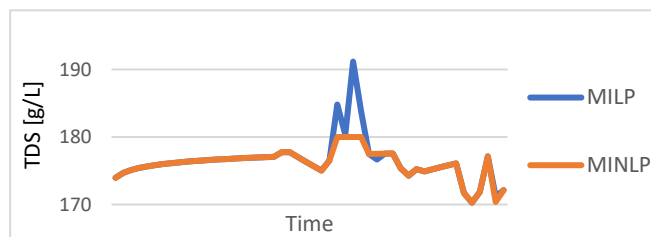
## PARETO PRODUCED WATER QUALITY MODELING APPROACH

In produced water network management, a key element involves the estimation of flowrates and water qualities. This process involves tracking various components within the network to efficiently manage and reuse water, tailored to the quality requirements of specific end uses. A notable consideration in modeling these networks is the non-convex and bilinear term in the mass balances emerging from the product of flow rates and concentrations of various components. Within its comprehensive management framework, PARETO adopts two distinct methods to estimate water quality. The main goals are to allow users to: 1) identify component/quality peaks within the produced water network; 2) simulate operating constraints based on water quality limitations; 3) enable the selection and sizing of treatment technologies to enable beneficial reuse.

**Quality Post-Processing:** In this case, PARETO solves an MILP model focusing on binary decisions and flow rates, without considering water quality. Once these flow rates are estimated, quality constraints are activated and flows and binary variables serve as parameters for subsequent quality assessment, simplifying these calculations to linear constraints. This approach is particularly valuable for pinpointing specific areas in the network that require targeted treatment and for identifying the most suitable technologies for the reuse of produced water (see Figure 5).

**Predictive Quality Management:** This more

advanced method begins with optimizing flow rates using an MILP, followed by the explicit integration of quality constraints. Subsequently, the system is initialized for a comprehensive MINLP model. While this method adeptly sets the stage for detailed analysis, the primary challenge emerges from the intricacies of the bilinear terms, particularly their nonconvex nature. This complexity hampers the achievement of a global optimum, highlighting the advanced analytical challenges inherent in accurately modeling and optimizing such systems. This approach is valuable for adding performance constraints, and/or to identify the best treatment technology for a given network. Figure 5 presents a case study demonstrating the application of an MILP model without quality level restrictions, resulting in elevated quality levels at specific locations. However, by integrating an MINLP model with explicit quality level constraints, it becomes feasible to precisely regulate and monitor the water flow and quality in these areas, ensuring that the quality does not exceed a specified threshold.



**Figure 5.** The TDS (g/L) level at a specific location in the PW network.

With both these methodologies in play, a pivotal aspect is the role of water treatment centers. The efficiency of these centers in component removal and water recovery is integral to determining the final quality and quantity of treated and residual streams. The PARETO framework provides flexibility by allowing users to define specific water recovery and removal efficiency targets for each treatment site and technology.

## PARETO Treatment and Desalination

Treatment systems are critical in achieving the required water quality for diverse applications such as beneficial reuse and critical minerals recovery. The costs linked to these treatment systems, varying with the purpose and intensity of treatment, play a substantial role in determining the overall investment in PW management strategies. It is, therefore, imperative to give due consideration to the cost and functionality of treatment models within the broader context of produced water management strategies.

**PARETO advanced water treatment analysis:** PARETO provides three methodologies (high-level cost analysis, semi-rigorous, and rigorous cost analysis):



- 1) discrete numerical inputs to enable flexible sizing of treatment plants (PARETO documentation includes a detailed literature review and vendor survey of treatment CAPEX and OPEX: [https://pareto.readthedocs.io/en/latest/model\\_library/water\\_treatment/index.html](https://pareto.readthedocs.io/en/latest/model_library/water_treatment/index.html)).
- 2) the deployment of detailed rigorous models for comprehensive process design and network optimization.
- 3) the application of surrogate models for streamlined computational analysis.

Notably, while discrete numerical inputs facilitate the estimation of treatment costs, they fall short in capturing the intricate interplay between water quality and cost implications. To bridge this gap, this work presents a PARETO surrogate modeling approach. These surrogate models adeptly establish correlations between water quality, plant capacity (measured as flow rate), and cost, thereby furnishing a more rigorous cost analysis framework.

PARETO effectively leverages the WaterTAP [5] library's rigorous models, which include a selection of advanced desalination technologies. Among these, Osmotically Assisted Reverse Osmosis (OARO), Mechanical Vapor Compression (MVC), and Membrane Distillation (MD) are prominent for their applicability in treating hypersaline produced water. The presented case study includes these technologies alongside primary treatment processes to produce both purified water and clean brine for beneficial reuse and network recycling, respectively.

The methodologies utilized in PARETO, as previously detailed, provide a platform for diverse approaches in water treatment analysis. To showcase and compare these capabilities, we evaluate three distinct cases, each reflecting a different strategy within the PARETO framework:

**Case 1: Discrete Input Values:** In this approach, the PARETO framework is applied to analyze a system with discrete input values, where cost data for discrete expansion sizes of the plants are provided. The model uses this data to evaluate and identify the most cost-efficient desalination method among Mechanical Vapor Compression (MVC), Membrane Distillation (MD), and Osmotic Assisted Reverse Osmosis (OARO).

**Case 2: MILP-NN Surrogate:** This approach involves the use of the MILP integrated with a surrogate neural network (NN). Operating under fixed feed quality assumptions typical of Permian produced water with 128,000 mg/L Total Dissolved Solids (TDS), the model allows for variability in feed flow rate, demonstrating the effectiveness of MILP-NN in systems with fixed quality and variable flow rates.

**Case 3: MINLP-NN Surrogate:** This case employs an MINLP model, incorporating a surrogate for the MVC

process. It enables analysis under varying inlet feed quality and flow rate, providing a comprehensive view of the system's performance under different conditions.

## Surrogate Models

Surrogate models are essential for simplifying complex processes in situations where computational limitations (i.e., large scale network problems such as produced water networks), lack of algebraic representation, or external functions/constraints are present.

In our study, we address the complexities of Mechanical Vapor Compression (MVC) produced water treatment technology, which features a non-linear, non-convex nature, making it unsuitable for direct representation in the large-scale Mixed-Integer Linear Programming (MILP) framework. To manage the substantial increase in problem size from the MVC model's 54 constraints/variables (per time period), we employ surrogate models based on simulations and empirical data, focusing on regression techniques.

Machine learning surrogate models, particularly Neural Networks with Rectified Linear Unit (ReLU) activation functions, are integrated into PARETO strategic model optimization frameworks. These models are applied to MVC plant scenarios, focusing on inputs like water quality and recovery, and outputs including CAPEX, OPEX, and energy consumption.

This work leverages the IDAES-PSE [6] machine learning toolset to train surrogate models for the treatment technologies and integrate the surrogate model within the PW network problem to determine the optimal selection of treatment technologies (case 2 and 3 mentioned above).

## PW Quality and Treatment Results

The results presented in Table 2 provide a comprehensive overview of the model's performance across the various cases.

The overarching goal here is not to compare these cases on absolute numbers but rather to demonstrate the PARETO framework's flexibility and adaptability. The basis of calculations and underlying assumptions varies for each method, reflecting the diverse scenarios and user requirements each case study aims to address. This framework is designed to cater to a range of user needs, offering tailored solutions for different operational conditions and types of analysis.

An important observation across all cases is the model's effectiveness in optimizing water management, particularly in its emphasis on recycling water for completions purposes. The data shows that a significant portion of produced water, approximately 45%, is recycled, which substantially reduces reliance on freshwater sources (only about 4%).

**Table 2:** PW quality and treatment results. <sup>a</sup> kbbl

	Case 1	Case 2	Case 3
Total Cost, k\$	18,603	18,470	18,306
Sourced Water <sup>a</sup>	515	515	515
Disposal Volume <sup>a</sup>	7,218	7,000	7,100
Reuse Volume <sup>a</sup>	8,805	8,805	8,805
Piping OPEX, k\$	15	14	13
Disposal CAPEX, k\$	142	140	140
Pipeline CAPEX, k\$	111	113	127
R03 treatment tech.	MVC	-	-
R03 inlet salinity, g/l	134	128	130
Run time, seconds	272	400	1600
Gap, %	0	0	13

As shown in Table 2, the model suggests investing in pipeline, treatment, and disposal infrastructure as an optimal long-term solution. This recommendation is based on a detailed cost-benefit analysis, indicating that such investments, despite their initial capital requirements, are beneficial in the long run for sustainable and efficient water management.

**Case 1: Basic Approach with Discrete Input Values:** Case Study 1 represents a fundamental yet effective approach, utilizing discrete input values for treatment costs and plant sizes. This method simplifies the analysis by not tracking water quality within the network, focusing instead on optimizing cost and size parameters. One of the significant advantages of this approach is its reduced computational intensity. As indicated in the results, this method can solve the network in just 272 seconds, demonstrating its efficiency and suitability for quick assessments or preliminary planning phases.

Another notable aspect of Case 1 is its suitability for scenarios where vendor or industrial data is available in a discrete format. This approach aligns well with situations where the costs and technologies for water treatment are not heavily dependent on the inlet water quality. It is particularly beneficial in cases where the variation in water quality across the network is minimal, allowing for a more straightforward optimization process without the need for intricate quality tracking mechanisms.

**Case 2 Advanced Approach with Surrogate Model Integration:** Case 2 incorporates a neural network based surrogate model that is responsive to both inlet flowrate and inlet water quality. To maintain the MILP structure while capturing the effects of varying costs associated with plant capacity and treatment, the water quality in this model is assumed to be a fixed value, representative of the conditions in the Permian Basin. A key attribute of Case Study 2 is its ability to strike a balance between computational efficiency and the adaptability of the model (reaching zero gap after 400s).

One of the most notable features of Case 2 is its capacity to effectively capture the nuances of centralized

versus decentralized treatment systems. By providing a finer-scale comparison that takes into account varying treatment cost and plant sizes in various location, this model delivers insights crucial for the consideration of transportation costs associated with water movement to the treatment centers.

Further enhancing the model's capabilities, Case 2 can incorporate an iterative procedure yielding to a dynamic and adaptable analysis by leveraging the post-processing approach (i.e., estimate cost/performance, update quality levels, and re-estimate cost/performance, until convergence is reached).

**Case 3 - Advanced Modeling with MINLP and Quality Tracking:** Case Study 3 employs an advanced MINLP model that intricately incorporates aspects of water quality, flowrate, and recovery. The scope of Case 3 extends beyond mere operational optimization. It is specifically designed to not only address environmental and regulatory considerations, but also efficiently predict treated water, recovery, and concentrated water. Which is critical to enable beneficial reuse (i.e., recovery of critical minerals from produced waters).

While Case Study 3 offers the most detailed and sophisticated analysis among the three models, it is also the most computationally demanding. As evidenced by the results, this model requires a significant amount of time to reach an acceptable optimality gap. This computational intensity reflects the model's complexity and the depth of analysis it provides. Users of Case Study 3 need to be cognizant of the trade-off between its comprehensive analytical capabilities and the time and computational resources required.

## QUANTIFYING EMISSIONS AND ENVIRONMENTAL JUSTICE IN PRODUCED WATER NETWORKS

Accurate estimation of emissions is not only critical for maintaining regulatory compliance but also for the economic considerations of PW systems, especially with the introduction of new emissions regulations. The Inflation Reduction Act, for example, established the Methane Emissions Reduction Program which introduces a charge for reported waste emissions beginning in 2024 [7]. While emissions measurements are generally acknowledged, there is also a growing interest in environmental justice impact. In 2021, the Justice40 initiatives in response to Executive Order 14008 outlined new guidance for environmental justice, including specific recommendation for the decrease of environmental exposure and burdens for Disadvantaged Communities. [8].

The PARETO framework includes three categories of environmental impact and environmental justice measures: (1) air pollutant metrics, (2) environmental exposure in disadvantaged communities, and (3) trucking

activity. Air emissions are tracked at different sources throughout the system: trucking, pipeline operations, pipeline installation, disposal, storage, and treatment. Five measures of air pollutants were identified for this case study: Greenhouse gas emissions (CO<sub>2</sub> equivalents), NH<sub>3</sub>, NO<sub>x</sub>, SO<sub>2</sub>, and PM<sub>2.5</sub>. Table 2 presents the respective coefficients. PARETO reports on air emissions at each source and across the full PW network for each air pollutant type. Total emissions are also calculated as the sum across all pollutant types. To incorporate emissions metrics into PARETO decision-making, users can select to optimize the PW network with an objective to minimize total emissions.

**Table 3:** Air emission coefficients.

Source	Unit	Type	Coefficient
Trucking	g/hour	CO <sub>2</sub>	2,035,000
		NH <sub>3</sub>	35
		NO <sub>x</sub>	12650
		SO <sub>2</sub>	770
		PM <sub>2.5</sub>	121
Pipeline Operations	g/bbl-mile	CO <sub>2</sub>	22
		NH <sub>3</sub>	0.0002
		NO <sub>x</sub>	0.014
		SO <sub>2</sub>	0.0071
		PM <sub>2.5</sub>	0.00086
Pipeline Installation	g/mile	CO <sub>2</sub>	310,000,000
		NH <sub>3</sub>	2,800
		NO <sub>x</sub>	190,000
		SO <sub>2</sub>	100,000
		PM <sub>2.5</sub>	12,000
Disposal	g/bbl	CO <sub>2</sub>	970
		NH <sub>3</sub>	0.13
		NO <sub>x</sub>	9.5
		SO <sub>2</sub>	2.8
		PM <sub>2.5</sub>	0.87
Storage	g/bbl-week	CO <sub>2</sub>	1,000
		NH <sub>3</sub>	0.012
		NO <sub>x</sub>	2
		SO <sub>2</sub>	0.65
		PM <sub>2.5</sub>	0.076

Table 3 and Table 4 present coefficients used to estimate air emissions output for unit of time, volume, distance, or combination, depending on the source. Air emissions coefficients and mathematical constraints are based on environmental impact modeling in Bartholomew & Mauter [7]. Ongoing work for emissions measurement includes establishing coefficients for technologies

beyond Mechanical Vapor Compression.

**Table 4:** Treatment technology air emission coefficients.

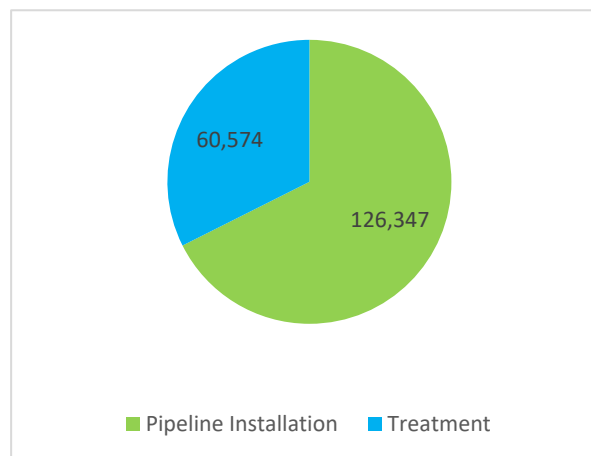
Treatment Technology	Type	Coefficient (g/week)
Mechanical Vapor Compression	CO <sub>2</sub>	9600
	NH <sub>3</sub>	1.45
	NO <sub>x</sub>	12.8
	SO <sub>2</sub>	13.5
	PM <sub>2.5</sub>	1.5

The second category included in the PARETO framework addresses environmental justice and focuses on Disadvantaged Communities (DAC) as defined by the Climate and Economic Justice Screening Tool [9]. This metric is a function of air emission measures and reports the air pollutants contributed from sources in the produced water system that fall within a DAC. Ongoing work for environmental justice measures includes incorporating penalties for building new PW infrastructure into an environmental objective function.

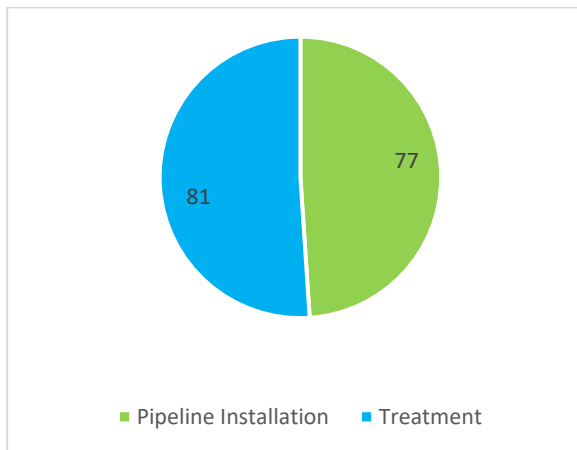
The third category reports trucking activity for the total volume of water trucked and total hours of trucking time with non-zero water load.

### Environmental Impact Assessment

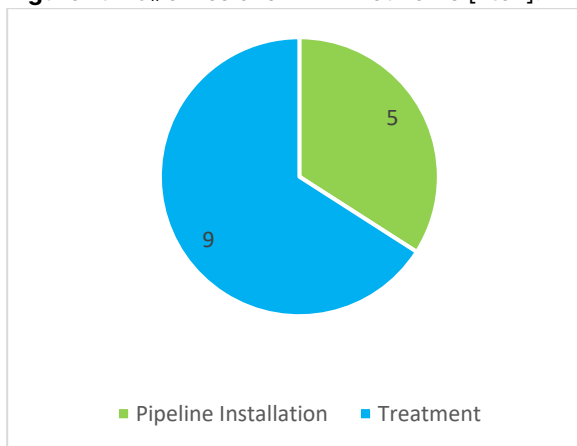
The above-mentioned metrics have been added to the PARETO framework. Figures 6 to 8 present detailed emissions for the given produced water network. The results presented correspond to Case 1 from the previous section (base case – minimize total cost). It is worth noting that, in the context of this case study and based on the specific data analyzed, the emissions originating from disposal, storage, and pipeline operations are minimal compared to those from treatment and pipeline installation. For this reason, they are not depicted in the figures but are instead included in the analysis.



**Figure 6:** CO<sub>2</sub> emissions in PW networks [kton].



**Figure 7.** NO<sub>x</sub> emissions in PW networks [kton].



**Figure 8.** PM<sub>2.5</sub> emissions in PW Networks [kton].

### Emissions Trade-offs in Produced Water Management

The emissions calculations presented in Table 5 of our study provide insightful findings on the environmental tradeoffs of various produced water management strategies. These results reveal the trade-off between minimizing costs and reducing total system emissions. The data indicates that while some strategies effectively reduce overall costs, they may result in higher total emissions. Conversely, approaches focused on minimizing emissions demonstrate a substantial decrease in environmental impact but at an increased cost. These findings underscore the complexities involved in balancing economic and environmental objectives in produced water management, highlighting the need for multi-faceted approaches that consider both financial and ecological sustainability.

For the given case study, PARETO framework reduced PW treatment in favor of increasing produced water disposal to reduce the overall emissions. This result can be seen as contradicting from the water sustainability perspective, presenting a potential opportunity to explore different environmental objectives.

**Table 5:** Results environmental KPIs.

Objective (minimize)	Cost	Emissions w/ limits
Total Cost (k\$)	18,603	20,410
Total Emissions (kton)	126,486	126,482
Sourced Water (kbbbl)	515	515
Disposal Volume (kbbbl)	7,218	9,312
Reuse Volume (kbbbl)	8,805	8,805
Piping operational costs (k\$)	15	12
Disposal CAPEX (k\$)	142	286
Pipeline CAPEX (k\$)	111	127

It is crucial to acknowledge the intrinsic limitations of the network studied, particularly in the context of the Permian case study which primarily relies on piping for water transportation. Piping, as demonstrated by prior research [10], offers emission savings compared to trucking, which is absent in our base case scenario, inherently limiting the scope for further emission reductions. Furthermore, the scope of the presented case study is constrained to a 52-week period, covering only 16 production and completion pads. This limited geographical and temporal scale suggests that the observed emission savings, while seemingly modest, may represent a larger potential for emission reduction across broader basins and over the operational lifetime of wells.

Lastly, it is pertinent to consider that the direct emission savings, while valuable, may not fully capture the broader environmental and health benefits. These benefits are more significantly recognized when emissions are translated into cost dollars in terms of Human Health Effects (HHE). Although the conversion to an HHE-focused objective is beyond the scope of this iteration of the PARETO framework, it represents a critical area for future exploration to comprehensively assess the value of emission reductions.

### CONCLUSIONS

This work presents an update on project PARETO's capabilities for advanced treatment modeling and quality tracking, and the quantification of environmental justice and emissions in produced water networks. We demonstrated the use of PARETO framework to determine optimal infrastructure buildout, PW management, and the selection of treatment technologies that will enable potential beneficial reuse options in oil and gas produced water networks.

Further work will focus on demonstrating multi-objective optimization tools for optimizing produced water networks under different objectives (i.e., water quality, cost, emissions and environmental justice, PW reuse,

etc.). Additionally, PW quality can be seen as a source of uncertainty in the decision-making process; process design under uncertainty is a promising topic in the area of PW desalination.

## ACKNOWLEDGEMENTS

We gratefully acknowledge support from the U.S. Department of Energy, Office of Fossil Energy and Carbon Management, through the Environmentally Prudent Stewardship Program.

**Disclaimer:** This project was funded by the Department of Energy, National Energy Technology Laboratory an agency of the United States Government, through a support contract. Neither the United States Government nor any agency thereof, nor any of its employees, nor the support contractor, nor any of their employees, makes any warranty, expressor implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## REFERENCES

1. Veil, John. "US produced water volumes and management practices in 2012." Groundwater Protection Council, 2015.
2. L. Capper, "Seismicity and class II SWD Disposal Well Risks. Induced Seismicity: Myth versus Reality and Impacts we can anticipate," AAPG Southwest Section, Midland, Texas, 2017.
3. J. Park, R. Chen, A. Datta-Gupta, S. Lele and T. Tyrrell, "Quantitative Assessment of Induced Seismicity from Hydrocarbon Production and Produced Water Disposal in Azle Area, North Texas," in Unconventional Resources Technology Conference, Houston, Texas, 2021.
4. Drouven, M. G., Calderon, A. J., Zamarripa, M. A., Beattie, K. 2022. PARETO: An Open-Source Produced Water Optimization Framework. Optimization and Engineering. <https://doi.org/10.1007/s11081-022-09773-w>
5. Beattie, K.S., Gunter, D., Knueven, B., Lee, A., Ladshaw, A., Drouven, M., Bartholomew, T. WaterTAP v1.0.0. Oak Ridge National Lab. (ORNL), Oak Ridge, TN (United States), 2021. [URL: <https://doi.org/10.11578/dc.20210527.7>]. Lawrence

Berkeley National Lab. (LBNL), Berkeley, CA (United States); National Energy Technology Laboratory (NETL), Pittsburgh, PA, Morgantown, WV, and Albany, OR (United States); National Renewable Energy Lab. (NREL), Golden, CO (United States).

6. Lee, A., Ghouse, J. H., Eslick, J. C., Laird, C. D., Siirola, J. D., Zamarripa, M. A., Gunter, D., Shinn, J. H., Dowling, A. W., Bhattacharyya, D., Biegler, L. T., Burgard, A. P., Miller, D. C. 2022. The IDAES process modeling framework and model library – Flexibility for process simulation and optimization. Journal of Advanced Manufacturing and Processing, 3(3), 1-30. <https://doi.org/10.1002/amp2.10095>
7. Inflation Reduction Act of 2022, Pub. L. No. 117-169, § 60113 (2022), enacting Clean Air Act § 136, to be codified at 42 U.S.C. § 7436.
8. United States, Executive Office of the President [Joseph Biden]. Executive Order 14008: Tackling the Climate Crisis at Home and Abroad. 1 Feb 2021. Federal Register, vol. 86, no. 19, pp. 7619-7633, <https://www.govinfo.gov/content/pkg/FR-2021-02-01/pdf/2021-02177.pdf>
9. Council on Environmental Quality, <https://screeningtool.geoplatform.gov/en/#3/33.47/-97.5>
10. Bartholomew, T. V., Mauter, M. S. 2016. Multiobjective optimization model for minimizing cost and environmental impact in shale gas water and wastewater management. ACS Sustainable Chemistry & Engineering 4.7: 3728-3735.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>







# Part 1

## Peer-Reviewed Articles

Section 6: Design Education and Future of Design



# Jacobian-based Model Diagnostics and Application to Equation Oriented Modeling of a Carbon Capture System

Douglas A. Allan<sup>ab\*</sup>, Anca Ostace<sup>ab</sup>, Andrew Lee<sup>ab</sup>, Brandon Paul<sup>ab</sup>, Anuja Deshpande<sup>ab</sup>, Miguel A. Zamarripa<sup>ab</sup>, Joshua C. Morgan<sup>ab</sup> and Benjamin P. Omell<sup>a</sup>

<sup>a</sup> National Energy Technology Laboratory, Pittsburgh, PA 15236, USA

<sup>b</sup> NETL Support Contractor, Pittsburgh, PA 15236, USA

\*Corresponding author: [douglas.allan@netl.doe.gov](mailto:douglas.allan@netl.doe.gov)

## ABSTRACT

Equation-oriented (EO) modeling has the potential to enable the effective design and optimization of the operation of advanced energy systems. However, advanced modeling of energy systems results in a large number of variables and non-linear equations, and it can be difficult to search through these to identify the culprit(s) responsible for convergence issues. The Institute for the Design of Advanced Energy Systems Integrated Platform (IDAES-IP) contains a tool to identify poorly scaled constraints and variables by searching for rows and columns of the Jacobian matrix with small  $L^2$ -norms so they can be rescaled. A further singular value decomposition can be performed to identify degenerate sets of equations and remaining scaling issues. This work presents an EO model of a flowsheet developed for post-combustion carbon capture using a monoethanolamine (MEA) solvent system as a case study. The IDAES diagnostics tools were successfully applied to this flowsheet to identify problems to improve model robustness and enable the optimization of process design and operating conditions of a carbon capture system.

**Keywords:** Pyomo, Optimization, Carbon Dioxide Capture, Jacobian, Modelling

## INTRODUCTION

In order to achieve carbon neutrality by 2050, as is the US Department of Energy's present goal [1], a wide variety of energy systems must be deployed. Process optimization can help allocate resources in the most efficient manner to effect the changes necessary in the US and world economies to achieve this goal. The Institute for the Design of Advanced Energy Systems (IDAES) was founded in 2015 to study such advanced energy systems and to develop the IDAES Integrated Platform (IDAES-IP) to facilitate their development and optimization. [2] IDAES-IP is based on the Pyomo modeling language [3-4] in Python and has been used to simulate a wide variety of chemical and energy process systems.

Because IDAES-IP is equation-oriented (EO), it can provide improved convergence when closing recycle loops over the sequential-modular (SM) approach that is used in popular commercial process modeling and simulation tools like Aspen Plus®. [5] However, it also requires more user skill to ensure that the model is well-

formulated to benefit from these theoretical convergence improvements. IDAES has developed and implemented a diagnostics toolbox leveraging several years of experience of many experts in EO modeling, debugging, and optimization, making it available to the public. [6] In this work, we detail some of the techniques that have been incorporated in that toolbox as applied to a flowsheet being developed for analysis of post-combustion carbon capture (PCC) systems.

## METHODS

When prototyping a model, it is typically best to solve a "square problem," i.e., one in which there are an equal number of free variables and equality constraints. The number of degrees of freedom can be checked by the function `idaes.core.util.model_statistics.degress_of_freedom`. There are zero degrees of freedom in a square problem. Variable bounds can be present, as they are often helpful to keep the nonlinear solver from exploring areas with non-physical solutions

or areas in which equations become undefined. For example, if a logarithm of a variable is taken, its lower bound should be set to zero.

However, variable bounds *should not* be active at a solution, because the inclusion of such a bound is effectively a constraint, reducing the number of degrees of freedom by 1. The system of equations then either becomes infeasible, which precludes a solution, or degenerate, which means that some of the equations are redundant. Different nonlinear solvers handle degeneracy differently, but it either prevents or dramatically slows convergence to a solution. For this reason, inequality constraints more complex than variable bounds should be avoided when formulating a square problem.

The principal diagnostic methods used in this work utilize the Jacobian. When solving a root-finding problem

$$\begin{bmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_n(x_1, \dots, x_n) \end{bmatrix} := \mathbf{f}(\mathbf{x}) = \mathbf{0} \quad (1)$$

the Jacobian matrix of  $\mathbf{f}(\mathbf{x})$  is given by

$$J(\mathbf{x}) := \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \dots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \quad (2)$$

For a square problem, the Jacobian matrix is square. Each row of  $J(\mathbf{x})$  corresponds to an equality constraint and each column corresponds to a free variable. Therefore, we can examine  $J(\mathbf{x})$  to find clues about problems in equations and constraints. In the past, calculation of the Jacobian was difficult because derivatives had to either be calculated by hand or by finite differences. However, advances in algorithmic differentiation (AD) allow for automatic, precise calculation of derivatives. Pyomo offers access to the AD capacities of the AMPL solver library (ASL) [7] through the PyNumero interface [8].

The most common methods of solving multivariate root-finding problems of the form of (1), like the one utilized in IPOPT [9], which is freely available and was used as a nonlinear solver in this paper, are variations on Newton's Method. The method approximates  $\mathbf{f}(\cdot)$  as linear and then repeatedly solves the linearized equation

$$\mathbf{f}(\mathbf{x}_k) + J(\mathbf{x}_k)\Delta\mathbf{x}_k = \mathbf{0} \quad (3)$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta\mathbf{x}_k \quad (4)$$

until the condition

$$\|\mathbf{f}(\mathbf{x}_{k+1})\|_2 \leq \varepsilon \quad (5)$$

in which  $\|\cdot\|_2$  is the vector two (Euclidean) norm, is satisfied for some chosen tolerance  $\varepsilon$ . In a practical implementation, the Newton step (4) may be truncated, either to account for variable bounds or because  $\|\mathbf{f}(\mathbf{x}_{k+1})\|_2 > \|\mathbf{f}(\mathbf{x}_k)\|_2$ , but it is desirable to take the full step when possible. It is a sign of a well-formulated method working

on a well-formulated problem when the full step (4) is frequently taken until a solution is reached.

A condition number of a matrix is a measure of the sensitivity of changes in the solution to a system of linear equations. For example, for the Newton step-finding problem

$$J(\mathbf{x}_k)\Delta\mathbf{x}_k = -\mathbf{f}(\mathbf{x}_k) \quad (6)$$

suppose we perturb the function output by  $\delta\mathbf{f}_k$  and want to estimate the perturbation in the resulting step  $\delta\mathbf{x}_k$

$$J(\mathbf{x}_k)(\Delta\mathbf{x}_k + \delta\mathbf{x}_k) = -(\mathbf{f}(\mathbf{x}_k) + \delta\mathbf{f}_k) \quad (7)$$

Use of the condition number for the 2-norm  $\kappa_2(J(\mathbf{x}_k))$  gives us the bound

$$\|\delta\mathbf{x}_k\|_2 \leq \kappa_2(J(\mathbf{x}_k))\|\delta\mathbf{f}_k\|_2 \quad (8)$$

In practice, such perturbations in the function output always exist from the roundoff errors in floating point arithmetic. The relative error inherent in double precision floating point arithmetic is on the order of  $10^{-16}$ . Additionally, a large condition number  $\kappa_2(J(\mathbf{x}_k))$  (greater than about  $10^8$ ) makes it difficult to solve the problem (6) numerically, increasing the time that Newton iterations take to solve.

The condition number of  $J(\mathbf{x})$  is given by

$$\kappa_2(J(\mathbf{x})) = \|J(\mathbf{x})\|_2 / \|J(\mathbf{x})^{-1}\|_2 \quad (9)$$

in which  $\|\cdot\|_2$  is the operator norm induced by the vector 2-norm, provided  $J(\mathbf{x})$  is full rank. A more useful formula comes from the singular value decomposition (SVD). We factorize

$$J(\mathbf{x}) = U\Sigma V^T \quad (10)$$

in which  $U$  and  $V$  are orthogonal matrices and

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_n \end{bmatrix} \quad (11)$$

in which  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$ . If  $J(\mathbf{x})$  is full rank, then  $\sigma_n > 0$  and

$$\kappa_2(J(\mathbf{x})) = \sigma_1/\sigma_n \quad (12)$$

Therefore, we can improve the problem formulation by reducing the condition number by bringing the values of  $\sigma_1$  and  $\sigma_n$  closer together. This can be done by appropriate scaling of variables and constraints.

## Scaling

When working in SI units, it is common to have variables and equations that vary over many orders of magnitude, from mole fractions of trace components with magnitudes  $10^{-6}$  to enthalpy flow values that can be up to  $10^6$ . IDAES-IP offers a variety of tools to assist in scaling both variables and constraints. Scaling serves two purposes: first, to ensure that the convergence criterion

(5) guarantees that all equations are satisfied without being unduly difficult to satisfy, and second, to reduce the condition number of the Jacobian. A method exists [10] for calculating scaling factors for variables and constraints that minimize the condition number of the Jacobian by solution of a convex program. While this method achieves the second goal, it has no guarantee of achieving the first goal.

Let there be two example equations that may occur in, for example, a heater:

$$y_{trace,in} - y_{trace,out} = f_1(\mathbf{x}) \quad (13)$$

$$H_{in} + Q - H_{out} = f_2(\mathbf{x}) \quad (14)$$

in which  $y_{trace}$  is the mole fraction of a trace component,  $H$  represents stream enthalpies, and  $Q$  is a heat duty. If the initial values were off by a relative error of 100%, the error in (13) would be of the order  $10^{-6}$  and the error in (14) of the order  $10^6$ . If equations of these magnitudes are loaded into the same  $f(\cdot)$ , significant error can exist in the values of mole fractions with the equation ostensibly satisfied while the solver would strain to keep reducing the error in the enthalpy equation far beyond what is significant. IDAES and Pyomo allow us to give the first equation a scaling factor of  $10^6$  and the second one of  $10^{-6}$  to bring them to the same basis.

Let us note the effect of such scaling on the function and the Jacobian. We would replace (13) and (14) with

$$10^6(y_{trace,in} - y_{trace,out}) = \tilde{f}_1(\mathbf{x}) \quad (15)$$

$$10^{-6}(H_{in} + Q - H_{out}) = \tilde{f}_2(\mathbf{x}) \quad (16)$$

and the corresponding rows in the Jacobian would be replaced by

$$J(\mathbf{x}) := \begin{bmatrix} 10^6 \frac{\partial f_1}{\partial x_1} & \dots & 10^6 \frac{\partial f_1}{\partial x_n} \\ 10^{-6} \frac{\partial f_2}{\partial x_1} & \dots & 10^{-6} \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \vdots \end{bmatrix} \quad (17)$$

Now we have coefficients in the Jacobian of extremely different values. Row 1 has norm  $10^6\sqrt{2}$  and row 2 has norm of  $10^{-6}\sqrt{3}$ . We can derive a lower bound for the condition number in terms of rows with largest and smallest 2-norm. Rearrange Jacobian rows so that they are sorted in descending order of 2-norm magnitude

$$J(\mathbf{x}) = \begin{bmatrix} r_{max} \\ \vdots \\ r_{min} \end{bmatrix} \quad (18)$$

Then we have that

$$\|r_{max}\|_2 = \left\| \begin{bmatrix} 1 & 0 & \dots \\ \vdots \\ r_{min} \end{bmatrix} \right\|_2 \leq \sigma_1 \quad (19)$$

and

$$\|r_{min}\|_2 = \left\| \begin{bmatrix} \dots & 0 & 1 \\ \vdots \\ r_{min} \end{bmatrix} \right\|_2 \geq \sigma_n \quad (20)$$

so, we have that

$$\kappa_2(J(\mathbf{x})) = \sigma_1/\sigma_n \geq \|r_{max}\|_2/\|r_{min}\|_2 \quad (21)$$

Presently, the Jacobian matrix has condition number greater than  $\frac{\sqrt{6}}{3} \cdot 10^{12}$ . To reduce it, we need to also set scaling factors for variables. We define

$$\tilde{y}_{trace,in} := 10^6 y_{trace,in}$$

$$\tilde{H}_{out} := 10^{-6} H_{out}$$

Suppose we had  $y_{trace,in} = x_1$  and  $H_{out} = x_n$ . The Jacobian now looks like

$$J(\mathbf{x}) := \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & 10^{12} \frac{\partial f_1}{\partial x_n} \\ 10^{-12} \frac{\partial f_2}{\partial x_1} & \dots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \vdots \end{bmatrix} \quad (22)$$

The entry involving  $10^{12}$  would be potentially problematic, but because  $\frac{\partial f_1}{\partial x_n} = 0$ , it is irrelevant. In short, scaling a constraint *multiplies* the corresponding Jacobian *row* by that scaling factor, and scaling a variable *divides* the corresponding Jacobian *column* by that scaling factor.

It is not necessary to manually create transformed variables and constraints because Pyomo allows the user to store scaling factors associated with variables and constraints as part of the model. The unscaled model can then be transformed into the scaled model before being passed to the nonlinear solver.<sup>1</sup> A scaled version of the model can also be retrieved for debugging purposes through a scaling transformation.

In a model with thousands of variables and constraints, however, it is difficult to keep track of which variables and constraints have been scaled, and whether the scaling factors assigned are appropriate. The IDAES diagnostics toolbox, described in [6], has the methods `display_constraints_with_extreme_jacobians` and `display_variables_with_extreme_jacobians` that iterate over Jacobian rows and columns, respectively, to display the ones with 2-norms larger and smaller than some user-specified tolerances (by default  $10^4$  and  $10^{-4}$ ). These methods highlight areas of the model that require work, but the user still needs to determine appropriate values for scaling factors, as well as which variables and constraints need to be scaled. For example, rows 1 and 2 would be highlighted in the Jacobian of (17),

<sup>1</sup> Scaling at the time of passing the model to the solver can be done by use of the `ipopt_v2` solver in Pyomo through the `SolverFactory` function.

when it is the variables, not the constraints, that require scaling.

Scaling is often an iterative process. Once appropriate scaling factors are assigned to variables, bringing the Jacobian into the form (22), large or small entries can appear in additional rows, corresponding to other constraints, that would then need to be scaled. Finally, a large entry can correspond to a model that is *intrinsically* ill-conditioned. For example, when modeling a process with a large recycle ratio, large terms can appear in the corresponding Jacobian. Small impurities in the feed stream can become concentrated by factors of 100 or 1000 as a feature of the process. Scaling might be able to hide this feature, but it can cause a loss in accuracy and will be revealed by performing SVD of the Jacobian.

## Singular Value Decomposition

Even once rows and columns with inappropriately large or small norms have been removed by scaling, the Jacobian's condition number, which can be revealed through the SVD of the Jacobian, can still be extremely large. Both  $\sigma_1$  and  $\sigma_n$  appear directly in the expression for the condition number, so reducing  $\sigma_1$  and increasing  $\sigma_n$  should help reduce it. In practice,  $\sigma_1$  is usually reduced to a reasonable level by scaling rows and columns, so we will focus on increasing  $\sigma_n$  instead.

Consider the solution to a Newton step problem

$$\mathbf{f}(\mathbf{x}_k) = -J(\mathbf{x}_k)\Delta\mathbf{x}_k \quad (23)$$

By decomposing  $J(\mathbf{x})$  using the SVD and using that to solve (23), we obtain

$$\mathbf{f}(\mathbf{x}_k) = -U\Sigma V^T \Delta\mathbf{x}_k \quad (24)$$

$$-\Sigma^{-1}U^T \mathbf{f}(\mathbf{x}_k) = \Sigma^{-1}U^T U \Sigma V^T \Delta\mathbf{x}_k = \Delta\mathbf{x}_k \quad (25)$$

So, in solving the perturbed problem, the constraint residual is projected by  $U^T$  into orthogonal components in the space of the singular values, scaled by the inverse singular values, and those scaled values are projected by  $V$  into the space of variables.

Looking at (25) another way

$$-U^T \mathbf{f}(\mathbf{x}_k) = \Sigma V^T \Delta\mathbf{x}_k \quad (26)$$

Each column of  $U$  (row of  $U^T$ ) is an orthogonal vector, associating elements of  $\mathbf{f}(\mathbf{x}_k)$ , i.e., particular constraints, with a singular value. Likewise, each column of  $V$  (row of  $V^T$ ) is a vector associating elements of  $\Delta\mathbf{x}_k$ , i.e., particular variables, with a singular value. The SVD shows which constraints are being satisfied using which variables in a neighborhood of  $\mathbf{x}_k$ , and the sensitivity of the variables with respect to constraints.

If the smallest singular value  $\sigma_n \ll 1$ , a small change in certain function values requires a large change in variable values. By looking at the constraints involved in the  $n$ th left-singular vector  $u_n$  and the variables involved in  $n$ th right singular vector  $v_n$  we can attempt to determine

what is causing this dysfunctional relationship. Because of the dense linear algebra involved, we expect most entries of  $u_n$  and  $v_n$  to be populated by nonzero numbers. A simple but crude method is to filter for indices that have values greater than some tolerance. We have found absolute values of 0.1 to 0.3 work well for this. If too many variables and constraints appear, raise the tolerance. If too few appear, decrease it.

Ideally, we would want the condition number of  $J(\mathbf{x}_k)$  to be relatively small. In practice, however, we typically accomplish a condition number on the order of  $10^6$ - $10^8$ . That means filtering for singular values smaller than  $10^{-8}$ - $10^{-6}$  and attempting to remedy them. In general, there are four causes of small singular values:

1. Incorrect scaling of variables or constraints
2. Redundant equations, hinting a problem that over-specifies some variables and under-specifies others, i.e., global degeneracy
3. A local singularity, caused by evaluating  $J(\mathbf{x})$  at a point where it locally loses rank, i.e., local degeneracy
4. Attempting to solve a problem that is inherently ill-conditioned.

Incorrect scaling is probably the most common cause of small singular values of values  $10^{-12}$ - $10^{-6}$ , followed by inherent ill-conditioning. For singular values smaller than  $10^{-12}$ , the cause is typically local or global degeneracy. The next section presents a carbon capture application in which these concepts can be demonstrated.

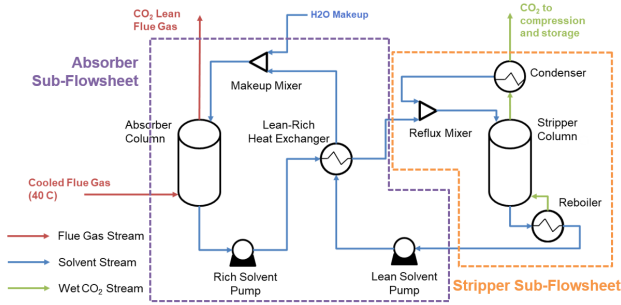
A similar technique for identifying sets of degenerate equations is given by the Degeneracy Hunter algorithm proposed in [11]. This algorithm is also incorporated into the IDAES diagnostics toolbox. It differs from this technique by performing a QR factorization instead of an SVD, then solving a mixed integer linear program (MILP) to find an irreducible degenerate set of equations. Which algorithm is appropriate for a given problem depends on the algorithms available for a sparse QR factorization, a sparse SVD, and MILP solution. An advantage the SVD provides, however, is that it shows which combinations of variables are being used to satisfy combinations of constraints. If a degenerate constraint is removed, then another constraint must be added to constrain that combination of free variables.

## CASE STUDY: POST COMBUSTION CARBON CAPTURE FLOWSHEET

A flowsheet in which  $\text{CO}_2$  is captured from flue gas (from a ~690 MWe Natural Gas Combined Cycle power plant) by absorption into monoethanolamine (MEA) provides several examples in which these diagnostic tests were of great assistance in model refinement and increasing robustness. This flowsheet was first presented in [12], but despite the techniques used to provide robust



solutions, such as a four-stage initialization routine for the absorption and stripping columns, it remained fragile and prone to failure. It was desirable to improve convergence and numerical robustness so the flowsheet could be used for robust design optimization. The flowsheet consists of an absorption column, stripper column, lean-rich heat exchanger, and balance-of-plant equipment. It is divided into two sub-flowsheets, an absorber section and a stripper section, that are solved independently before being linked and solved together.



**Figure 1.** MEA carbon capture flowsheet, divided into stripper and absorber sections.

We begin with a flowsheet that has some partial scaling applied to the column model, but does not have scaling for any other models. The column model has already been reformulated to remove division in constraints where possible, which helps avoid bad numeric behavior when the denominator of an expression is nearly zero at an intermediate iteration. For an unscaled or partially-scaled model, the output can be hundreds of lines long. The output length is mostly the result of scaling issues in indexed variables or constraints resulting in those variables or constraints being printed for each value of their indices. Since the column model is discretized into 40 finite elements along its length, 10 badly-scaled equations results in 400 entries. Therefore, the output cannot be fully displayed here, but some representative entries can be shown.

## Jacobian Analysis

Here, we run the diagnostics tools after a successful solution of the flowsheet. They can be used after failures to solve a flowsheet, but care should be taken as variables may not have realistic values, which is reflected in the model Jacobian. The following is one line (out of hundreds) produced by the `display_variables_with_extreme_jacobians` command from the `DiagnosticsToolbox`:

```
fs.strip_section.reflux_mixer.rich_solvent_state[0.0].temperature: 1.997E+08
```

This entry indicates a column norm of  $2 \cdot 10^8$ . We can inspect the associated column of the Jacobian to find which rows are associated with this large value. In IDAES-

IP, we can do this using the `display_constraints_including_variable` function in the `SVDToolbox`.

```
fs.strip_section.reflux_mixer.enthalpy_mixing_equations[0.0]: 1.997e+08
```

```
fs.strip_section.reflux_mixer.rich_solvent_state[0.0].log_k_eq_constraint[bicarbonate]: 5.364e+00
```

```
fs.strip_section.reflux_mixer.rich_solvent_state[0.0].log_k_eq_constraint[carbamate]: 6.551e+00
```

```
fs.strip_section.reflux_mixer.rich_solvent_state[0.0].log_conc_mol_phase_comp_true_eq[Liq,MEACO_0-]: 1.514e+00
```

```
fs.strip_section.reflux_mixer.rich_solvent_state[0.0].log_conc_mol_phase_comp_true_eq[Liq,HCO3-]: 2.333e-01
```

```
fs.strip_section.reflux_mixer.rich_solvent_state[0.0].log_conc_mol_phase_comp_true_eq[Liq,MEA+]: 1.748e+00
```

```
fs.strip_section.reflux_mixer.rich_solvent_state[0.0].log_conc_mol_phase_comp_true_eq[Liq,H2O]: 3.183e+01
```

```
fs.strip_section.reflux_mixer.rich_solvent_state[0.0].log_conc_mol_phase_comp_true_eq[Liq,MEA]: 1.129e+00
```

```
fs.strip_section.reflux_mixer.rich_solvent_state[0.0].log_conc_mol_phase_comp_true_eq[Liq,CO2]: 5.685e-03
```

```
fs.rich_temperature: 1.000e+02
```

From inspection, the problematic Jacobian entry is associated with a mixer enthalpy mixing equation that is not yet scaled.

To fix the problem, we need to determine an appropriate scaling factor. We can calculate such a value from scaling factors for both the molar flow and the specific molar enthalpy. Finding the molar flow scaling factor is easy; it is generally observed that the molar flow rates are on the order of  $10^3$ - $10^4$  throughout the flowsheet, so we choose  $3 \cdot 10^{-4}$  as a scaling factor for the molar flow rate. Determining a scaling factor for the specific molar enthalpy is harder. Enthalpy can be either positive or negative in different areas of the flowsheet, so its order of magnitude does not make a good scaling factor. By evaluating the liquid-phase enthalpy in the reflux mixer, the water makeup mixer, and the reboiler, we get values of  $-42200$ ,  $-43600$ , and  $-38800$ . With an apparent range of 4800, a scaling factor of  $3 \cdot 10^{-4}$  is also a good choice for this variable. Values for scaling factors do not have to be exact: almost counts in horseshoes, hand grenades, and model scaling.

The process of removing this scaling issue took us through at least three Pyomo sub-models (and associated Python files) to determine a process through which the scaling factor for one equation was calculated. However, had reasonable default scaling factors for the property sub-model been set by the user ahead of time, this issue would never have arisen.

Next, we consider scaling the equations describing

heat transfer in the column

$$h_V a_e Ack = -(\underline{C}_{pCO_2,V} N_{CO_2,V} + \underline{C}_{pH_2O,V} N_{H_2O,V}) \quad (28)$$

$$h'_V a_e (1 - \exp(Ack)) = \underline{C}_{pCO_2,V} N_{CO_2,V} + \underline{C}_{pH_2O,V} N_{H_2O,V} \quad (29)$$

in which  $h_V$  is the heat transfer coefficient from liquid to vapor,  $a_e$  is the effective area of heat transfer per unit column volume,  $\underline{C}_{p,i,V}$  is the heat capacity of species  $i$ ,  $N_{i,V}$  is the molar flux of species  $i$  from vapor to liquid, and  $Ack$  is the Ackerman factor, which is a measure of how much diffusive heat fluxes are distorted by convective heat transfer. It was discovered that  $|Ack| \leq 10^{-2}$  for operating conditions of interest and was frequently on the order of  $10^{-5}$ . This causes ill-conditioning in both equations because the factors  $Ack$  and  $1 - \exp(Ack)$  became close to zero. Because  $Ack$  could vary over several orders of magnitude depending on location and operating conditions, assigning a consistent scaling factor is difficult.

If we first rearrange (28) and (29) to make the relationship between  $h_V$  and  $h'_V$  more clear

$$h'_V = h_V \frac{Ack}{(\exp(Ack)-1)} \quad (30)$$

several solutions are possible. The function

$$\theta(x) = \frac{x}{\exp(x)-1} \quad (31)$$

has an indeterminate form at  $x = 0$ , but a well-defined Taylor expansion

$$\theta(x) \approx 1 - \frac{x}{2} + \frac{x^2}{12} + O(x^4) \quad (32)$$

A Taylor approximation could be substituted for  $\theta(Ack)$  in (30), but  $\theta(x)$  could also be implemented as an external grey-box function in IDAES, switching between the full form (31) and Taylor form (32) based on  $|x|$ . Both options were eventually implemented. An external function for  $\theta(x)$  is implemented in IDAES with a sixth-order Taylor approximation, but the trivial Taylor approximation  $\theta(x) \approx 1$  is presently used in the MEA flowsheet because not every solver in Pyomo supports external functions.

## SVD Analysis

The SVD also tells us valuable information about the state of the flowsheet. Its condition number is  $9.2 \cdot 10^{17}$ , so the matrix is singular to machine precision. The smallest singular value is  $\sigma_n = 3.1 \cdot 10^{-10}$ . Using a tolerance for the singular vectors of 0.1, we find the following variables involved:

```
fs.stripper_section.reflux_mixer.re-
flux_state[0.0].log_conc_mol_phase_comp_true[Liq,H
CO3_-]
```

```
fs.stripper_section.reflux_mixer.re-
flux_state[0.0].log_conc_mol_phase_comp_true[Liq,M
EA_+]
```

```
fs.stripper_section.reflux_mixer.re-
flux_state[0.0].log_conc_mol_phase_comp_true[Liq,M
EA]
```

```
fs.stripper_section.reflux_mixer.re-
flux_state[0.0].log_conc_mol_phase_comp_true[Liq,M
EACOO_-]
```

and the following constraints involved:

```
fs.stripper_section.condenser.liquid_phase[0.0].ap-
pr_to_true_species[Liq,MEA]
```

```
fs.stripper_section.condenser.liquid_phase[0.0].tr-
ue_mole_frac_constraint[Liq,HCO3_-]
```

```
fs.stripper_section.condenser.liquid_phase[0.0].tr-
ue_mole_frac_constraint[Liq,MEA_+]
```

```
fs.stripper_section.condenser.liquid_phase[0.0].tr-
ue_mole_frac_constraint[Liq,MEA]
```

```
fs.stripper_section.condenser.liquid_phase[0.0].lo-
g_conc_mol_phase_comp_true_eq[Liq,HCO3_-]
```

```
fs.stripper_section.condenser.liquid_phase[0.0].lo-
g_conc_mol_phase_comp_true_eq[Liq,MEA_+]
```

```
fs.stripper_section.condenser.liquid_phase[0.0].lo-
g_conc_mol_phase_comp_true_eq[Liq,MEA]
```

```
fs.stripper_section.condenser.liquid_phase[0.0].ap-
pr_to_true_species[Liq,HCO3_-]
```

```
fs.stripper_section.condenser.liquid_phase[0.0].ap-
pr_to_true_species[Liq,MEA_+]
```

All these variables and constraints occur in the condenser, which makes interpretation of the problem easier. However, if multiple degeneracies are present, they can mix between different tiny singular values, so sometimes additional analysis is necessary to separate different degeneracies.

Because the property model in this flowsheet does not account for amine volatility, the mole fraction of MEA in the condenser is effectively zero. All the variables implicated here are dissociation species of MEA, which similarly have concentrations of effectively zero. The constraints are likewise the disassociation equations for MEA. Thus, the absence of MEA causes degeneracy in the system of equations governing the dissociation reactions. Different strategies can be employed to overcome this degeneracy. The easiest is to increase the mole fraction of MEA to around  $10^{-4}$ . However, this merely mitigates the ill-conditioning—it does not remove it—and the addition of extra MEA in the system could cause problems with material balances converging.

IDAES-IP allows the user to define property packages that bundle together thermodynamic calculations into a single sub-model that can then be employed in different unit models. The solution we employed was to create a duplicate property package without the dissociation reactions and use that for the condenser and reflux mixer. That solution works only because the liquid phase property sub-model does not rely on ion concentrations for the calculation of enthalpy. For one that requires ion concentrations to calculate the enthalpy of mixing, like

eNRTL, another solution would have to be devised.

## Limits of Scaling and Reformulation

Not every problem that can be discovered by these diagnostic tools has a nice solution. Such is the case with the system of equations for the enhancement factor for mass transfer in reactive systems. Taken from [13] for the MEA-CO<sub>2</sub> system, the full enhancement factor model has ten tightly coupled numerical expressions in it. However, the core numerical issues derive from two equations:

$$E = 1 + (E_{\infty}^* - 1) \frac{1 - Y_{MEA}^i}{1 - Y_{CO_2}^b} \quad (33)$$

$$E = Ha \sqrt{Y_{MEA}^i \frac{1 - Y_{CO_2}^*}{1 - Y_{CO_2}^b}} \quad (34)$$

in which  $E$  is the enhancement factor,  $E_{\infty}^*$  is the “instantaneous enhancement factor,” a theoretical maximum value for the enhancement factor,  $Y_{MEA}^i$  is a dimensionless concentration of MEA at the vapor/liquid interface,  $Y_{CO_2}^*$  is the dimensionless concentration of CO<sub>2</sub> at the interface,  $Y_{CO_2}^b$  is the dimensionless concentration of CO<sub>2</sub> in the bulk liquid, and  $Ha$  is the Hatta number, the ratio of reaction film to the rate of diffusion in the film. Absorption happens when  $Y_{CO_2}^b < 1$ , desorption happens when  $Y_{CO_2}^b > 1$ , and equilibrium occurs when  $Y_{CO_2}^b = 1$ .

The problem occurs near equilibrium, because the expression  $1 - Y_{CO_2}^b$  is nearly equal to zero, and the quotients in (33) and (34) are nearly singular. At actual solutions to the system of equations, numerical tests show that  $1 - Y_{CO_2}^*$  and  $1 - Y_{MEA}^i$  also approach zero, so the quotient remains defined. However, since we now must deal with a multivariate function, we have been unsuccessful at removing the singularity by use of a Taylor series, like we did with the Ackerman factor. The expressions  $1 - Y_{CO_2}^b$  and  $1 - Y_{CO_2}^*$  will have the same sign at any solution to the activity factor system. However, contrary to what is stated in [13], there exist at least one scenario in which  $1 - Y_{CO_2}^b$  has the opposite sign as  $1 - Y_{MEA}^i$ . When CO<sub>2</sub> is desorbing at low temperatures, the rate of reaction can become slow enough that the rate of desorption is lower with reaction than it would be without reaction, and the enhancement factor drops below one. That fact limits possible reformulations of these two equations.

The equations were ultimately reformulated in several steps. First, a log variable  $S_{CO_2}$  was introduced for one of the singular ratios:

$$(1 - Y_{CO_2}^b) \exp(S_{CO_2}) = 1 - Y_{CO_2}^* \quad (35)$$

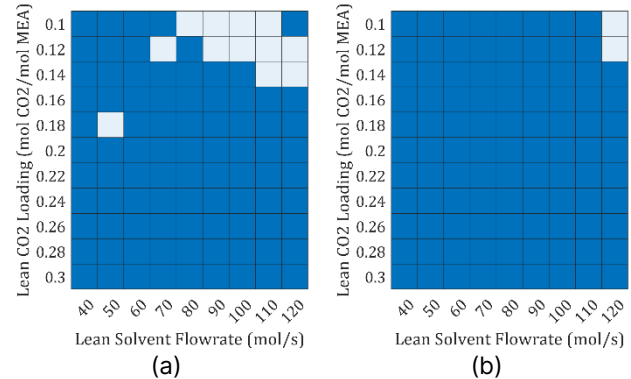
so that (34) could be rewritten as the linear equation

$$\log E = \log Ha + \frac{1}{2} \log Y_{MEA}^i + S_{CO_2} \quad (36)$$

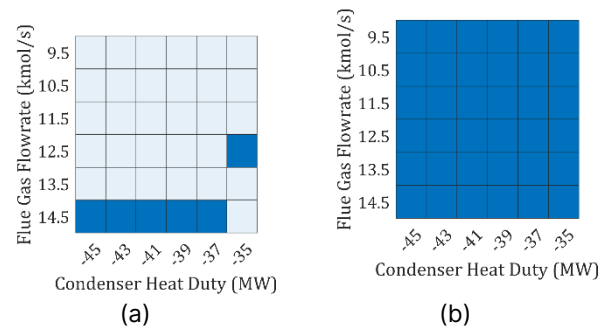
in which  $\log E$ ,  $\log Ha$ , and  $\log Y_{MEA}^i$  are all additional log variables defined using equations of the form

$$\exp(\log x) = x \quad (37)$$

To be completely clear,  $\log x$  is a single variable, not an expression denoting the logarithm of the variable  $x$ , which we denote  $\log(x)$ . We have that  $\log x = \log(x)$  at solutions of the model, i.e., when  $f(x) = 0$ , but equality does not necessarily hold at intermediate Newton iterates. The benefit of implicitly taking the logarithm in equations like (35) and (37) instead of using a log function in (36) is that the variable  $\log x$  can maintain numerical accuracy even when the values of  $x$  are extremely small or extremely large.



**Figure 1.** Convergence of the pilot-scale stand-alone column model depending on inlet parameters (a) before and (b) after the scaling and reformulations detailed here. A filled square indicates it converged, an unfilled square indicates it did not converge. In the reformulated model, divergence typically is a result of  $Y_{CO_2}^*$  incorrectly converging to 1, resulting in (35) and (38) becoming degenerate.



**Figure 2.** Convergence of the plant-scale full flowsheet model depending on inlet parameters (a) before and (b) after the scaling and reformulations detailed here. A filled square indicates it converged, an unfilled square indicates it did not converge.

Unfortunately, both (35) and (39) become degenerate when  $Y_{CO_2}^b = 1$ . Because the enhancement factor system can be largely decoupled from the remaining equations, a surrogate model was created to

give the log enhancement factor as a function of liquid phase CO<sub>2</sub> and H<sub>2</sub>O loading, vapor phase CO<sub>2</sub> partial pressure, temperature, and liquid and vapor mass transfer coefficients. However, too much accuracy was lost in the first attempt at a surrogate model, and the reformulated enhancement factor model is reliable enough to keep. Figures 1a and 1b show the convergence of the original and reformulated stand-alone column models, while Figures 2a and 2b show the convergence of the original and reformulated full flowsheet models (with all changes, not just to the enhancement factor calculations), respectively. The numerical robustness of the reformulated and scaled models is significantly improved over that of the original model.

It is possible to use the same technique as (35) for the ratio of MEA to CO<sub>2</sub> in (33). However, that implicitly assumes that  $1 - Y_{CO_2}^*$  and  $1 - Y_{MEA}^i$  have the same sign, which is true for most operating conditions of interest, but is not true for desorption at low temperatures. The resulting reformulation of (33) is then

$$\exp(\log E)(1 - Y_{CO_2}^b) = \exp(\eta)(1 - Y_{MEA}^i) \quad (38)$$

$$\exp(\eta) = E^* - 1 \quad (39)$$

in which the log variable  $\eta$  has been introduced for the instantaneous enhancement factor minus one because there is a simple expression for  $\eta$  (not given here) in terms of logarithms of other variables.

## CONCLUSIONS

We have outlined model diagnostic methods using tools available in Pyomo and IDAES-IP. As illustrated through the example of the MEA flowsheet, these tools can be of great assistance to the user by serving to point out likely problems. However, they do not solve them for the user. Finding problems in a model with tens of thousands of constraints and variables is a major service, but the user's insight and modelling expertise is still necessary to solve them. Nevertheless, these tools and techniques can help make EO modeling frameworks more accessible and facilitate the use of advanced optimization techniques in process design.

## Acknowledgements

This work was conducted as part of the U.S. Department of Energy's Institute for the Design of Advanced Energy Systems (IDAES) supported by the Office of Fossil Energy and Carbon Management through the Simulation-based Engineering/Crosscutting Research and Carbon Capture Programs.

## Disclaimer

This project was funded by the Department of Energy, National Energy Technology Laboratory an agency

of the United States Government, through a support contract. Neither the United States Government nor any agency thereof, nor any of its employees, nor the support contractor, nor any of their employees, makes any warranty, expressor implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## REFERENCES

1. US Department of Energy. <https://www.energy.gov/articles/how-were-moving-net-zero-2050>
2. Lee, A., Ghouse, J. H., Eslick, J. C., Laird, C. D., Siirola, J. D., Zamarripa, M. A., Gunter, D., Shinn, J. H., Dowling, A. W., Bhattacharyya, D., Biegler, L. T., Burgard, A. P., & Miller, D. C. The IDAES process modeling framework and model library—Flexibility for process simulation and optimization. *J Adv Manuf Process* 3:e10095. (2021)
3. Bynum, M.L., Hackebeil, G.A., Hart, W.E., Laird, C.D., Nicholson, B.L., Siirola, J.D., Watson, J.-P., Woodruff, D.L. Pyomo— Optimization Modeling in Python. Springer (2021).
4. Hart, W.E., Watson, J.-P., Woodruff, D.L. Pyomo: modeling and solving mathematical programs in Python. *Math Program Comput* 3 219–260. (2011)
5. ASPENTech. ASPEN Engineering Suite.
6. Lee, A., Dowling, A. W., Parker, R., Poon, S., Gunter, D., Nicholson, B. Model Diagnostics for Equation Oriented Models: Roadblocks and the Path Forward. *Proc of FOCAPD*. (2024)
7. Gay, D.M., Hooking Your Solver to AMPL. (2017)
8. Rodriguez, J.S., Parker, R.B., Laird, C.D., Nicholson, B.L., Siirola, J.D., Bynum, M.L. Scalable Parallel Nonlinear Optimization with PyNumero and Parapint. *INFORMS J Comput* 35:509–517. (2023)
9. Wächter, A., Biegler, L.T. On the Implementation of an Interior-Point Filter Line-Search Algorithm for Large-Scale Nonlinear Programming. *Math Program* 106:25–57. (2004)
10. Braatz, R.D., Morari, M. Minimizing the Euclidean Condition Number. *Siam J Control Opt* 32:1763–1768 (1994)
11. Dowling A.W., Biegler L.T. (2015). Degeneracy Hunter: An Algorithm for Determining Irreducible

Sets of Degenerate Constraints in Mathematical Programs. *Comput Aided Chem Eng.* 37:809–814 (2015)

12. Akula, P., Eslick, J., Bhattacharyya, D., Miller, D.C. Model Development, Validation, and Optimization of an MEA-Based Post-Combustion CO<sub>2</sub> Capture Process under Part-Load and Variable Capture Operations. *Ind Eng Chem Res* 60:5176–5193 (2021)
13. Gaspar, J., Fosbøl, P.L. A general enhancement factor model for absorption and desorption systems: A CO<sub>2</sub> capture case-study. *Chem Eng Sci* 138:203–215 (2015)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Integration of Process Design and Intensification Learning via Combined Junior Course Project

Madelyn R. Ball, Oishi Sanyal, and Yuhe Tian

West Virginia University, Department of Chemical and Biomedical Engineering, Morgantown, WV, United States

\* Corresponding Authors: [madelyn.ball@mail.wvu.edu](mailto:madelyn.ball@mail.wvu.edu), [oishi.sanyal@mail.wvu.edu](mailto:oishi.sanyal@mail.wvu.edu), [yuhe.tian@mail.wvu.edu](mailto:yuhe.tian@mail.wvu.edu).

## ABSTRACT

We present the implementation of combined junior course projects encompassing three core courses: reaction engineering, separations, and process simulation and design. The combined project aims to enhance the vertical integration of process design learning through all levels of the curriculum. We design the projects to utilize novel modular process technologies (e.g., membrane separation) and to emphasize new process design goals (e.g., sustainability, decarbonization). Two example projects, respectively on green methanol synthesis and ethylene oxide production, are showcased for project implementation. Feedback from junior and senior students is also presented to motivate the development of such joint project in CHE curriculum. We will also discuss the challenges we hope to address to maximize student learning from this unique project.

**Keywords:** Education, Modelling and Simulations, Process Design, Process Intensification, Carbon Capture.

## INTRODUCTION

Chemical Engineering curriculum is well-recognized for its broad scope covering many core concepts with a wide range of engineering applications. Students typically acquire the fundamentals distributed over courses in the first three years while applying the entire body of knowledge to a capstone design project during the senior year. Given this, vertical integration of process design experiences through the curriculum<sup>1-2</sup> offers promise to help students develop a connected ChE knowledge map and foster critical thinking before starting senior design.

In this paper, we present our efforts on advancing junior-level design education, in particular, leveraging combined projects joint across three courses that students are taking in the Spring semester, namely Reaction Engineering, Separations, and Process Simulation and Design. The major goals of the combined project include:

1. Students work in teams to solve open ended chemical engineering problems.
2. Integrate topics and knowledge from multiple courses towards plant design and optimization.
3. Learn emerging new modular units (e.g., membrane) and integrate to existing process options.
4. Learn new process design goals including decarbonization, sustainability, etc.

## EXAMPLES OF COMBINED PROJECT

### Project 1: Methanol Synthesis using Captured CO<sub>2</sub> and Green Hydrogen

The Spring 2022 combined project is shown in Fig. 1, which involved the modeling, simulation, and optimization of an industrial-scale methanol synthesis process using captured CO<sub>2</sub> and green hydrogen. The CO<sub>2</sub> capture unit, which was the focus of Separations course, involved designing and optimizing a membrane-based process for CO<sub>2</sub> capture from natural gas combined cycle plant<sup>3</sup>. This enabled the students to be familiar with a relatively new separation technology as opposed to traditional sorbent-based processes. For the CO<sub>2</sub> hydrogenation reactor, the focus of Reaction Engineering course, students analyzed various reactor designs and operating conditions with an emphasis on the impact of recycle stream. This enabled the students to model a more complex reactor while investigating how the reactor design was influenced by the up/down process streams. Students also simulated and analyzed the entire flowsheet<sup>4</sup> as part of Process Simulation and Design course. Process design optimization strategies were proposed and investigated by students, assessed against multiple economics and sustainability metrics.

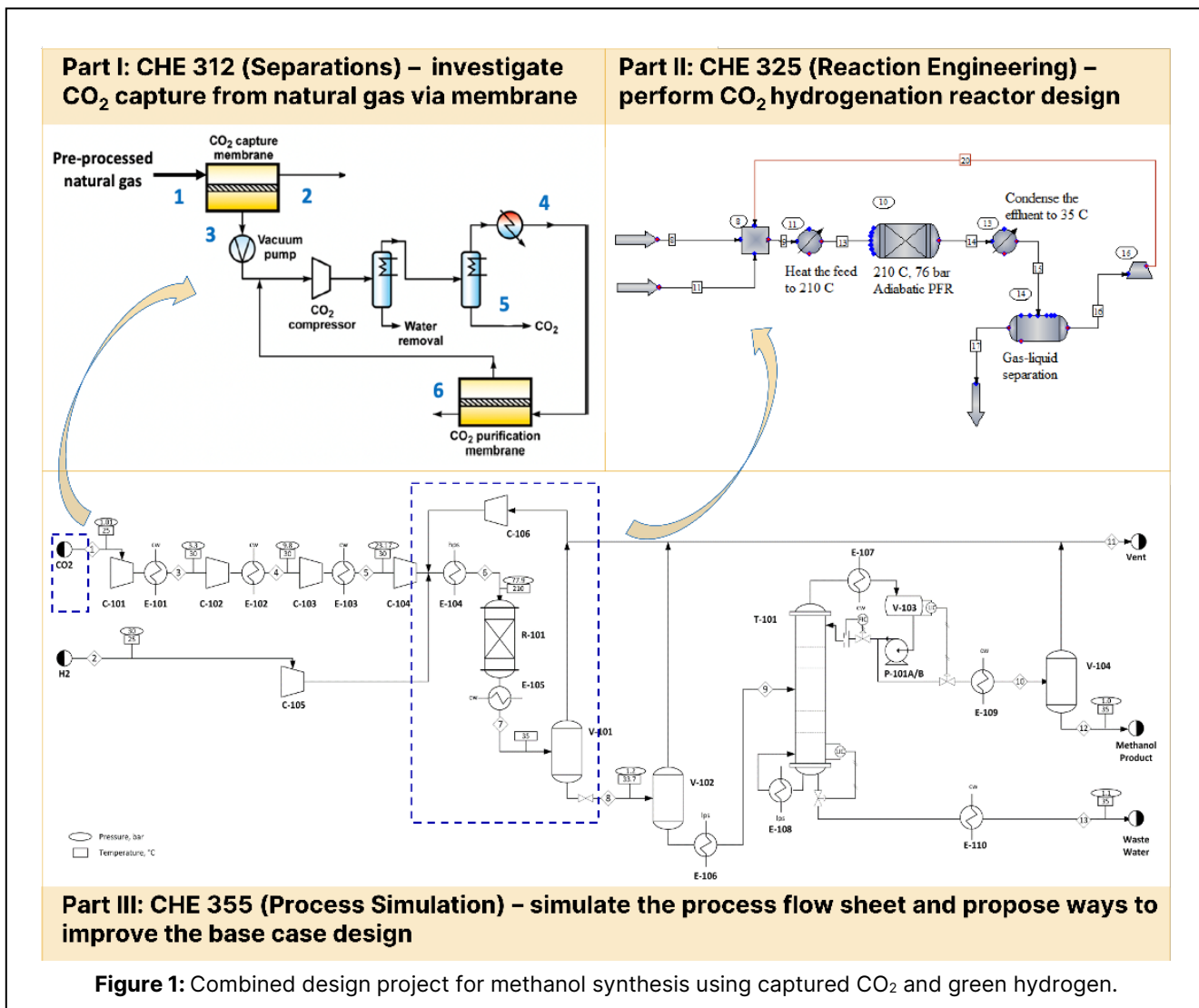


Figure 1: Combined design project for methanol synthesis using captured CO<sub>2</sub> and green hydrogen.

## Project 2: Ethylene Oxide Production

Fig. 2 illustrates the Spring 2023 combined project to design an ethylene oxide production plant via the direct oxidation of ethylene. The base case process design starts with compressing a mass amount of dry air which includes both nitrogen and oxygen. The multi-stage compressors are the major energy consumer in this process. The Separations course focuses on analyzing an alternative route in which oxygen is first separated from air via membranes before being sent to compressors. This enabled the students to quantitatively compare two process options, including a well-established process with high energy intensity and an emerging modular process where the large-scale commercialization is at an early stage. For the ethylene oxidation reactor, the focus of Reaction Engineering course, students compared the simulation using equilibrium-based reactor and kinetic reactor to understand the characteristics of this reaction. Students also searched for alternative catalysts from open

literature, such as the use of promoters on the Ag catalyst to improve selectivity to the ethylene oxide product. This enabled the students to learn the impact of reaction kinetics on the design and optimization of this reactor unit as well as the integrated process systems. Students simulated and optimized the entire flowsheet as the focus of Process Simulation and Design course. They also analyzed the bottleneck of base case process design based on cost breakdown and investigated systems-level integration of the upstream membrane units.

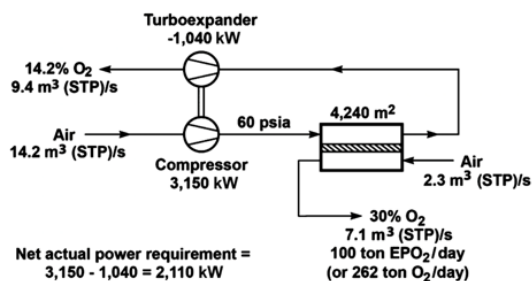
## IMPLEMENTATIONS AND OUTCOMES

This section presents several examples to showcase student implementations for the green methanol synthesis project (Fig. 1).

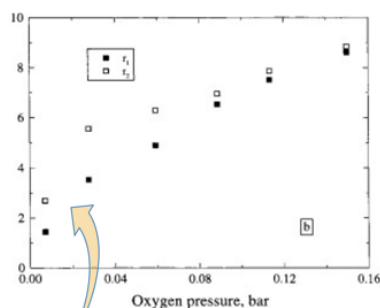
### Separations

The Separations design project was focused on simulating a hollow fiber membrane-based carbon capture

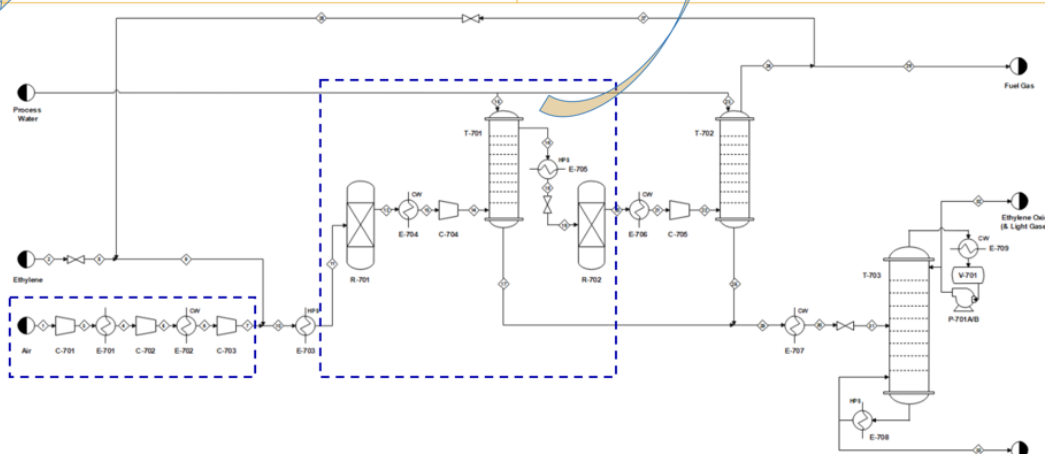
**Part I: CHE 312 (Separations) – investigate O<sub>2</sub> separation from air via membrane**



**Part II: CHE 325 (Reaction Engineering) – study the impact of catalyst and optimize reactor**



**Part III: CHE 355 (Process Simulation) – simulate the process flow sheet and propose ways to improve the base case design**



**Figure 2:** Combined design project for ethylene oxide production.

process, in conjunction with natural gas combined cycle (NGCC), i.e., primarily aimed at CO<sub>2</sub>/N<sub>2</sub> separation. The project involved analysis of the effects of intrinsic membrane parameters (CO<sub>2</sub> permeability and CO<sub>2</sub>/N<sub>2</sub> selectivity) on the overall system performance such as product quality, recovery, and membrane area. The membrane properties were chosen from prior experimental studies. Some preliminary calculations were expected to be done by students using the membrane separation principles covered during regular lectures. The students were also required to identify the input parameters (e.g., membrane performance, upstream/downstream pressure) which significantly affect the final product purity. Finally, the student groups were asked to identify improvement and possible modifications to the proposed system. Results from one representative group are shown in Fig. 3.

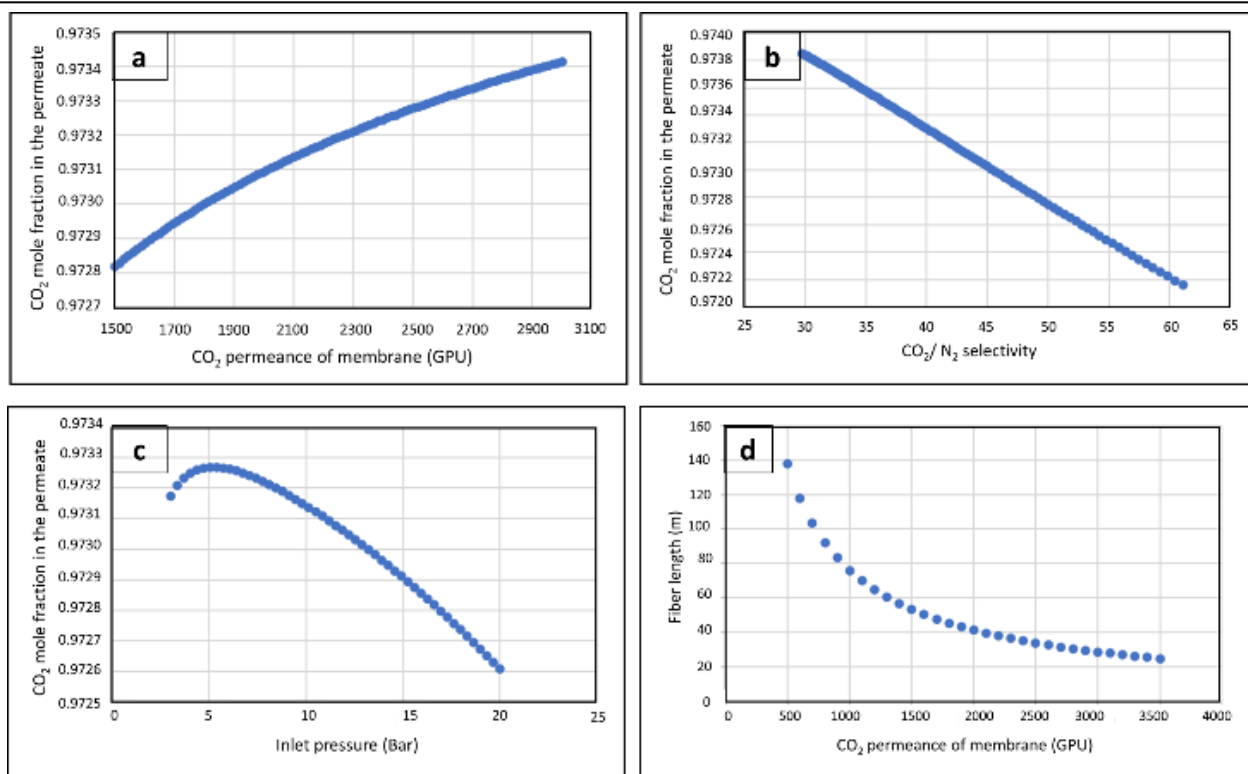
### Reaction Engineering

The Reaction Engineering design project section was focused on simulating the CO<sub>2</sub> hydrogenation reactor and investigating the impact of the recycle stream on reactor performance. The rate expressions for both methanol

synthesis and the reverse water gas shift reactions were provided along with the corresponding kinetic parameters. These kinetic expressions were implemented into CHEMCAD by the students and compared the reaction section - including reactor, heat exchanger, flash vessel, and recycle - modeled with these kinetics vs a reactor at equilibrium. Students were asked to report selectivity and yield, as well as fractional conversion as a function of the reactor volume to optimize the reactor sizing. Additionally, they were asked to propose a reactor improvement based on either economic or performance metrics. As an example of the results, one project group results are shown in Fig. 4. This specific group used profit as the metric by which to conduct their optimization and determined the reactor volume based on maximum profit.

### Process Design and Simulation

The Process Simulation and Design project section was focused on simulating the entire flowsheet using CHEMCAD, which included (multi-stage) compression of the CO<sub>2</sub> and H<sub>2</sub> feed streams, CO<sub>2</sub> hydrogenation reactor, and a series of separation units (flash and distillation



**Figure 3:** Effects of (a) CO<sub>2</sub> permeance, (b) CO<sub>2</sub>/N<sub>2</sub> selectivity and (c) Inlet pressure on permeate quality and (d) effect of CO<sub>2</sub> permeance on the fiber length.

columns). Students were also required to perform equipment sizing and calculate the equivalent annual operating cost (EAOC) to identify the process design bottlenecks. On this basis, students would propose and implement two major areas for process improvements. Fig. 5 presents an example of the improvements by one of the design groups, where they applied heat integration utilizing the high temperature reactor effluent to heat the inlet streams respectively to reactor and distillation column. Additionally, they added a membrane module to recover hydrogen from the vent stream. As this project assumed the use of green hydrogen, the raw material cost took up more than 80% of base case EAOC. The recovery of remaining H<sub>2</sub> would effectively reduce the cost. The students proposal of using a membrane for downstream gas separation also well justified the efficacy of this combined project to bring emerging modular intensified process technologies to the toolbox of next-generation chemical engineers. To address the trade-off between sustainability and economic competitiveness as illustrated by this green methanol process, one design group explored the impact of carbon tax.

## STUDENT FEEDBACK

After the second year of project implementation, surveys were conducted for both project cohorts of

students. The survey administered to the junior students who had just completed the project (2023 project on ethylene oxide production) focused on how the project helped them develop skills and apply technical content from the courses. The survey administered to the senior students, who had completed the joint project approximately one year prior (2022 project on methanol synthesis), focused on how the project prepared them for senior design.

### Junior Survey on Combined Project

The primary survey question investigated whether students felt the joint project helped them to practice material from each of the three courses, shown in Fig. 6.

Most students felt that the joint project helped them practice material from the Process Design course, with a strong majority answering “very much”. For the Separations and Reaction Engineering courses, however, responses were more distributed with approximately equal numbers of students responding 3 (neutral), 4, or 5 (very much).

Additionally, students were asked to compare their skills before and after working on the junior design project across a range of areas, including technical skills such as process design and optimization as well as skills

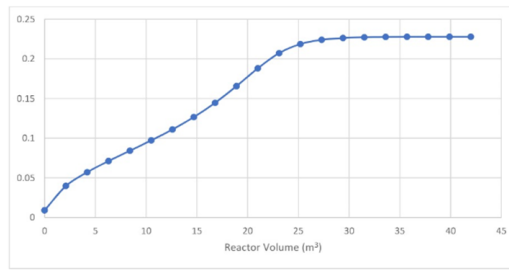


Figure 2: Base Case Methanol Yield vs. Reactor Volume

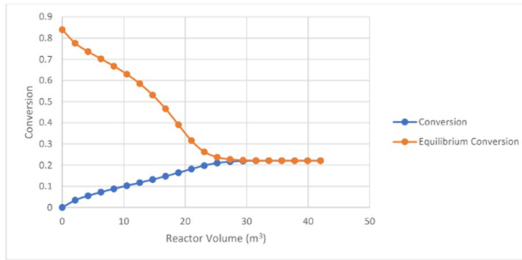


Figure 3: Base Case Carbon Dioxide Conversion vs. Reactor Volume

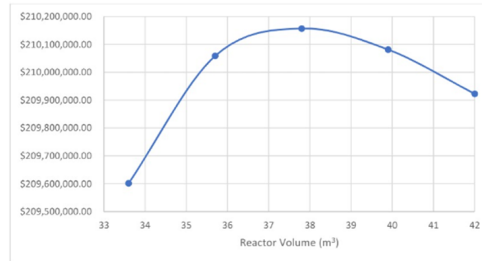
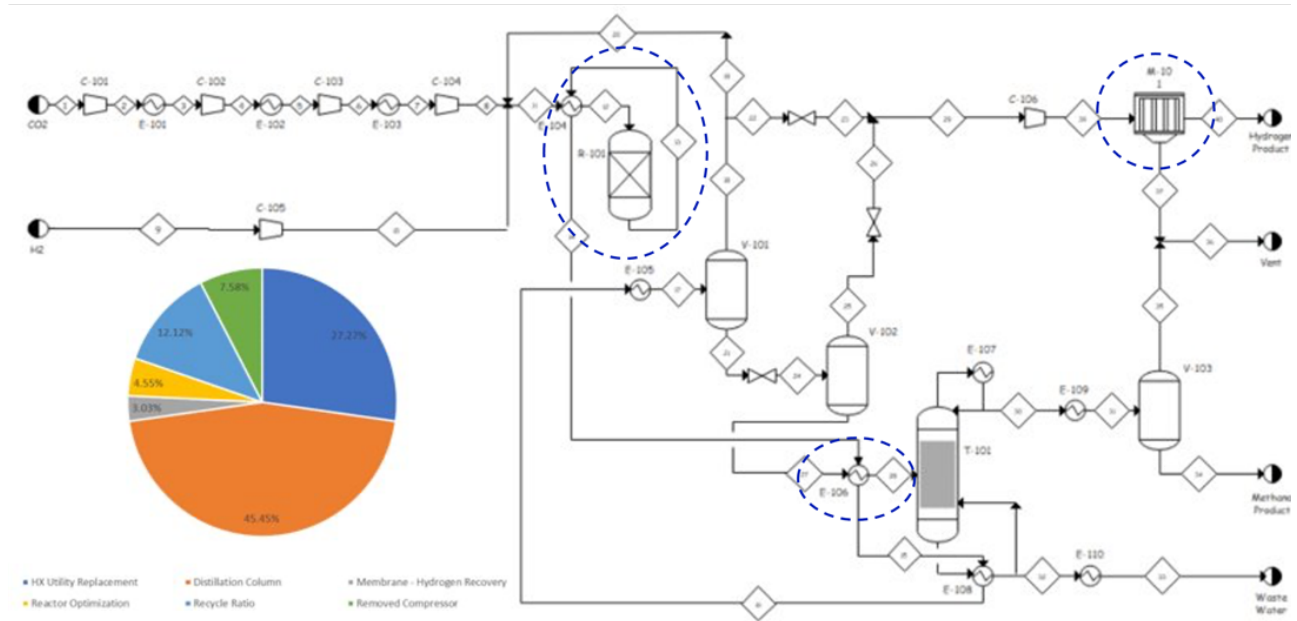


Figure 8: Closer Examination of "Profit" Plot to Find Ideal Reactor Volume

**Figure 4:** (Top) Methanol yield as a function of reactor volume, (bottom left) conversion and equilibrium conversion as a function of reactor volume, and (bottom right) process profit as a function of reactor volume to determine the optimum reactor size.



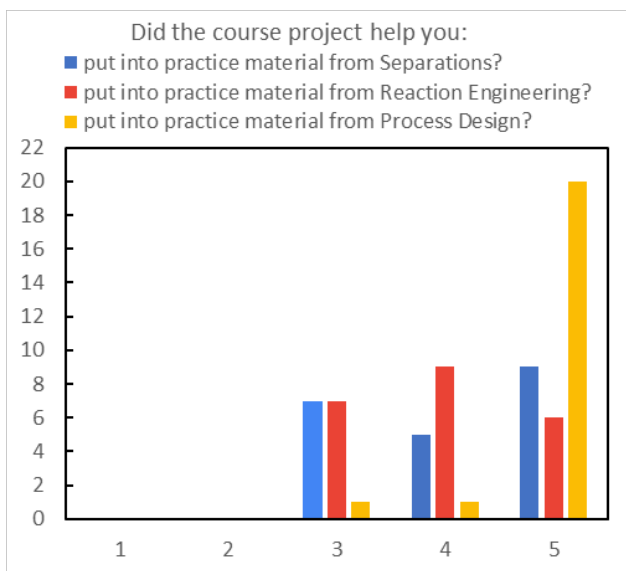
**Figure 5:** Process design improvements via heat integration and membrane-based hydrogen recovery.

such as teamwork, communication, and creativity. The results are shown in Fig. 7, with the left of each bar indicating before the project and the right of each bar indicating after the project.

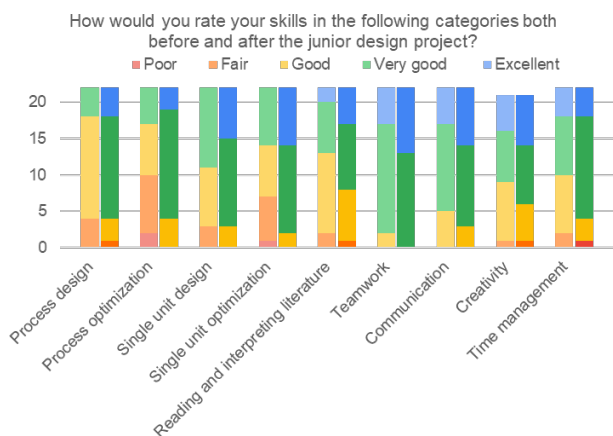
While all skills showed an improvement after participating in the design project (indicated by more green and blue and less red/orange/yellow), the most significant improvements were observed for the technical skills. In reflecting on the beginning of the class, at least half the

class rated themselves as having poor, fair, or good skills in process/single unit design and optimization. After participating in the project, the vast majority of the students reported their skills in these areas as very good or excellent. Improvements to group work, communication, and other similar skills were less drastic.





**Figure 6:** Responses from junior students to “Did the course project help you put into practice material from Separations (blue), Reaction Engineering (red), and Process Design (yellow)?” Responses were given on a scale of 1 (not at all) to 5 (very much).



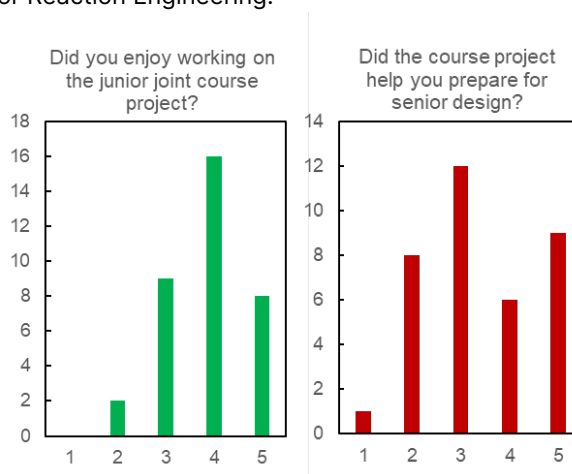
**Figure 7:** Responses from junior students on their skills in a variety of topic areas before (left, faded) and after (right, darker) working on the junior design project. Responses were given on a scale of poor, fair, good, very good, and excellent.

### Senior Survey on Preparation for Capstone Design

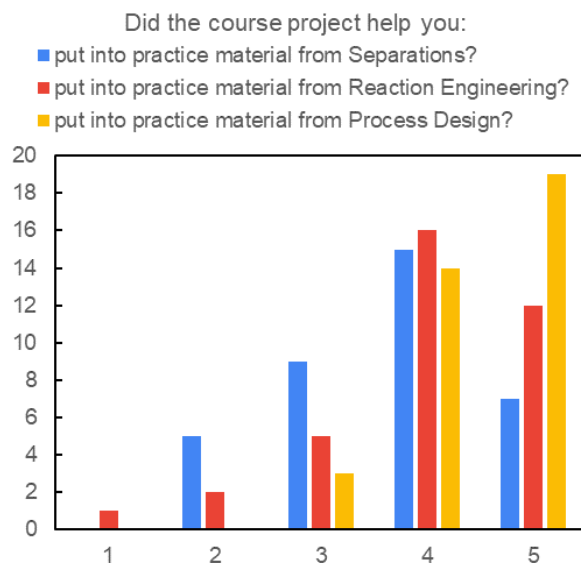
The senior students who participated in the first implementation of the joint project in Spring 2022 were surveyed about their experience in Spring 2023. This survey served to assess their perception of preparation for their senior design course, as well as their retrospective perception of how well the project helped them practice material from each of the junior level courses. Most students reported a neutral or positive experience of working on the project, however the perception of preparation for

senior design was mixed. Students reported an average of 3.4 (out of 5) indicating they felt that the junior course project somewhat prepared them for senior design.

Additionally, when asked if the joint course project helped students practice material from each of the three courses, a range of responses were given. Notably, the students reported that the project helped them practice course material to a lesser extent than the junior responses, indicating that the project modifications between implementation in 2022 and 2023 better connected the course content and project. Similar to the responses from the junior class above, the students found that the project helped them put into practice material from Process Design to a greater extent than Separations or Reaction Engineering.



**Figure 8:** Survey results for seniors, responses collected approximately one year after completing the joint junior course project. Scale of 1 (not at all) to 5 (very much).



**Figure 9:** Survey results for seniors, responses collected approximately one year after completing the joint junior course project. Scale of 1 (not at all) to 5 (very much).

## Select Student Comments

### Comments from seniors who took CHE 355 in Spring 2022 (surveyed a year later)

**Q → What was the most interesting/enjoyable part of the junior course project?**

**Response:** The integration of a single topic between all of our core classes (since it reflected more of a real-world process design project)

**Q → What was the most challenging part of the junior course project?**

**Response:** Finding a true optimum case because so many factors + process units depend on each other

### Comments from juniors who took CHE 355 in Spring 2023 (surveyed at the end of semester)

**Q → What was the most interesting and/or enjoyable part of the project?**

**Response 1:** The most interesting part of the project was finally integrating all of the information we've learned into one project instead of a bunch of little bits and pieces.

**Response 2:** Seeing how much money that can be saved with implementation of simple optimizations was very interesting.

## CONCLUDING REMARKS

In this paper, we have introduced our development of combined junior process design project across reaction engineering, separations, and process simulation and design. Two project examples are presented which strive to integrate the knowledge from different courses and highlight the increasingly important process design goals such as sustainability and decarbonization. It is worth highlighting that the implementation of such combined course projects can be greatly benefited by, while not restricted to the availability of, a specialized process design course at junior level. Based on student feedback, we aim to address the following key points to continuously improve the combined project in future semesters:

- Emphasize systems-level optimization and its interactions (or trade-offs) with single unit optimization.
- Emphasize process design and optimization with simultaneous considerations of multiple objectives (e.g., economics, carbon footprints, water usage).
- Arrange training workshops on Aspen and custom modeling to prepare students for senior design with an enriched set of simulation tools and skills.
- Integrate graduate student teaching practice to train next-generation process design educators.

## ACKNOWLEDGEMENTS

The authors acknowledge financial support from WV Research Challenge Grant (RCG23-009) and Department of Chemical and Biomedical Engineering at West Virginia University, United States.

## REFERENCES

1. Strachan, S. M., Marshall, S., Murray, P., Coyle, E. J., & Sonnenberg-Klein, J. (2019). Using Vertically Integrated Projects to embed research-based education for sustainable development in undergraduate curricula. *International Journal of Sustainability in Higher Education*, 20(8), 1313-1328.
2. Fletcher, A., & Boon, S. (2013). Proposed vertical integration of prior learning to support students undertaking Chemical Engineering Design. *Education for Chemical Engineers*, 8(2), e72-e85.
3. Pérez-Fortes, M., Schöneberger, J. C., Boulamanti, A., & Tzimas, E. (2016). Methanol synthesis using captured CO<sub>2</sub> as raw material: Techno-economic and environmental assessment. *Applied Energy*, 161, 718-732.
4. Merkel, T. C., Wei, X., He, Z., White, L. S., Wijmans, J. G., & Baker, R. W. (2013). Selective exhaust gas recycle with membranes for CO<sub>2</sub> capture from natural gas combined cycle power plants. *Industrial & Engineering Chemistry Research*, 52(3), 1150-1159.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Model Diagnostics for Equation-Oriented Models: Roadblocks and the Path Forward

Andrew Lee<sup>a,b,\*</sup>, Robert B. Parker<sup>c</sup>, Sarah Poon<sup>d</sup>, Dan Gunter<sup>d</sup>, Alexander W. Dowling<sup>e</sup>, and Bethany Nicholson<sup>f</sup>

<sup>a</sup> National Energy Technology Laboratory, Pittsburgh, PA 15236, USA

<sup>b</sup> NETL Support Contractor, Pittsburgh, PA 15236, USA

<sup>c</sup> Los Alamos National Laboratory, Los Alamos, NM 87545, USA

<sup>d</sup> Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

<sup>e</sup> University of Notre Dame, Notre Dame, IN 46556, USA

<sup>f</sup> Sandia National Laboratories, Albuquerque, NM 87185, USA

\* Corresponding Author: [andrew.lee@netl.doe.gov](mailto:andrew.lee@netl.doe.gov).

## ABSTRACT

Equation-Oriented (EO) modeling techniques have been gaining popularity as an alternative for simulating and optimizing process systems due to their flexibility and ability to leverage state-of-the-art solvers inaccessible to many procedural modeling approaches. Despite these advantages, adopting EO modeling tools remains challenging due to the significant learning curve and effort required to build and solve models. Many techniques are available to help diagnose problems with EO process models and reduce the effort required to create and use them. However, these techniques still need to be integrated into EO modeling environments, and many modelers are unaware of sophisticated EO diagnostic tools. To survey the availability of model diagnostic tools and common workflows, the U.S. Department of Energy's Institute for the Design of Advanced Energy Systems (IDAES) has conducted user experience interviews of users of the IDAES Integrated Platform (IDAES-IP) for process modeling. The interviews reveal a gap between the availability and utilization of model diagnostic tools driven primarily by a lack of awareness of and lack of standard interfaces among different tools. To address this gap, the IDAES team has developed a recommended workflow for integrating diagnostics into the model development process and an IDAES Model Diagnostics Toolbox that provides a standard interface for many of these best practices. This paper identifies barriers to the widespread adoption of diagnostic tools for EO models and reduces these barriers by providing a standard, user-friendly interface for many different tools.

**Keywords:** Modelling and Simulations, Optimization, Simulation, Pyomo, Education

## INTRODUCTION

Equation-Oriented (EO) modeling techniques are powerful tools for solving complex process engineering models, providing modelers with access to powerful gradient-based solvers and greater flexibility than more traditional sequential-modular tools [1]. Several EO-based modeling tools are now available for use, including ABA-CUS [2], AIMMS [3], AMPL [4], Aspen Custom Modeler [5], DAEPACK [6], GAMS [7], gPROMS [8], JuMP [9,10], and Pyomo [11].

However, EO-based tools come with several challenges, especially when applied to large models, as

model tractability significantly depends on the formulation and the provision of good initial guesses to the solver. The number of potential issues in EO models is extensive, and understanding how to identify and remedy them during model development requires significant expertise from the user. Some possible problems include:

- inconsistencies in units of measurement,
- numerical evaluation errors and singularities,
- structural singularities in the problem matrix [12,13],
- degenerate equations and rank deficiency in the

Jacobian of the model [14],

- poorly scaled variables and constraints, and
- bound violations and infeasible constraints.

These issues are easily overlooked when manually inspecting large, complicated models such as those required for sophisticated process systems.

In 2015, the U.S. Department of Energy established the Institute for the Design of Advanced Energy Systems (IDAES) to develop a next-generation platform for process modeling: the IDAES Integrated Platform (IDAES-IP) [15,16]. IDAES aims to make state-of-the-art EO modeling techniques accessible to the process modeling community. A significant advantage of the IDAES-IP is that it is built on Pyomo [11] and Python, allowing the development of algorithms that directly interrogate a model's structural and numerical properties.

## UNDERSTANDING CURRENT APPROACHES

To understand the challenges faced by modelers, the IDAES team started by performing a user experience study, interviewing a range of users of both IDAES-IP and Pyomo with varying backgrounds and levels of modeling experience. These interviews sought to discover each modeler's most common issues and the workflows and tools they used to help diagnose them. Additionally, the interviews sought to understand how modelers developed these workflows and learned about the different tools they applied to better inform the education of new modelers.

The interviews were structured as free-form conversations to capture the range and nuance of users' experiences. With encouragement, interviewees often expanded on their responses to an initial set of questions and provided relevant anecdotes. While much of each interview was open-ended, **Table 1** summarizes the responses to three interview questions. Note that the categorization of issues in responses to **Q1** is imperfect due to imprecise definitions of various modeling issues. For example, what one user may refer to as a "degeneracy" due to a flow rate approaching zero, another may refer to as a "scaling issue." In responses to **Q2** and **Q3**, "Basic 'model statistics'" refers to simple utility functions for identifying variables and constraints with useful properties, *e.g.*, variables close to (or exceeding) their bounds or constraints with large residuals, that are provided by the `idaes.core.util.model_statistics` module.

The most frequently identified issue was poor model scaling. This is partly because the IDAES-IP (and Pyomo) tools are designed to be as general as possible to cover a wide range of operating conditions. For best performance, modelers should specify scaling factors to tailor

the problem to the specific operating range of interest.

Another common issue IDAES-IP users encounter is infeasibility, usually due to bound violations, partly due to the modular nature of the IDAES model libraries. Modelers frequently rely on pre-built sub-models, *e.g.*, for thermodynamics, but often do not check that the ranges of applicability, and thus bounds, for those sub-models match the needs of their application.

A third common issue was numerical evaluation errors, such as divisions by zero or logarithms of negative numbers. Nearly a quarter of the interviewees reported numerical evaluation errors related to zero-flow conditions in their models. These frequently arise in process systems involving reactions where not all components are present in every process stream. Suppose care is not taken to eliminate these components in streams where they do not exist or to ensure a minimum positive lower bound on flow/concentration. In that case, these values can cause numerical errors in complicated thermodynamic expressions involving logarithms or divisions.

Perhaps surprisingly, degenerate equations were not a common issue identified by users. However, this might indicate a need for more awareness of degeneracies rather than their infrequent occurrence. Most process flowsheets include at least one recycle stream, and these can easily introduce degeneracies into a model unless care is taken to identify an appropriate tear stream. Degeneracies are often tricky to identify unless the modeler knows what to look for. Degenerate models have a family of feasible solutions; thus, the solver frequently reports "optimal solution" termination messages. Hence, if the modeler does not intentionally inspect the solver logs, they may miss that their model is degenerate.

Qualitatively, the interviews revealed that most users had no formal training in diagnosing modeling issues and that most learning had come through experience. As a result, there was little uniformity in the workflows and tools used, and the interviewees tended to only think about issues they had previously encountered. Additionally, while most modelers interviewed used a bottom-up approach, starting with small models and gradually adding complexity and resolving issues as they appeared, some of those interviewed used a top-down approach, building the full complex model first and then trying to work backward to find the resulting issues.

A second finding was that there were two types of modelers within the IDAES user base: those who focused on developing models for unit operations and those who generally constructed flowsheets from pre-existing unit models in model libraries. Modelers in the first category were generally more aware of the different modeling issues they might encounter and how to fix them. At the same time, those who used existing models tended to struggle with diagnosing modeling issues. One key aspect of this is awareness of the model equations; unit

**Table 1:** Questions asked and responses received during 13 user experience interviews.

**Q1.** What are the issues you most commonly encounter with equation-oriented modeling and optimization in IDAES?

Response	Number of users
Scaling issues	8
Infeasibility	4
Incorrect degrees of freedom	4
Evaluation errors	4
Initialization issues	3
Zero-flow issues	3
Not understanding library models	3
Degeneracy	1
Incompatible assumptions	1

**Q2.** How do you recognize when these problems occur in your model or workflow?

Response	Number of users
Solver log	8
Basic “model statistics”	4
Inspect Jacobian	2
Incidence graph analysis	1
Test failure	1
Physical sanity check	1

**Q3.** What tools do you use to diagnose the cause of the issue or propose a fix?

Response	Number of users
Basic “model statistics”	8
Degeneracy Hunter	8
Inspect Jacobian	4
Change solver options	4
Solver log	3
Incidence graph analysis	2
Ask for help	2
Parameter sweep	2
Pyomo model viewer	2
Solver callback	1
Singular value decomposition	1
Serialize model state	1

model developers were generally fully aware of the equations in their models as they had written most equations themselves, while model users typically treated the library models as black boxes and were thus unable to resolve modeling issues that arose from them.

Finally, while solver logs were the most common method of identifying that a problem had occurred, many interviewees were not familiar with the full content of these. They focused mainly on apparent signs such as the final solver status, number of iterations, and obvious warning messages. This practice reveals a need to educate modelers about the data provided by solvers and how it can be used to identify potential modeling issues. Indeed, many modelers interviewed assumed that once a solver reports an “optimal solution,” that their model is

well-behaved; it is crucial to educate modelers that just as one should always check that the solution of their model makes sense, they should also check the solver logs to ensure that no warnings were encountered.

In terms of tools used to help diagnose modeling issues, the modelers identified existing tools within the IDAES-IP toolset, Pyomo, or other Python packages (*e.g.*, the SciPy packages for linear algebra [17]) to diagnose most of the modeling issues encountered. These tools fell into three categories:

1. *Structural analysis* tools that analyze constraints, variables, and bounds without depending on specific variable values. These tend to be computationally cheap and do not require the model to be initialized.



2. *Numerical analysis* tools that analyze constraints, bounds, and their Jacobian at specific variable values. These tend to be slightly more computationally expensive and require that all variables have values from a previous solution or initialization.
3. *Advanced techniques*, such as singular value decomposition or meta-modeling techniques, to identify irreducible degenerate sets. These tend to be more computationally expensive.

However, as these tools were scattered across different packages with different interfaces and output formats, no single modeler who was interviewed was aware of all the tools and how to use them. When asked about other tools, the most common response was that modelers were unaware of them, followed by comments about the lack of user-friendliness of some tools.

The interviews reveal that the main barrier to effective debugging of EO models was awareness of different issues modelers might encounter, the tools available to diagnose them, and methods to remedy them when they occur. To address this, the IDAES team developed and documented a comprehensive workflow and set of best practices for diagnosing modeling issues, accompanied by an integrated toolbox interfacing with most of the existing tools mentioned by the interviewees.

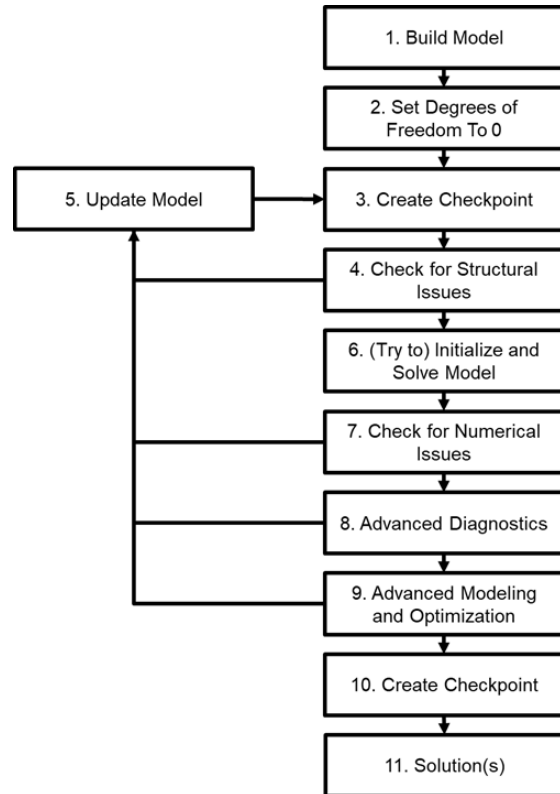
## PROPOSED DIAGNOSTICS WORKFLOW

The interviews revealed that issues can arise at many points during the modeling process. Therefore, modelers must continuously iterate between model development and diagnostic testing. Almost all the modelers interviewed agreed that the best approach to model development was to start from the bottom up, starting from simpler sub-models and gradually adding complexity, checking for and resolving issues at each step. In this way, the source of the problems can be easily traced to the most recent changes made to the model, and the smaller these changes, the easier it is to identify the root cause. As such, model development and diagnosis should always be viewed as an incremental process of gradual development and refinement, not a “one-shot” fix-all operation.

Additionally, the consensus across the modelers interviewed was that diagnosis should start with a “square” simulation model with no degrees of freedom; all optimization models derive from an underlying deterministic model by releasing some degrees of freedom. Adding degrees of freedom to the problem can often make it easier to solve, but this can obscure issues in the model. If the simulation model is ill-posed or poorly formulated, then these issues remain in the optimization model even if they are not apparent.

Based on this incremental, iterative approach and

the “hierarchy” of diagnostic tools, a sequence of common steps for model development was proposed, which is shown in **Figure 1** and explained below.



**Figure 1.** Proposed Model Diagnostics Workflow.

1. The workflow begins with an existing model, which should be the simplest possible model to start with (or an existing model that has been thoroughly tested for issues).
2. Next, the degrees of freedom necessary to form a square simulation model should be fixed. If the modeler is uncertain of which degrees of freedom to choose, they should take their best guess and allow the subsequent checks to identify any potential singularities.
3. As the model development and diagnosis process is iterative and often involves trial and error, modelers should always take the time to create a checkpoint so that they can easily revert changes to earlier states in case the current thrust of development becomes a dead-end. This is especially important results are being generated and recorded in parallel with the model development; there is nothing more frustrating than to discover that you can no longer reproduce a previous set of results with no idea of how or why this occurred. (Tip: version control systems such as git make it much easier to systematically track incremental model changes during

iterative diagnostics and refinement.)

4. When checking for potential modeling issues, the modeler should first apply structural analysis tools. As these only rely on model structure, they do not require an initial guess for the model solution. Fixing these structural issues before calling a solver decreases the chance of a critical solver failure.
5. Modelers should resolve issues as they are identified and should generally only make changes to address one issue at a time. Issues can often be interwoven, and fixing one issue may create or reveal other issues (or resolve other known issues). As each change is made to the model, modelers should return to the top of the workflow and follow all the steps again to make sure they are not accidentally introducing new issues.
6. Only once all (known) structural issues are resolved should modelers attempt to solve their model (initializing as necessary). Modelers should always review the solver output logs, as these can contain helpful information for diagnosing modeling issues. Hopefully, the solver will return a solution, even if it is infeasible, which can be used to continue the model diagnostics workflow. (Tip: If a solution is not returned, then the modeler can re-run the solver with a lower iteration limit or use a block triangular decomposition solve (see Section 3.6 of [12]) to obtain a partial solution.)
7. Once an initial solution is obtained, even if it is infeasible, the modeler can check for potential numerical issues at the current state. Again, problems should be resolved as they are identified, and modelers should return to the top of the workflow after each change. Numerical issues inherently depend on the current state of the model. As optimization involves exploring the behavior of the model over an entire design space, it is essential to perform numerical checks across this design space.
8. If issues remain (*e.g.*, failure to converge or signs of poor behavior in the solver logs), modelers should then apply the advanced diagnostics tools to try to find the root cause of the issues. As these techniques are computationally intensive, they are reserved for the final step to address challenging issues.
9. Once a reliable model has been identified, modelers can then move on to optimization studies and other more complex modeling tasks. Modelers should continue monitoring the solver logs for signs of potential issues and should apply the model diagnostics workflow if these signs appear.
10. Finally, as solutions and results are generated,

modelers should always maintain a checkpoint and log of the model state (*e.g.*, git commits) to ensure that they can reproduce these results in the future and to provide a point of comparison in case future changes to the model change these results.

Finally, modelers should keep in mind that their model may be used by others in the future who may not have the same level of familiarity with the model equations and formulation; this is especially important for developers of model libraries such as those in IDAES-IP. As the original modeler has the most experience with the formulation of the model, they are best suited for ensuring the model is as robust as possible to minimize the challenges faced by future users.

## IDAES DIAGNOSTICS TOOLBOX

Whilst the workflow described above provides modelers with a series of steps to take when developing their models and diagnosing any issues that arise, they also need access to a suite of tools to perform these checks. To limit the number of methods and interfaces modelers need to know, a new Diagnostics Toolbox has been added to the IDAES-IP (v2.2.0, released September 2023) [16], which incorporates tools to check for a wide range of common modeling issues.

To assist users with the model diagnostics workflow, the Diagnostics Toolbox guides them through the diagnostics process through a series of “reporting methods.” These reporting methods call a series of tools to check for different types of issues and then provide the modeler with a summary of the issues found and a set of next steps for further investigation. Each of these “next steps” corresponds to a “display” method to call one of the underlying methods and provides a detailed description of the issue and, if possible, its source.

Currently, the Diagnostics Toolbox contains three reporting methods—one for structural issues, one for numerical issues, and one for irreducible degenerate sets—and approximately 22 associated display methods. To assist modelers with identifying the most likely causes of model failures, problems are further divided into “warnings” representing issues that will most likely result in solver issues or incorrect results (*e.g.*, singularities and degeneracies), and “cautions” that warrant checking but may be expected model behavior (*e.g.*, variables which do not appear in any constraint). The Diagnostics Toolbox may be imported from IDAES as:

```
from idaes.core.util import DiagnosticsToolbox
```

The remainder of this section gives a brief overview of the structural and numerical issues that can be diagnosed by the Diagnostics Toolbox.

### Structural Issues

The Diagnostics Toolbox contains three methods that check a model for unexpected variable status:

- variables which do not appear in any constraints,
- variables which are not declared as part of the model being diagnosed,
- and variables with a value fixed to exactly 0.

If any such variables are encountered, the modeler is cautioned. Additionally, the Diagnostics Toolbox contains methods to check constraints and objectives for consistent units, verify that the over and under-constrained subsystems of the model's Dulmage-Mendelsohn decomposition [12,13] are empty, and check for potential evaluation errors. If any of these checks fail, a warning is displayed.

### Numerical Issues

The Diagnostics Toolbox contains several checks performed with simple variable and expression evaluations. These check for:

- variables with values at or outside their declared bounds,
- variables with no values,
- variables with values near zero,
- variables with extremely large or small values,
- variables with values approaching their bounds,
- and constraints with large residuals.

These checks display a warning or a caution depending on the magnitude of *e.g.*, variable value or bound proximity/violation. A variable with no value always yields a caution. The Diagnostics Toolbox also contains methods that analyze the model's Jacobian matrix to detect the following indicators of ill-conditioning:

- values in the Jacobian matrix with extremely large or small values,
- rows and columns in the Jacobian matrix with extreme 2-norm values,
- and condition number of the Jacobian.

Whether a warning or a caution is displayed again depends on the magnitude of the value detected.

### Advanced Diagnostics Tools

For cases where the above tools are not sufficient to fully diagnose the issues in a model, the Diagnostics Toolbox also contains an implementation of Degeneracy Hunter [14] and an interface for performing Singular Value Decomposition (SVD) analysis on the model (by default using the SciPy linear algebra libraries [17], but with support for modeler-provided SVD routines).

## CASE STUDY

To better demonstrate the application and utility of the Diagnostics Toolbox, consider the following toy problem with variables  $v_1$ [units:  $m$ ],  $v_2$ [ $m$ ],  $v_3 \in [0,5]$ ,  $v_4$ ,  $v_5 \in [0,1]$ ,  $v_6, v_7 \in [0m, 1m]$ , and  $v_8$ , and constraints:

$$v_1 + v_2 = 10 \quad 1)$$

$$v_3 = v_4 + v_5 \quad 2)$$

$$2v_3 = 3v_4 + 4v_5 + v_6 \quad 3)$$

$$v_7 = 1 \times 10^{-8}v_1 \quad 4)$$

One common issue facing modelers relying on pre-built EO models is identifying suitable degrees of freedom to fix for their application. Poor choices can often result in infeasible models or structural singularities. To demonstrate this, consider setting the following conditions for fixed variables:  $v_4 = 2$ ,  $v_5 = 2$ , and  $v_6 = 0$ .

While the above problem may appear relatively simple and involve only linear constraints, the model will yield an infeasible result after fixing the specified degrees of freedom and attempting to solve it. This is because there are several issues in the model as posed. As this model is small, an experienced modeler can likely identify most or all these issues by manually inspecting the variables and constraints; however, it will likely take a few minutes. The Diagnostics Toolbox, however, can automatically detect all these issues in under a second and direct the modeler to the root causes.

Following the diagnostics workflow, the first step is to check the model for structural issues. This is done by calling the `report_structural_issues` method from the Diagnostics Toolbox. This executes the series of checks described in the previous section, which reveals two warnings (critical issues) and two cautions.

*Structural warning 1:* Checking unit consistency reveals an issue in constraint 1):  $v_1$  and  $v_2$  were assigned units of meters; however, the constant in the constraint was not assigned units. If this constant was in units other than meters, then the solution to the model would be incorrect. This should be corrected by rewriting constraint 1) to include the units for the constant and applying any necessary unit conversion.

*Structural warning 2:* Performing a Dulmage-Mendelsohn partition of the incidence matrix reveals a structural singularity. After fixing  $v_4$ ,  $v_5$ , and  $v_6$ , constraints 2) and 3) include only a single free variable. Similarly, constraints 1) and 4) include three free variables, and thus this sub-system is under-constrained. Addressing this issue requires the modeler to rethink the degrees of freedom they have chosen or the constraints they have chosen to write. The Dulmage-Mendelsohn partition helps the modeler choose degrees of freedom, as variables that are candidates for degrees of freedom must be in the under-constrained subsystem.

*Structural caution 1:* A simple inspection routine reveals that  $v_8$  does not appear in any constraint. This could be considered a modeling error. Either it was meant to be included in a constraint, but a mistake was made, or this variable is extraneous and should be removed from the model definition.

*Structural caution 2:*  $v_6$  has been assigned a value of exactly 0. While this will not cause an issue in the current model, modelers should be careful when assigning values of zero to ensure this will not cause numerical issues (e.g., appearing in a denominator or logarithm).

For this example, assume that the units of the constant in constraint 1) were meters and that the modeler decides to use  $v_2 = 5$  instead of  $v_4 = 2$ . After reviewing the two cautions,  $v_8$  should be removed from the model, while  $v_6$  can remain fixed to 0 for now. After making these changes and confirming that the structural issues have been resolved, the next step is to try to solve the model. The solver reports that the model is infeasible, so the next step is to check the model for numerical issues by calling the `report_numerical_issues` method from the Diagnostics Toolbox. This will execute the second set of checks described above, which will reveal two additional warnings and four cautions.

*Numerical warning 1:* At least one of the constraints has a large residual. The specific constraint may depend on the solver used but will be constraints 2) or 3). This should not be unexpected due to the infeasible solution returned by the solver.

*Numerical warning 2:* Iterating over all the variables in the model and comparing their values to any bounds specified will reveal that  $v_3$  and  $v_5$  have values at or outside their bounds. In the case of  $v_3$ , this is because the value required to satisfy both constraints 2) and 3) lies below the lower bound of 0, while for  $v_5$  the issue is that it was fixed to a value outside of its bounds (it was assigned a value of 5 but had bounds of [0,1]). Addressing these issues requires the modeler to consider why these bounds were placed on the model in the first place (e.g., to indicate a range over which an empirical correlation is reliable or to protect against a numerical singularity) and to decide whether the bounds can be relaxed or if changes to the model constraints are required.

*Numerical cautions 1-4:* All four cautions arise because  $v_7 = 5 \times 10^{-8}$ , which is likely close to the solver tolerance. Additionally, while it does not violate the bounds, this value is very close to the lower bound. Analysis of the Jacobian will reveal a large value for  $v_1$  in constraint 4). This points to poor scaling of  $v_7$  and constraint 4) and addressing this will likely improve solver performance.

As mentioned previously, an experienced modeler would likely discover most of these issues by manual inspection. However, this becomes more challenging for more complex (and especially nonlinear) models or modelers with less experience. The Diagnostics Toolbox runs

these checks automatically and quickly, providing the modeler with immediate guidance on which constraints and variables to investigate to resolve any issues observed with solver behavior or solutions.

A full demonstration of applying the Diagnostics Toolbox to this example can be found here:

[https://idaes-examples.readthedocs.io/en/latest/docs/diagnostics/diagnostics\\_toolbox\\_doc.html](https://idaes-examples.readthedocs.io/en/latest/docs/diagnostics/diagnostics_toolbox_doc.html)

## ONGOING WORK

Since the release of the IDAES Diagnostics Toolbox, the team has been working to advertise these capabilities to users and solicit feedback on how they can be improved. The IDAES team have begun using these tools to debug modeling issues, and future work will involve a second round of user interviews to determine whether the tools have been successful. The team is also continuing to identify additional modeling issues and tools to diagnose them, as well as visualization tools to improve user-friendliness and aid in interpreting results.

Anecdotally, initial responses have been uniformly positive, and including a handful of users who have successfully applied the tools in practice during the past few months. Users have reported that the Diagnostics Toolbox has greatly reduced the time required to get their models to solve. One user reported that the toolbox helped them resolve a structural singularity due to mis-specified initial and boundary conditions in a discretized PDE model that they described as “like searching for a needle in a haystack.”

However, diagnosing issues with nonlinear optimization problems remains challenging with many open questions. To quote an experienced researcher in process systems engineering consulted for this study:

*Automatically diagnosing arbitrary issues with nonlinear optimization problems is a Holy Grail of Operations Research.*

While identifying the presence of an issue in a model can be automated, tracing this back to the true root cause and identifying a fix often requires a degree of engineering knowledge that can only be gained by experience. Documentation to help users in this process is still required, and a long-term goal for IDAES is to create a library of examples for resolving different issues.

## CONCLUSIONS

The IDAES Diagnostics Toolbox and workflow have been developed to assist modelers with addressing issues encountered when developing, debugging, and solving equation-oriented models. A standard workflow for systematically identifying issues has been created.



This workflow is supported by a toolbox that includes methods for automatically detecting common modeling issues and identifying the associated variables and constraints for the modeler. This workflow and toolbox will significantly reduce the time, effort, and expertise that modelers require for diagnosing issues when working with large, complicated models.

## ACKNOWLEDGEMENTS AND DISCLAIMER

This work was conducted as part of the U.S. Department of Energy's Institute for the Design of Advanced Energy Systems (IDAES), supported by the Office of Fossil Energy and Carbon Management's Simulation-based Engineering/Crosscutting Research Program.

This work was supported by the U.S. Department of Energy through the Los Alamos National Laboratory (LANL) LDRD program and the Center for Nonlinear Studies (CNLS). Los Alamos National Laboratory is operated by Triad National Security, LLC, for the National Nuclear Security Administration of the U.S. Department of Energy (Contract No. 89233218CNA000001).

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525.

This project was funded by the Department of Energy, National Energy Technology Laboratory an agency of the United States Government, through a support contract. Neither the United States Government nor any agency thereof, nor any of its employees, nor the support contractor, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## REFERENCES

1. Dowling A. W., Biegler L. T., A Framework for Efficient Large Scale Equation-Oriented Flowsheet Optimization. *Comput. Chem. Eng.* 72:3-20 (2015) <https://doi.org/10.1016/j.compchemeng.2014.05.013>
2. Tolsma J. E., Barton P. I., Process Simulation and Analysis with Heterogeneous Models. *Comput Aided Chem Eng.* 9(C):309-314 (2001) doi:10.1016/S1570-7946(01)80047-X
3. Bisschop J., AIMMS Optimization Modeling (2006) [http://download.aimms.com/aimms/download/manuals/AIMMS3\\_OM.pdf](http://download.aimms.com/aimms/download/manuals/AIMMS3_OM.pdf)
4. Fourer R., Gay D. M., Kernighan B. W., A Modeling Language for Mathematical Programming. *Manage Sci.* 36(5):519-554 (1990) doi:10.1287/mnsc.36.5.519
5. ASPENTech. ASPEN Engineering Suite
6. Tolsma J., Barton P. I., DAEPACK: An Open Modeling Environment for Legacy Models. *Ind Eng Chem Res.* 39(6):1826-1839 (2000) doi:10.1021/ie990734o
7. G.D. Corporation, General Algebraic Modeling System (GAMS) Published online (2013)
8. Barton P., Pantelides C., gPROMS - a Combined Discrete/Continuous Modelling Environment for Chemical Processing Systems. *Simul Ser.* 25:25-34. (1993) <https://yoric.mit.edu/gproms-combined-discretecontinuous-modelling-environment-chemical-processing-systems>
9. Bezanson J., Edelman A., Karpinski S., Shah V. B., Julia: A Fresh Approach to Numerical Computing. *SIAM Rev.* 59(1):65-98 (2017) doi:10.1137/141000671
10. Dunning I., Huchette J., Lubin M., JuMP: A Modeling Language for Mathematical Optimization. *SIAM Rev.* 59(2):295-320 (2017) doi:10.1137/15M1020575
11. Bynum M. L., Hackebeil G.A., Hart W. E., Laird C. D., Nicholson B. L., Siirola, J. D., Watson J.-P., Woodruff D. L., Pyomo - Optimization Modeling in Python. 3rd Edition. Springer (2021)
12. Parker R. B., Nicholson B. L., Siirola J. D., Biegler L. T., Applications of the Dulmage-Mendelsohn Decomposition for Debugging Nonlinear Optimization Problems, *Comput. Chem. Eng.* 178:108383 (2023) <https://doi.org/10.1016/j.compchemeng.2023.108383>
13. Dulmage, A. L., Mendelsohn, N. S., Coverings of bipartite graphs. *Can. J. Math.* 10:517-534. (1958) doi:10.4153/cjm-1958-052-0
14. Dowling A. W., Biegler L. T., Degeneracy Hunter: An Algorithm for Determining Irreducible Sets of Degenerate Constraints in Mathematical Programs. *Comput. Aided Chem. Eng.* 37:809-814 (2015)
15. Lee A., Ghouse J. H., Eslick J. C., Laird, C. D., Siirola J. D., Zamarripa M. A., Gunter D., Shinn J. H., Dowling, A. W., Bhattacharyya D., Biegler L. T., Burgard A. P., Miller D. C., The IDAES Process Modeling Framework and Model Library—Flexibility



for Process Simulation and Optimization. *J. Adv. Manuf. Process.* 3(3):e10095 (2021)

<https://doi.org/10.1002/amp2.10095>

16. IDAES. <https://github.com/IDAES/idaes-pse>
17. Virtanen P., Gommers R., Oliphant T. E., Haberland M., Reddy T., Cournapeau D., Burovski E., Peterson P., Weckesser W., Bright J., van der Walt S. J., Brett M., Wilson J., Millman K. J., Mayorov N., Nelson A. R. J., Jones E., Kern R., Larson E., Carey C. J., Polat I., Feng Y., Moore E. W., VanderPlas J., Laxalde D., Perktold J., Cimrman R., Henriksen I., Quintero E. A., Harris C. R., Archibald A. M., Ribeiro A. H., Pedregosa F., van Mulbregt P., and SciPy 1.0 Contributors, SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17(3), 261-272 (2020)

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Analysis of Chemical Engineering Curricula Using Graph Theory

Blake R. Lopez<sup>a</sup> and Victor M. Zavala<sup>a\*</sup>

<sup>a</sup> Department of Chemical Engineering and Biological Engineering, University of Wisconsin-Madison

\* Corresponding Author: [victor.zavala@wisc.edu](mailto:victor.zavala@wisc.edu)

---

## ABSTRACT

Chemical engineering is a highly complex interconnected major. Just as chemical engineers have broken complex processes into unit operations, the chemical engineering curriculum has been broken up into courses. The organization of these courses vary among institutions and are based on years of prior teachings and research. Despite this, there have been calls to reevaluate the curriculum from both industry and academia. We propose a graph-based representation of curricula in which topics are represented by nodes and topic dependencies are represented by directed edges forming a directed acyclic graph. This enables using graph theory measures and tools to provide formal ways of evaluating a curriculum. Additionally, the abstraction is readily understandable meaning conversations between instructors regarding the curriculum can occur within a department and even across institutions. This abstraction is explained with a simplified curriculum and applied to the undergraduate chemical engineering curriculum at University of Wisconsin-Madison. Highly and lowly connected topics are identified and approaches for grouping the topics into modules are discussed.

---

**Keywords:** Education, Curriculum, Graph Theory

## INTRODUCTION

The early days of chemical engineering focused on tools such as unit operations, material and energy balances, process control, transport phenomena, and more to develop petrochemicals. Since then, these tools have been increasingly applied to more applications, and chemical engineers can be found in diverse industries and roles. Climate change has also placed pressure on industries to decrease the environmental impact of their products and processes. However, there are proven difficulties in embedding sustainability in engineering curricula, and industry has identified ability shortages in graduating chemical engineers [1]. For these reasons and more, there have been calls to evaluate the chemical engineering curriculum, and some departments have elected to no longer seek accreditation to allow students more personalized paths towards different careers [2].

Process design is often treated as the culmination of the undergraduate chemical engineering curriculum. This course draws on knowledge from previous chemical engineering courses including transport, balances,

controls, and more. Students are often reminded of what they learned in previous courses and are for the first required to pull knowledge from multiple previous courses. This means students are likely seeing the interconnectivity of the curricula for the first time at the very end of their education. Additionally, process design tends to focus on the design of a single process typically within the chemicals or petrochemical sector. This means that students interested in other fields of such as biological systems or semiconductors may be less motivated in this course. For these reasons, there is discourse on implementing systems thinking and design throughout the curriculum to instead tying the curriculum together at the very end of a student's learning experience.

The ability shortages and the changing field of chemical engineers means the curriculum needs to evaluate to improve outcomes. There are some tools to analyze curricula, such as curriculum prerequisite maps to identify bridge and source-hub courses, tree-structured topic modeling, and concept maps [3,4]. Concept maps have been applied to chemical engineering courses and curriculum to enhance student learning and explore the

ontology of topics covered [5]. Other work has applied statistical tools to study how chemical engineering curricula has changed focus and analyzed different sequencing of transport courses at institutions [6].

We propose a graph-based representation in which topics are represented by nodes and dependencies between topics are represented by directed edges. Unlike in concept maps, the directed edges do not have words attached to them. This enables the use of graph tools guide curriculum analysis with numerical measures to formalize discussion in improving a curriculum.

## GRAPH ABSTRACTION

To explain graph theory applied to a curriculum, a simplified Curriculum A will be used. Curriculum A is composed of six courses which are named ABC 101-106. These courses depend on one other through prerequisites. For example, ABC 101 is a prerequisite to complete ABC 103. The full list prerequisites and the credits for each course can be found in **Table 1**.

**Table 1:** Courses making simplified Curriculum A.

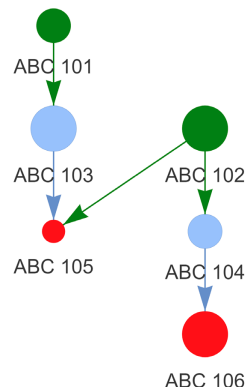
Course	Credits	Prerequisite(s)
ABC 101	3	None
ABC 102	4	None
ABC 103	4	ABC 101
ABC 104	2	ABC 102
ABC 105	2	ABC 102, ABC 103
ABC 106	3	ABC 105

This table looks like one a student would find in their handbook or when registering for the courses for the term. This table can be represented with a graph abstraction in which courses are represented by nodes and directional edges connecting nodes represent dependencies as seen in **Figure 1**. The edges are directional to represent the sequence of the courses and follows the thinking of "I have to take this course before I take that course." With the nature of course progression there are no cycles. For example, ABC 103 would not and cannot be a prerequisite for ABC 101 because then there would be no way take either of them. This makes the graph abstraction both directional and acyclic. Within this representation, the number of credits for each course is reflected by the node size.

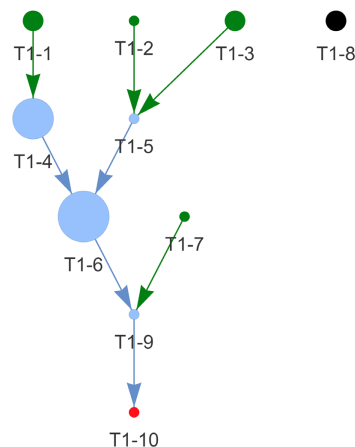
The course level graphical abstraction can allow for the scheduling one's term and visualizing how courses interact with one another. However, these dependencies rely on the accuracy of the prerequisites which could have been established years prior and the courses and, more broadly, the curriculum could have changed since. For example, students may be encouraged by their adviser to take ABC 103 and ABC 104 in the same term. From the graph abstraction, this need is not clear perhaps because ABC 103 has topics that are required for

ABC 106 despite not showing up in the prerequisites. This means looking more closely into the courses and their interconnectivity is required to fully understand the curriculum.

To look more closely at the curriculum, we can develop topic level graphs for each course within Curriculum A. This could be generated from the schedule of topics in a syllabus, recommended textbook readings, and more. The selection of topics and their connectivity takes some expertise and refinement, but the generation of the abstraction can provide invaluable insights. The graph abstraction for ABC 101 is presented in **Figure 2**.



**Figure 1.** Course level graph abstraction of Curriculum A in which courses are represented as nodes and directed edges represent course dependencies. Node size represents the credit hours for the course. Courses without prerequisites are green, ternary courses are red, and courses with both prerequisites and subsequent dependencies are blue.



**Figure 2.** Graph abstraction of course ABC 101 in which nodes represent topics and topic dependencies are represented as directed edges. Node size represents the relative resources use in terms of credit hours.

In this topic level graph abstraction of ABC 101, topics are represented by nodes and the size of the nodes

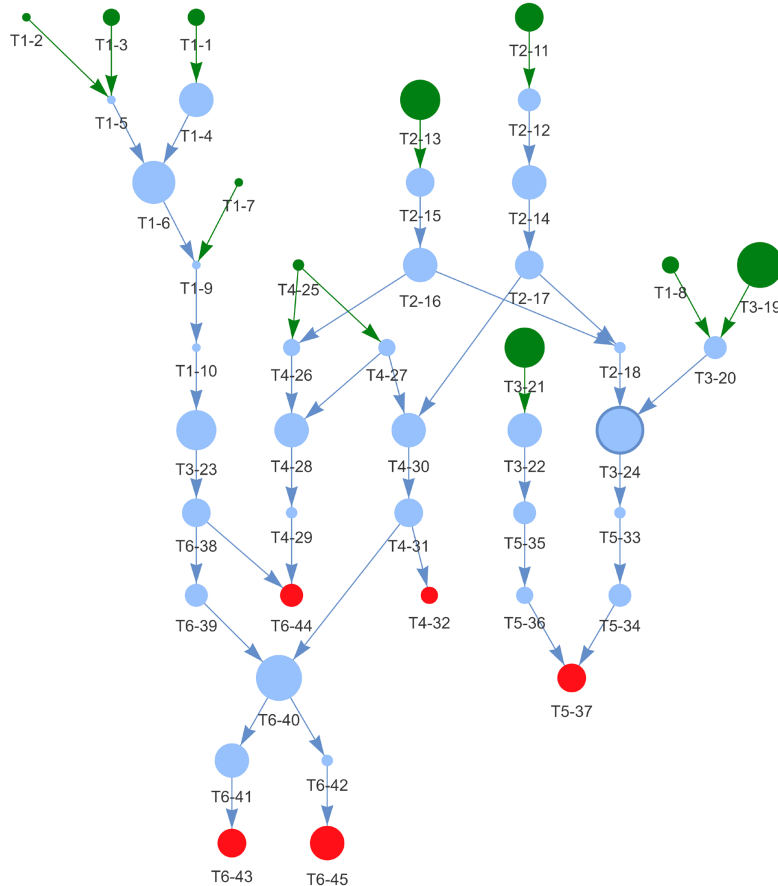
represent the relative time spent covering the topic in terms of the credit hours for the course. This means the sum of all the node sizes is the number of credit hours for the course. Topic connectivity is represented by the directional edges. A topic without any preceding topics, such as T1-1, means it can be taught without any other prior knowledge. These can be thought of as introductory topics like how one would think about an introductory course. A topic with preceding topics, such as T1-6, means it requires the preceding connected topics (T1-4 and T1-5) to be taught. This follows the logic of "I need to know numbers and counting to learn addition." These dependencies can also be used to help plan the course by noticing topic T1-10 requires all other topics besides T1-8. Topics not connected to any others, such as T1-8, means it can be taught at any time during the course because it does not build on previous topics, nor is it required to understand other topics.

While the graph abstraction of ABC 101 provides a lot of valuable insights, it is not the complete picture. ABC 101 is a prerequisite for ABC 103, but this is not yet indicated. Expanding the topic level abstraction to include all courses will provide the complete picture of the

curriculum and enable understanding of the inter-connectivity of the topics and courses. This is accomplished by constructing a graph abstraction of each course and then connecting them where appropriate. This results in the topic level graph abstraction of Curriculum A in **Figure 3**.

Topics can be introduced in one course and later reviewed in another. Reviewing a topic and then building off it (connecting to another topic) is different than just building off one topic to another because of the time spent on the review. To reflect this, if a topic is introduced and later reviewed, then the size of the topic is the sum of both times. To distinguish between, a border will be added to node. For example, T3-24 is introduced in ABC 103, but is later reviewed in ABC 105. As such the inner circle represents the resources first spent learning T3-34 and the outside circle represents the resources spent reviewing.

Reviewing topics in later courses could seem to lead to a cycle being created. The simplest case of a reviewing a topic is to remind students of what they were taught due to the time between the topic was first presented and when it needs to be built off from. For this case, no



**Figure 3.** Topic level graph abstraction Curriculum A in which topics are represented by nodes and topic dependencies are represented by directed edges forming a directed acyclic graph. Node size reflects the relative resources required for the topic in terms of the credit hours for the course.

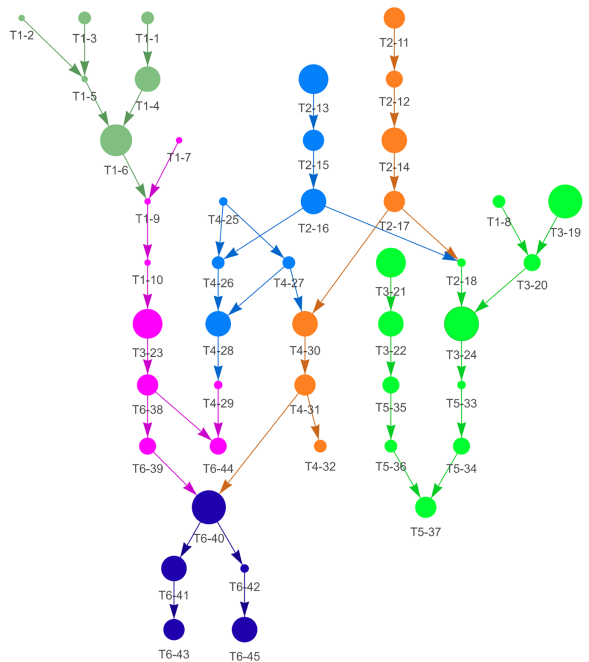
cycle is formed because it is just a reminder of what was taught with no other context or methods. A topic may also be reviewed to provide more detail in a new framework or using a new skill that has since been taught. For example, phase equilibrium may be introduced in the context of pure components in a first course on thermodynamics. Then phase equilibrium may be introduced in a multicomponent context in a second course on thermodynamics. This scenario would then be composed of three nodes: “Phase Equilibrium of Pure Components”, “Mixtures”, and “Phase Equilibrium of Mixtures”. In this scenario to teach “Phase Equilibrium of Mixtures” one would review “Phase Equilibrium of Pure Components” and combine that with “Mixtures”. This example demonstrates that revisiting a concept in a new context can be treated as a review of the original topic plus a dependency on the context.

The topic level graph of Curriculum A can be very useful on its own to understand the scope of the curriculum and discuss properties amongst instructors. However, the analysis can be improved by using tools from graph theory to provide formalized measures to discuss ways of changing the curriculum. Measures like the degree of node (the number of edges connected to a node), determining paths (getting from one node to another without any repeat nodes), the number of components (subsets of nodes that do not have edges to other nodes), connected vs. strongly connected graphs, and more can be considered. For this analysis, grouping of nodes through community detection and modularity and the degree of a node will be critical for analysis.

For grouping of nodes, a common approach is known as community detection which seeks to maximize “modularity” using optimization. The “modularity” is function of the interconnectivity of communities that make up the graph. These algorithms can be useful to identify which topics form areas of dense interconnectivity, but it does not consider the size of the nodes. This means it may result in communities of varying sizes. For example, applying the community detection algorithm known as greedy modularity maximization to Curriculum A results in the communities in **Figure 4**. Here the number of communities was restricted to be equal to the number of original courses (six). It is important to note that the topics forming a community are not always from the same course. The community sizes are 5.8, 3.4, 3.2, 2.7, 2.65, and 2.25 credit hours which when turned into courses could lead to an unbalanced course load.

To improve upon the community detection, a different approach to modularity will be considered. This approach described in [7] can restrict the sizes of the module in addition to the number of modules. This approach maximizes a different measure of modularity with constraints to set the number of modules and the size of modules are within a given size. This is cast a binary

quadratic optimization program. Additionally, we can use the framework to provide a definition of a modular curriculum. A curriculum will be considered modular if: (a) the topics in a module form a cluster of dense dependency, (b) connectivity between modules is sparse, and (c) the modules are within a set resource use (credit hours). These two approaches enable evaluating the modularity of the curriculum using modules to guide course organization and overall curriculum planning.



**Figure 4.** Topic level graph abstraction Curriculum A grouped into six communities (equal to the number of original courses) using greedy community detection.

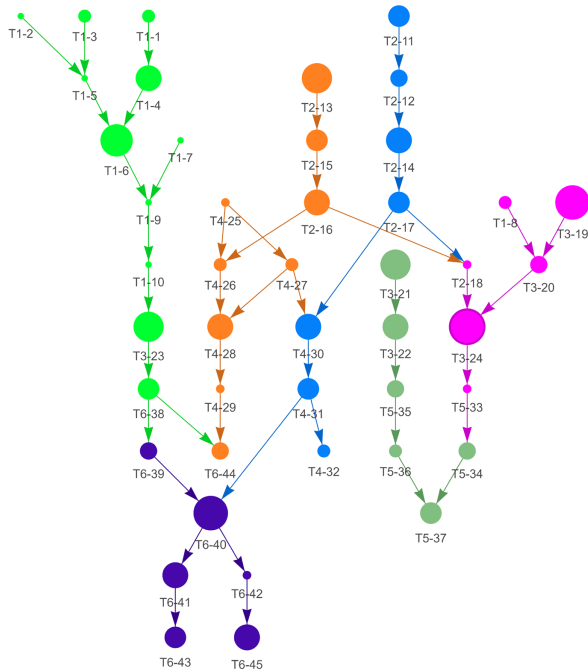
Restricting the module size between 2.9 and 4.1 credit hours, the new topic grouping is presented in **Figure 5**. This results in community sizes of 3.9, 3.8, 3.4, 3.1, 2.9, and 2.9 credit hours. This approach leads to more balanced course load implementable sizes.

## CASE STUDY

To demonstrate some the capabilities and insights that can be gained from this graph abstraction, this analysis will be applied the undergraduate chemical engineering curriculum at University of Wisconsin-Madison. This curriculum is composed of 19 credit hours of math, 10 of physics, 20 of chemistry, 6 of life sciences, 6 of communication skills, 16 of liberal studies, 6 of professional depth, 40 of core chemical engineering courses, and 9 chemical engineering electives for a total of 132 semester credit hours. The focus of this analysis will be on the core chemical engineering courses listed in **Table 2** along with their already defined prerequisites. Most of



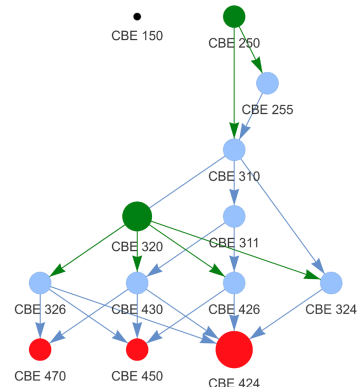
these courses are similar to courses required by other undergraduate chemical engineering programs.



**Figure 5.** Topic level graph abstraction Curriculum A grouped into six communities (equal to the number of original courses) when restricting modules size.

A graph abstraction of the core chemical engineering courses is presented in **Figure 6**. One interesting note of the curriculum is CBE 250 can be taken concurrent with CBE 255 and CBE 255 can be taken concurrent with CBE 310, but CBE 250 is a prerequisite for CBE 310. Since this graph abstraction is made to model the flow of topics, CBE 250 will be treated as a prerequisite CBE 255 and CBE 255 will be treated as a prerequisite for CBE 310.

A similar scenario exists for CBE 320 and CBE 324 and CBE 430 and CBE 470. From this course level abstraction, one can easily see just how many courses depend on CBE 320 based on the connectivity.

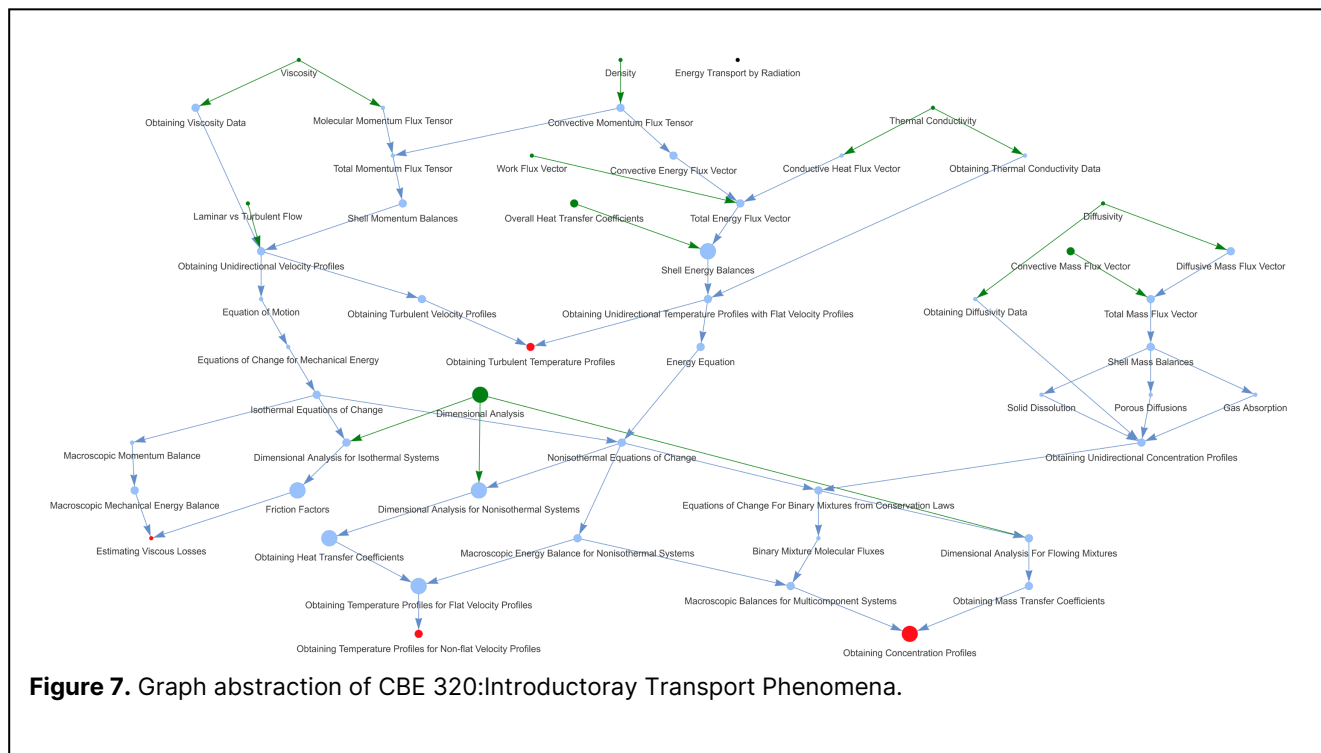


**Figure 6.** Course level graph abstraction of the core chemical engineering courses.

To construct the topic level graph abstraction of the curriculum, the syllabuses, course schedules, course readings, and lab documents were collected from recent occurrences of the courses. Then, a graph abstraction was developed for each course. For example, the graph CBE 320: Introductory Transport Phenomena is shown in **Figure 7**. This course covers momentum, heat, and mass transport at a theoretical level. The selection of topics and the “resolution” (how detailed the topics are e.g., Energy Flux Vector vs Convective, Conductive and Work Flux Vector) of the topics had to be decided for each course. This was an iterative process to ensure the graphs for each course were similar in terms of the level of resolution. While the graphs made and presented with this approach could look different depending on who

**Table 2:** Core chemical engineering courses at University of Wisconsin-Madison. Term taken is the term a student would normally take the course where 1F means fall of year one and 4S means spring of year four.

Course	Name	Term	Credits	Prerequisite(s)
CBE 150	Intro to Chemical Engineering	1F	1	None
CBE 250	Process Synthesis	2F	3	None
CBE 255	Intro to Chemical Process Modeling	2S	3	CBE 250 (co)
CBE 310	Chemical Process Thermo	2S	3	CBE 250 (co)
CBE 311	Thermodynamics of Mixtures	3F	3	CBE 310, CBE 255
CBE 320	Intro Transport Phenomena	3F	4	None
CBE 324	Transport Phenomena Lab	3S	3	CBE 310, CBE 320
CBE 326	Momentum and Heat Transfer	3S	3	CBE 320 (co)
CBE 426	Mass Transfer Operations	4F	3	CBE 311, CBE 320
CBE 430	Chem Kinetics and Reactor Design	4F	3	CBE 311, CBE 320
CBE 450	Process Design	4S	3	CBE 326, CBE 430, CBE 426
CBE 470	Process Dynamics and Control	4S	3	CBE 326, CBE 430 (co)
CBE 424	Operations Process Lab	4Su	5	CBE 324, CBE 326, CBE 426, CBE 430

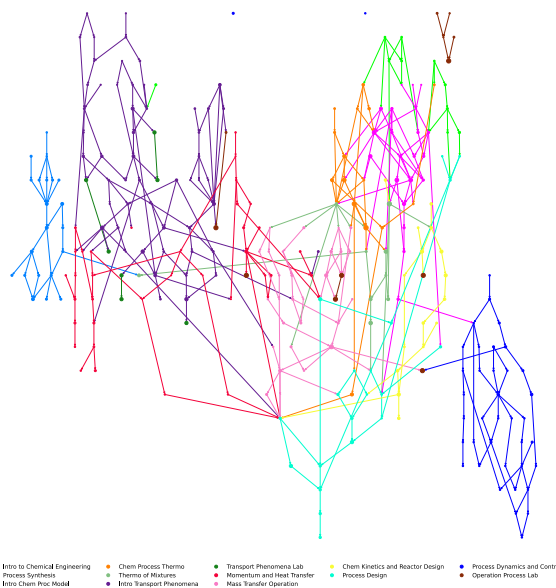


makes them, it opens a more formal discussion of what is being taught to students in a course. This could be used alongside learning goals to get the full picture of a course.

With a graph abstraction of each course, the topic level graph abstraction of the core chemical engineering courses can be constructed. To do this, the dependencies between courses and review of topics in subsequent courses were identified. For example, there is review of some material from CBE 310: Chemical Process Thermodynamics in CBE 311: Thermodynamics of Mixtures and there are topics introduced in CBE 320: Introductory Transport Phenomena that are required to understand topics in CBE 426: Mass Transfer Operations. This was also an iterative process to ensure topics of the same name had the same scope and to ensure the interconnectivity of courses was fully captured. Again, the interconnectivity of the topics between courses could be up for debate. However, this abstraction encourages conversation beyond “Make sure you introduce the Peng-Robinson Equation of State in Thermodynamics I because I need it for Thermodynamics II” that may occur between instructors teaching a series.

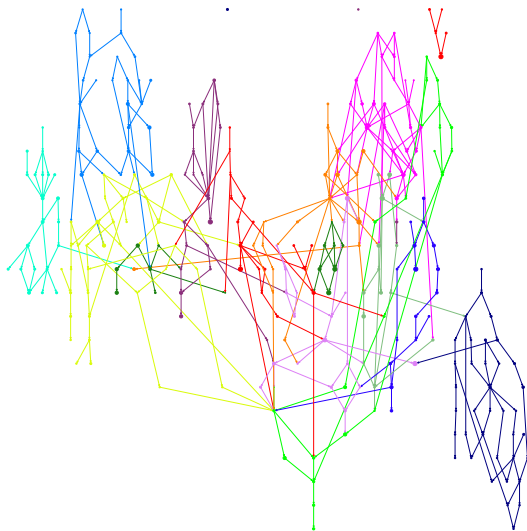
The topic level graph abstraction is presented in **Figure 8**. Due to the size of the graph, interactive versions of the graph are available in the supplementary material. Across the 13 courses and 40 credit hours, there are 286 topics and 398 edges. 43 topics are introduced in one course and reviewed later in another totalling 4.96 credit hours. Introductory topics often involve a physical property like density and entropy while topics without successors are often lab experimentation or extensions

into more complex application such as multiphase reactions.



The grouping of topics into modules was conducted restricting the module size to three to four credit hours across 13 modules resulting in modules in **Figure 9** (unfortunately, due to the size of the problem, the optimization was solved to 4.4% optimality gap in one hour and

future work will need to be done to optimize an entire curriculum with reasonable solver times). The modules here would reflect if the department wanted to keep the same number of courses but rearrange the topics to create more even courses rather than being between one and five credits. Some of the modules are close to what they were prior. For example, topics covered within CBE 470 and CBE 450 are largely still with the same topics. This demonstrates these courses are distinct from other parts of the curriculum. For CBE 450, this means that implementing systems thinking and design throughout the curriculum would require a different approach than just rearranging the topics from process design. As such, to teach systems design throughout a curriculum we would have to revisit the entire curriculum and identify gaps and areas where concepts from systems design could be introduced using this framework.



**Figure 9.** Graph abstraction the core chemical engineering courses with topics grouped into modules.

An additional note is on the transport related courses of CBE 320, CBE 324, CBE 326, CBE 426, and CBE 424. Previous work has discussed how different institutions sequence transport topics. The curriculum is currently designed to introduce the theoretical aspects of transport in one course and equipment and experimentation in subsequent courses. The modules determined by the optimization, on the other hand, tend to have the theoretical, equipment, and experimentation of each transport type (momentum, heat, mass) together due to the higher connectivity of topics within the same transport type compared to between the transport types.

## CONCLUSION

This work presented a graph abstraction in which courses are represented by nodes and connections between courses are represented by directed edges. Additionally, a graph abstraction in which topics are represented by nodes and topic dependencies are represented by directed edges was presented. This graph abstraction was applied to the core chemical engineering courses required for undergraduate chemical engineering students at University of Wisconsin-Madison. This abstraction identified lowly and highly connected topics and discussed ways of grouping topics. This abstraction encourages instructors to discuss how topics and curricula are presented. This work improves upon past efforts by formalizing the node attributes (topic size, review size) and the direction of edges between nodes. This formalization allows the exploration of topic dependencies and grouping with measurable features from graph theory.

## FUTURE WORK

While the core chemical engineering courses for the undergraduate chemical engineering program at University of Wisconsin-Madison was presented here, this analysis can easily and will be expanded to include the entire undergraduate chemical engineering curriculum. This will enable further study into the interconnectivity between chemical engineering courses and background courses from chemistry, math, and physics. With this ability, the determination of which courses should be taught by the chemical engineering department and which courses are appropriate to be taken from other departments.

For example, the chemical engineering undergrads at University of Wisconsin-Madison are required to take the three-credit STAT 324: Introductory Applied Statistics for Engineers course while there also exists CBE 562: Statistics for Chemical Engineers and can be used towards the chemical engineering elective requirement. This graph abstraction could be used with connectivity measures to make a case for CBE 562 to replace STAT 324 entirely to improve the connectivity of the topics with more relevant chemical engineering examples and improve reinforcement of topics.

In the current framework, some additional constraints could be added to account for additional attributes of topics. For example, a topic could be categorized as being a lab or lecture topic and a constraint could be implemented to ensure modules are made such that they are either lab or lecture modules. This would make the modules more accurately reflect more traditional course organization.

While the module sizes were set to the sizes of typical courses in this analysis, the module size can be smaller. Setting the module size smaller would reflect smaller units within a course. This would enable writing

specific learning objectives for each module and package the curriculum into more easily presented sizes. This would further encourage discourse surrounding the curriculum and what exactly is and should be taught within it. The creating of smaller modules could have the added benefit of increased flexibility to move modules around and replace modules with different concepts such as adding and electrochemistry module to the curriculum.

Another use for this graph abstraction is to embed important concepts or ideas throughout it such as safety, sustainability, and systems thinking and design. For example, implanting systems thinking and design throughout the curriculum can be accomplished by creating case studies. These case studies would involve systems thinking and design that depend on topics from different stages in the curriculum. In doing so, the case studies would have to be designed in a way such that all topics are used to understand and solve the case studies. A similar approach could be used for implementing safety and sustainability.

## DIGITAL SUPPLEMENTARY MATERIAL

Interactive HTML files of Figures 8 and 9 are available at <https://github.com/zavalab/ML/tree/master/CurriculaGraphs>.

## ACKNOWLEDGEMENTS

We acknowledge funding from NSF CAREER award CBET-1748516

## REFERENCES

1. Rampasso IS, Anholon R, Silva D, Cooper Ordoñez RE, Quelhas OLG, Leal Filho W, Santa-Eulália LA. An analysis of the difficulties associated to sustainability insertion in engineering education: Examples from HEIs in Brazil, *J Clean Prod*, 193:363–371 (2018)
2. Armstrong RC. A Vision of the Chemical Engineering Curriculum of the Future. *Chem Eng Educ* 40:104–109 (2006)
3. Rouly JM, Rangwala H, Johri A. What Are We Teaching? Automated Evaluation of CS Curricula Content Using Topic Modeling. *ICER '15* 189–197 (2015)
4. Toral SL, Martínez-Torres MR, Barrero F, Gallardo S, Durán MJ. An electronic engineering curriculum design based on concept-mapping techniques. *Int J Technol Des Educ* 17:341–356 (2007)
5. Bussemaker M, Trokanas N, Cecelja F. An ontological approach to chemical engineering curriculum development. *Comput Chem Eng* 106:927–941 (2017)

6. Voronov RS, Basuray S, Obuskovic G, Simon L, Barat RB, Bilgili E. Statistical analysis of undergraduate chemical engineering curricula of United States of America universities: Trends and observations. *Educ Chem Eng* 20:1-10 (2017)
7. Shao Y and Zavala VM. Modularity measures: Concepts, computation, and applications to manufacturing systems. *AIChE J* 66:e16965 (2020)

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Laying the foundations of Machine Learning in Undergraduate Education through Engineering Mathematics

Pavan Kumar Naraharisetti <sup>a, \*</sup>

<sup>a</sup> Newcastle University in Singapore, Faculty of Science, Agriculture & Engineering, Newcastle University, Newcastle Upon Tyne NE1 7RU, UK

\* Corresponding Author: [pavan-kumar.naraharisetti@ncl.ac.uk](mailto:pavan-kumar.naraharisetti@ncl.ac.uk)

---

## ABSTRACT

Some educators place an emphasis on the commonalities between engineering mathematics with process control, among others and this helps students see the bigger picture of what is being taught. Traditionally, some of the concepts such as diffusion and heat transfer are taught with a mathematical point of view. Now-a-days, Machine Learning (ML) has emerged as topic of greater interest to both educators and learners and new and disparate modules are sometimes introduced to teach the same. With the emergence of these new topics, some students (falsely) believe that ML is a new field that is somehow different and not linked to engineering mathematics. In this work, we show the link between the different topics from engineering mathematics, that are traditionally taught in UG education, with ML. We hope that educators and learners will appreciate the treatise and think differently, and we further hope that this will further increase the interest to improve ML models.

---

**Keywords:** Education, Machine Learning, Optimization, Numerical Methods, Modelling, Process Control

## INTRODUCTION

It is important that students understand the similarities and differences between the various modules in UG education so that they become well rounded engineers. To this end, mathematics, kinetics, thermodynamics, and process control allow us to explain the link between the different modules. While each module is designed to be disparate for the convenience of teaching, without placing due emphasis on the commonalities, it is easy to forget that all the topics are linked. Some educators place an emphasis on the commonalities between a.) engineering mathematics (ordinary differential equations, eigen values and eigen vectors) with process control (Laplace Transforms, poles, and state space modelling), b.) engineering mathematics and diffusion/heat transfer via partial differential equations, among others and this helps students see the bigger picture of what is being taught. Further, the concept of empirical equations and models is introduced much earlier to students through topics in fluid mechanics and heat and mass transfer.

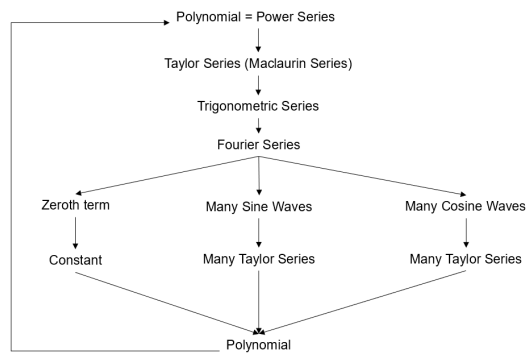
Considering UG Chemical Engineering education at the Newcastle University in Singapore, the different modules that teach applied mathematics are Engineering Mathematics – I & II, Computing & Simulation, Process Measurement Dynamics and Control, Process Control 2, and Chemical Process Optimization. In these modules, the following topics are taught. Linear algebra, ordinary differential equations (ODEs), numerical methods – solution to linear and nonlinear systems, series (Taylor, Fourier, Trigonometric), partial differential equations (PDEs), introduction to artificial neural networks (ANN), and Laplace Transforms. In addition, students are also taught to program some of the topics using Matlab and Python.

A focused student and an experienced educator can readily see that the solution to some common ODEs is either in terms of an exponential or in terms of a sine or cosine function. Further, some common PDEs can be written as a combination of exponential, sine, and cosine functions using the Fourier Law. Through the study of Taylor series, we know that any function (exponential, sine, cosine, or others) can be written as a polynomial and



hence we can hypothesize that the solution to any ODE or PDE is a polynomial. We know that the trigonometric functions such as sine and cosine can be written as a polynomial using Maclaurin series. Thus, a system represented as a trigonometric series can be shown to be a polynomial. Fourier series is a special case of trigonometric series in which the coefficients can be calculated through the given formulae. Thus, Fourier series is also a polynomial (Fig. 1).

Further, transfer functions in ANN have its foundations in exponential functions. Albeit the exponential is in the denominator, we will be able to convert this into a polynomial using Taylor series. Exploring such commonalities, we aim to bring forth a common thread between engineering mathematics and machine learning so that readers will better appreciate what they are learning in the new and emerging topics.



**Figure 1.** System (One Dependent and One Independent Variable) – Polynomial

Through some examples, this paper will lay the foundations of the inter-links between the different topics and show that “any (simple) system can be modelled as a polynomial”. It is believed that the motivational examples will lay the foundations for understanding machine learning better. To this end, we start with univariate problems, as simple polynomials (and straight lines which are first degree polynomials) are taught to students at the beginning are most courses as they are easy to comprehend.

## SINGLE-INPUT SINGLE OUTPUT MODELS

Single-input, single-output (SISO) models are the earliest and easiest to understand, especially so in UG education. We begin first by introducing the power series whose parameters must be estimated. A simple example is that of a solution to a 1<sup>st</sup> order ordinary differential equation (ODE) shown in Eqn. 2. This can be re-arranged and shown that the solution is an exponential function i.e.,  $e^{x+c}$ , where ‘c’ is a constant. If  $e^c = a_0$  the solution is  $a_0e^x$ . While it is known that  $e^x$  can be written as a

polynomial using Maclaurin series (derivation not shown), the solution to Eqn. 2 becomes a polynomial. Separately, we can also calculate the coefficients of the power series by substituting it (Eqn. 1) in Eqn. 2 and prove that the power series is the same as the polynomial obtained through direct integration.

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 \dots \quad (1)$$

$$\frac{dy}{dx} = y \quad (2)$$

Substituting Eqn. 1 in Eqn. 2 gives,

$$(a_1 - a_0) + (2a_2 - a_1)x + (3a_3 - a_2)x^2 + \dots = 0 \quad (3)$$

Thus,  $(a_1 - a_0) = 0$ ;  $(2a_2 - a_1) = 0$ ;  $(3a_3 - a_2) = 0$ . This shows that  $a_1 = a_0$  and  $a_n = \frac{a_0}{n!}$  and hence,  $y = a_0e^x$

## Modelling differential equations as polynomials

The solution to a simple 1<sup>st</sup> order ODE has been presented from Eqn. 1 onwards. Here, let us look at some other examples where the solution to an ODE can be presented as a polynomial. Consider a 2<sup>nd</sup> order ODEs which can be written in homogeneous form as,

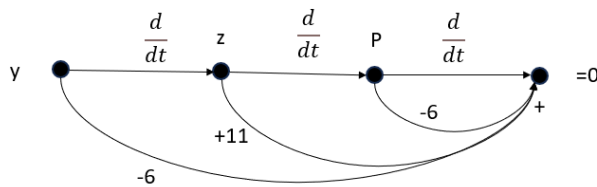
$$a \frac{d^2y}{dx^2} + b \frac{dy}{dx} + cy = 0 \quad (4)$$

The solution to this can be obtained by solving for the roots of the characteristic equation. If the roots are real, the solution only has exponentials, however, if the roots are complex, the solution has trigonometric functions which causes the system to oscillate (relevant to stable and unstable systems in process control). And since trigonometric functions can always be written polynomials, the solution to the above ODE is a polynomial. Here, we should note (details not shown) that we can show that the roots of the characteristic equation are the same as the poles of second order equation solved using Laplace Transforms (in process control) and that the roots are also same as the eigenvalues if we were to solve the second order equation by reducing it into a system of two first order equations written in homogeneous form. As a side note, we need to remember that eigenvalues and eigenvectors are also useful in state-space modelling in process control. While it is possible to obtain the solution for homogeneous form of ODEs, the concept of system as a polynomial can also be extended to non-homogeneous forms by obtaining the complementary function (solution to homogeneous form which is a polynomial) and the particular-integral (where solution is assumed as either being a polynomial or a trigonometric function which in turn is a polynomial).

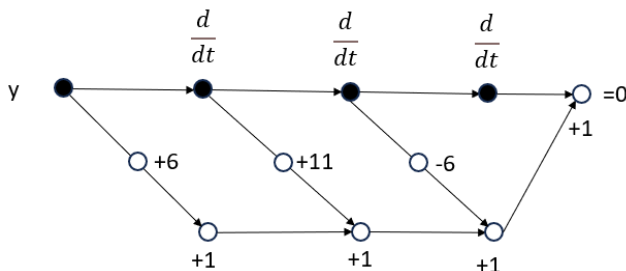
If we consider higher order ODEs, which is still a SISO system, it can be reduced to a system of first order

linear ODEs that can be represented in a homogeneous form ( $\frac{dx}{dt} = A$ ) and can be solved using Eigenvalues and Eigenvectors of the matrix A. The solution would be of the form,  $x = \sum_{i=1}^n c_i v_i e^{\lambda_i t}$  where,  $\lambda_i$  are the eigenvalues,  $v_i$  are the eigenvectors, and  $c_i$  are the constant of integration which can be obtained using the initial conditions. Since an exponential can be written as a polynomial, we can conclude that the solution to a system of first order linear ODEs is a polynomial as well. Whether this structure can be represented as a graph, or a neural network would be of interest.

Consider the third order ODE,  $\frac{d^3y}{dt^3} - 6\frac{d^2y}{dt^2} + 11\frac{dy}{dt} - 6y = 0$ . Fig. 2 shows the ODE represented as a graph. Readers who are familiar with Simulink would readily see that this graph has similarities with it. Simulink would have an inverse representation and solve the system using integration while the representation in Fig. 2a, shown as a graph, uses differentiation. The same is further elaborated in Fig. 2b as activity on node so that it is easier to see the similarities with a neural network. Hence, it is not entirely surprising that ODEs have been incorporated into neural networks. In our syllabus, we cover the famous Lotka-Volterra model to explain solution to system of ODEs.



**Figure 2a.** Representation of an ODE as a graph.



**Figure 2b.** Representation of an ODE as a network with activity on node, which is akin to a neural network.

More in-depth studies incorporating solution of ODEs had been addressed by the development of neural differential networks where researchers have commented that “neural differential equations (NDEs) demonstrate that neural networks and differential equation are two sides of the same coin” [1]. While the details of NDEs are beyond the scope of this work, we believe that Fig. 2 aptly captures this statement. A simplistic view

is that the neural network in NDEs is a series of first order ODEs resulting in a higher order system.

### Use of SISO models in neural networks

When we move to artificial neural networks (ANN), the logistic regression, a special case of the sigmoid function, is used as the activation function (Eqn. 5). The sigmoid function, given below, can also be represented as a polynomial. There are three ways of deriving the polynomial. One, we could use the Maclaurin series to derive the same. Two, one can assume a higher order polynomial and estimate the parameters using matrix operations. Three, estimate the parameters using optimization (error minimization).

$$y = \frac{1}{1 + e^{-(x-\mu)/s}} \quad (5)$$

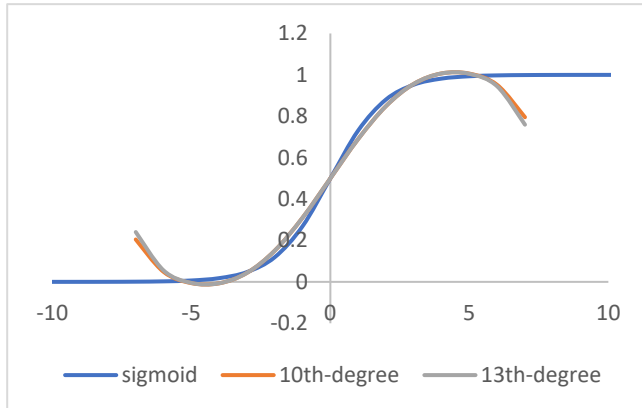
where,  $\mu$  is a location parameter used to recenter ‘x’ around ‘0’ and ‘s’ is scaling parameter used to normalise the value of ‘x’. For simplicity, we can assume that  $\mu = 0$  and  $s = 1$ .

A comparison of the 10<sup>th</sup>-degree and 13<sup>th</sup>-degree polynomials generated by using matrix operations for parameter estimation with a sigmoid function is shown in Fig. 3 (method of estimation and parameters not shown). We observe that it is rather easier to represent the data/system as a sigmoid than a polynomial and that the polynomials generated do not completely represent the sigmoid. Whether a higher order polynomial is required remains to be seen. Thus, for use in ANN, it is prudent to use sigmoid instead of an equivalent polynomial. For more on the equivalent polynomials used in ANN and other advanced versions of neural networks, the reader is referred to [2]. Another point to note here is that We can clearly see here that we have plotted only [-7, +7] range for the polynomials which already resulted in poor accuracy whereas we can plot for any range  $[-\infty, +\infty]$  when sigmoid is considered. While it is known that Polynomial Neural Networks are difficult to implement because polynomials have ‘exploding gradients’ i.e., steep inclines in the plot, this simple single input, single output system clearly demonstrates the pitfalls of using a polynomial as an activation function. Will a trigonometric series give a better fit? Probably (see next section on sawtooth wave alludes to it), however, we refrain from doing it for the SISO system.

Although our attempt to represent sigmoid function has failed (miserably), one can readily see why. Sigmoid function is in fact a ratio of two polynomials, and we can see this if we try to bring the exponential to the numerator. It may be worthwhile try to model sigmoid as the ratio of two different polynomials, but we refrain from doing it here.

To conclude this section, while any SISO system can be modelled as a polynomial, the question remains as to

whether it is necessary to do so in neural networks especially since many hidden layers with the activation function in each will result in a nonlinear system and may sufficiently capture the system dynamics.



**Figure 3.** Comparison of sigmoid function with a 10<sup>th</sup> and 13<sup>th</sup> degree polynomial.

### A side note

Using the Maclaurin series, the formula for the binomial series, which is  $(1 + x)^n$ , can be derived, and by using the binomial series, the expanded equation for  $(a + b)^n$  can be found by assuming that  $\frac{b}{a} = x$ . As an example, we can expand  $(a + b)^3$  as,

$$a^3(1 + x)^3 =$$

$$a^3(1 + 3(1 + x_{x=0})^2x + \frac{3 \times 2(1 + x_{x=0})x^2}{2!} + \frac{3 \times 2 \times 1 \times x^3}{3!})$$

which becomes,

$$a^3(1 + 3x + 3x^2 + x^3) = (a^3 + 3a^2b + 3ab^2 + b^3).$$

We must note here that this derivation is normally not taught using series and it is hence interesting that algebraic formulae can also be derived using series and represented as a polynomial.

## DOUBLE-INPUT SINGLE OUTPUT MODELS

Double input, single output systems can also be called Two/Twin input, single output models (TISO). Let us consider two problem types for this analysis. The first is that of geometric shapes and the second is solution of 2<sup>nd</sup> order partial differential equations used in heat and mass transfer.

### Modelling topology using neural networks

We first consider the case of a cone and a paraboloid. They can be represented by Eqn. 5a & 5b

respectively.

$$z^2 = x^2 + y^2 \tag{5a}$$

$$z = x^2 + y^2 \tag{5b}$$

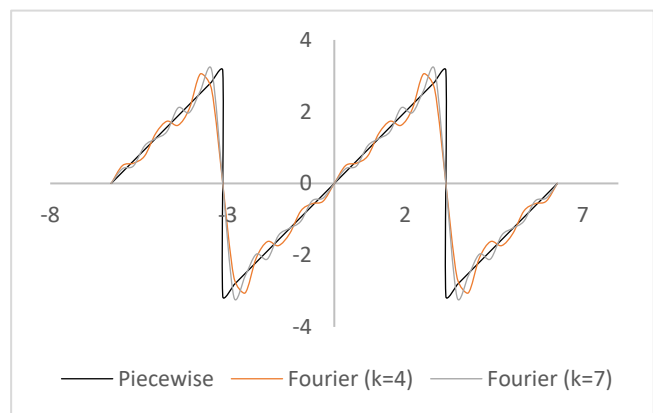
In theory, these two surfaces can be modelled as an ANN with two inputs ( $x$  and  $y$ ) and one output ( $z$ ). Once the ANN model is generated (using data generated by Eqn. 5), we should be able to calculate  $z$  by giving different values of  $x$  and  $y$  and regenerate the surfaces using the ANN model. The question then is that if surfaces can be modelled simply as in Eqn. 5, should we resort to using neural networks to do the same? Or, a more philosophical question is, how can topology be modelled as a neural network? The answer does not seem to be simple or straight forward, and it appears that when the physics is well understood, there is no need to use a neural network, or is that so? Let us understand this visually by taking a step back by looking at a SISO problem.

Consider the sawtooth wave which can be represented by simple piecewise linear mathematical representation as given in Eqn. 7. While this is simple, engineers however are interested in a continuous function that can be differentiated and is easy to implement (especially in electronics and signal processing). Fourier transforms allow us to generate an equivalent continuous function which allows us to do the same. The equation for the Fourier series is given in Eqn. 8.

$$f(x) = x \quad \forall 0 \leq x < \pi \tag{7a}$$

$$f(x) = x - 2\pi \quad \forall \pi \leq x < 2\pi \tag{7a}$$

$$f(x) = 2 \sum_{k=1}^{\infty} ((-1)^{k+1} \frac{\sin(kx)}{k}) \tag{8}$$

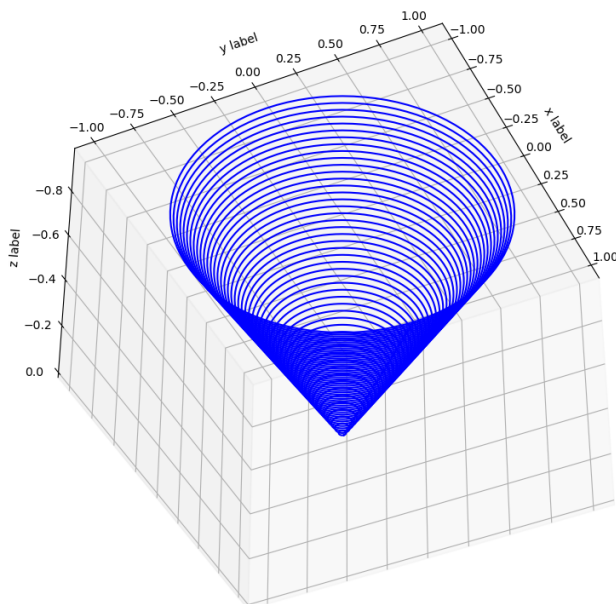


**Figure 4.** Sawtooth wave. A comparison of piecewise linear representation vs. Fourier series ( $k = 7$ ).

It can be seen from Fig. 4 that although the Fourier series will have increasing accuracy with increasing 'k', the 'edge-effects' will continue to be present. While the

application of this simple example is easy to understand, and we are willing to compromise accurate (accuracy beyond a point is not perceived to the end-user) for easy of use, we do not fully understand the implications of compromising accuracy over easy of use in multidimensional problems where neural networks are typically used. Note: It is possible to represent the sigmoid function, presented in the earlier section, as a trigonometric series or Fourier series, but we will refrain from doing it here.

Another way to look at the issue of accuracy is by considering a cone or a paraboloid. While Eqns. 5a & 5b give a perfectly smooth representation, objects in the real world are not nearly as perfect at the microscopic level. We can only achieve certain level of smoothness using machining processes and objects are always non-smooth when we see them under magnification. Thus, we can probably say that neither of Eqn. 5a & 5b are perfect representations of the reality. We do not however know what the implications are if we can represent this 'imperfect reality' through imperfect neural networks! Fig. 3 shows both the smoothness and the gaps between the data points in a cone ('imperfections' arising from discrete datasets).



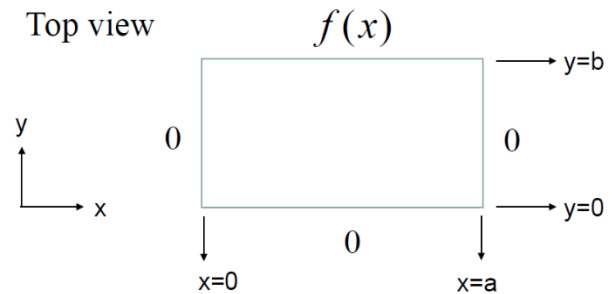
**Figure 5.** A cone, with both smoothness and gaps (imperfections) is shown.

Having said that, developing neural networks that understand or replicate topology may open new doors for the field of artificial intelligence in its quest for developing cognitive capabilities. Here, we should emphasize that we must not confuse 'topology of neural networks' with 'modelling topology using neural networks' – we are alluding to the later and not the former. Readers who are

more interested in 'topology of neural networks' with applications to 'modeling topology using neural networks' are referred to what is known as 'Geometric Deep learning', more specifically 'Geodesic Convolutional Neural Networks' [3], which are beyond the scope of a typical undergraduate programme in chemical engineering.

## Modelling partial differential equations using neural networks

Solution to partial differential equations with two independent variables and one dependent variable can also be derived by using Fourier series and the same is taught in our undergraduate courses. Consider the case of a hotplate shown in Fig. 6. The solution, temperature vs. location, of this problem which can be derived from first principles is given in Eqn. 9. We can see from Eqn. 9 that the solution is an infinite series that involves trigonometric function sine as well as exponentials. The data for change in temperature with location can be obtained by direct integration of the constant,  $c_n$ , and substituting different values of  $x$  and  $y$ . We must note here that, just as in SISO Fourier series, the accuracy increases as 'n' becomes large.



**Figure 6.** Hotplate that is heated at  $y = b$ .

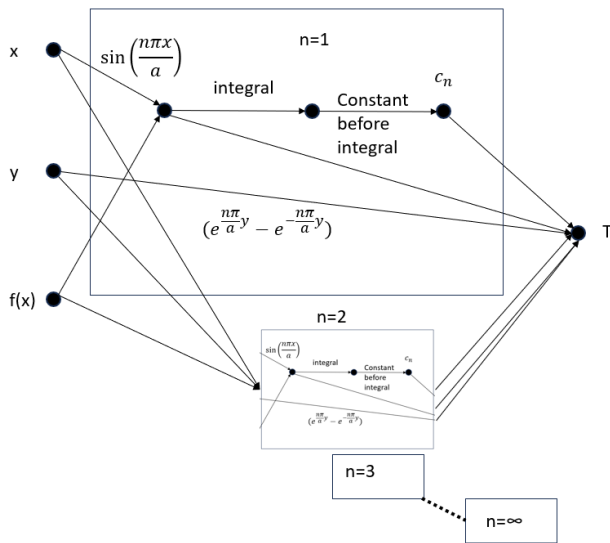
$$T(x, y) = \sum_{i=1}^{\infty} c_n \sin\left(\frac{n\pi x}{a}\right) \left(e^{\frac{n\pi}{a}y} - e^{-\frac{n\pi}{a}y}\right) \quad (9)$$

$$\text{where, } c_n = \frac{2}{a\left(e^{\frac{n\pi b}{a}} - e^{-\frac{n\pi b}{a}}\right)} \int_0^a f(x) \sin\left(\frac{n\pi x}{a}\right) dx$$

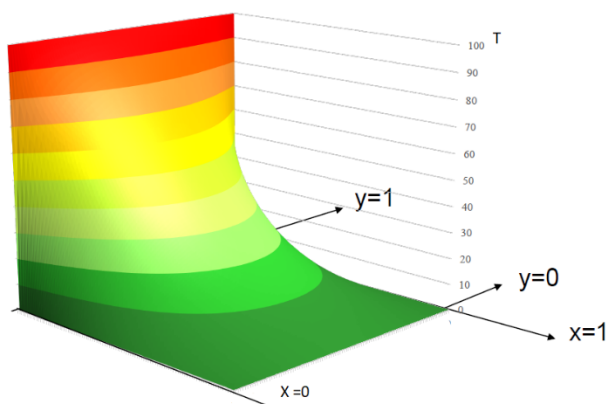
This equation can be visualised as being equivalent to a graph or a network through Fig. 7. Here, the node represents a product (multiplication) and the arc represents the transformation of the input node to the output. This is not exactly like the traditional neural network, but it is important to understand that the equation can be represented as a graph. When  $n=1$ , only the equations in big rectangle are used. As 'n' increases, the number of such rectangles can be increased as the structure within the rectangle remains the same.

Separately, the graphical representation of solution of Eqn. 9 is shown in Fig. 8a as a 3D image. On the other hand, obtaining the neural network model of the solution using data either from experiments or data generated

using Eqn. 9 is also possible. While it may not be clear from Fig. 8a that we can model this as a neural network, Fig. 8b makes it clearer. It being a 2D image with different colours, we can clearly see that generating multiple such images with different values of  $f(x)$  and calculating 'T' with respect to 'x' and 'y' will clearly help us generate an equivalent neural network model. We can say this with reasonable confidence since making sense of 2D images is one of the earliest applications of neural networks. While four sub-models are required if we are to obtain the solution to a hotplate that is heated on all four sides, Fig. 8c (equations not shown), it may be possible to generate an equivalent neural network model that can model a hot plate heated on all sides within one model.



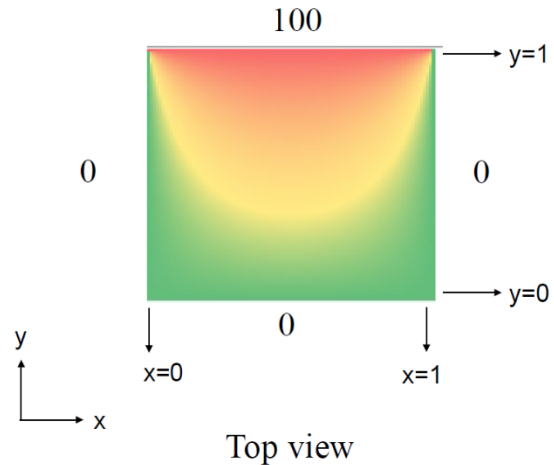
**Figure 7.** Graph representation of Eqn. 9 with hotplate heated on one side.



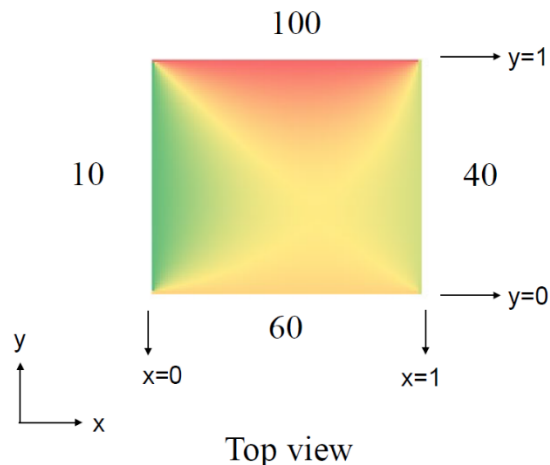
**Figure 8a.** Solution to PDE (in 3D) of a hotplate that is heated on one side. *Note: Fig. 6a-c are from lecture notes, CHE1021 Engineering Mathematics II by Liu Wen Paul, former faculty at Newcastle University.*

To elaborate, in a simple image analysis problem,

the value in each pixel is vectorized and several such pictures are used to train (parameter estimation) the model to determine if the image is that of a cat or a dog i.e., the training data consists of  $1024 \times 1024$  inputs (assuming that image dimension is  $1024 \times 1024$ ) vs. one output (cat or dog). On the contrary, if a hot plate is to be modelled as a neural network, there will be six inputs (location  $(x, y)$  and temperature,  $T$ , on the four sides) and one output (temperature at the given pixel/location). Even though the image dimensions are large ( $1024 \times 1024$ ), the problem can be modelled with fewer number of input nodes in the neural network.



**Figure 8b.** Solution to PDE (in 2D) of a hotplate that is heated on one side.



**Figure 8c.** Solution to PDE (in 2D) of a hotplate that is heated on all four sides.

We should note here that, just as accuracy increases with increasing 'n' in Eqn. 9, the accuracy of the neural networks also increases with increasing hidden layers and nodes. We should note here that parameters and hence model size increases with 'n' in Eqn. 9 and with the number of hidden layers and nodes in neural networks.



Hence, we may allude that similarities exist between first principles models and neural networks that more the parameters, the more accurate representation of the reality.

We should also note that parameter estimation becomes more tedious in this TISO model if we were to formulate is as a pure mathematical model with no understanding of Fourier series. We can envision that higher dimensional problems can also be modelled as a trigonometric series and exponentials. We must however remember that the issue with using trigonometric functions in mathematical modelling and then performing parameter estimation is that trigonometric functions create 'ripples' which have multiple (bad) local minima and hence parameter estimation may become difficult, if not impossible. On the contrary, neural networks have been extensively used to model higher dimensional problems. So, unless there is an easier way to estimate the parameters when using trigonometric series, such as in Fourier series, it is probably better to use neural networks in the near term. Having said that, it may be worthwhile to explore use of trigonometric series and exponentials in neural networks. One rather interesting use of trigonometric functions in neural networks is the use of wavelets and the treatise of the same is beyond the scope of this work.

## CONCLUSIONS

In this work, we have demonstrated through simple examples how engineering mathematics course in undergraduate education can be used to lay the foundations of neural networks. The author is of the view that other topics such as numerical methods, optimization, parameter estimation can also be taught with a machine learning flavour.

To conclude, although any SISO system may be a polynomial or may be reduced to be a polynomial, it is rather difficult to model (and perform parameter estimation on) all systems by assuming that they are multi-input, multi-output polynomials or as combination of trigonometric functions. Hence, modellers should be conscious to incorporate the physics of the system into the modelling framework where possible. Alternatively, an appropriate existing model must be used, typically with an understanding of the problem under consideration which incorporated known knowledge into the model. This must be emphasized early on in mathematics education so that future research and education can be directed at bridging the gap between traditional mathematics and machine learning, possibly leading to explainable AI.

## REFERENCES

1. Kidger P. On Neural Differential Equations. <https://doi.org/10.48550/arXiv.2202.02435>
2. Anonymous Authors. Polynomial activation functions. Under review as a conference paper at International Conference on Learning Representations (2020). <https://openreview.net/pdf?id=rkxsgkHKvH>
3. Bronstein MM, Bruna J, Choen T, Veličković P. Geometric Deep Learning Grids, Groups, Graphs, Geodesics, and Gauges. <https://doi.org/10.48550/arXiv.2104.13478>

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Part 2

## Non Peer-Reviewed Articles

Extended Abstracts



# Accelerating Discovery in Consumer Product Design

**Brian T. Gettelfinger**

Procter & Gamble, Mason, OH, USA  
[gettelfinger.bt@pg.com](mailto:gettelfinger.bt@pg.com)

**Keywords:** Consumer Products, Modeling, Design

## DESIGNING IRRESISTIBLE SUPERIORITY

At Procter and Gamble, innovation is based on a framework of Irresistible Superiority [1] that utilizes five complementary components – Products, Packages, Communication, Retail Execution, and Value. Increasingly, the computational techniques heavily leveraged within the Process Systems Engineering community are playing a leading role in delivering these five vectors, and they become increasingly valuable as we aim to deliver products in more exploratory consumer spaces – where combining high-volume data, advanced modeling, and quantified uncertainty will allow us to discover and deliver better products faster than ever before.

### Superior Product & Package Design

In the product design space, we have a long history of utilizing traditional modeling & simulation to deliver superior products [2]. In today's technology landscape, we see potential in both product and ingredient discovery through the complementary use of traditional modeling with machine learning techniques for high-throughput screening applications. This approach allows us leverage digital twins to explore a wide variety of environments in which our products must perform, but it also allows for a simultaneous exploration of numerous – and often competing – constraints on ingredient safety, sourcing, and performance.

Packaging presents different opportunities for computer aided design. We continue to use modeling & simulation to design our physical packages to be resilient in the supply chain and to give novel and superior consumer experiences. These same models – and specifically the material constitutive equations therein – can help us identify more sustainable materials and can improve our internal processing of our packages. In more recent developments, we are also using package models combined with formulation understanding to improve product/package interactions, enabling the delivery of sustainability without sacrificing product performance in our consumers' daily use.

### Superior Communications & Retail Execution

While the above two areas seem somewhat conventional domains for computational design, perhaps some of our biggest opportunities lie in consumer communication. We are constantly challenged to translate our products' technical superiority to consumer-noticeable superiority, and measuring that superiority is very well-suited to Bayesian inference from small testing and larger consumer behaviors data. For example, we have long utilized Bayesian methods for estimating latent effects in sensory tests, and we leverage learnings from these small experiments in well-established methods for claims and communication testing (i.e. A/B tests). These techniques help us ensure that the best aspects of our technical performance are meaningfully visible to our consumers.

Retail execution presents still more opportunities for computer-aided design – especially as an increasing amount of retail activity is conducted via digital platforms. We learn through online channels and ML models which aspects of brand propositions work well with consumers, and we are then able to leverage those best-performing elements into successful retail executions. Conversely, we often supplement our retail displays with digital information that is relevant to consumers based on environmental & behavioral trends – enabling models to present the most meaningful content to a consumer who is searching for more about our brands [3].

### Superior Value to Customers & Consumers

Finally, it is essential that we continue to deliver the above superiority at an exceptional value – to both the consumer and the customer. This value delivery is readily enabled by productivity-enhancing AI in data-rich areas like ingredient screening, but in other areas of design, we rely heavily on Process Systems Engineering experts across the whole of our supply chain. Increasingly, we are using process systems & machine learning together to prioritize & screen new ingredients not based solely on their performance properties, but also on their possible

supply chains & their actual and perceived safety & regulatory risks. This allows us to deliver new products quicker and more affordably than could have been imagined even 10 years ago.

## COMPUTATIONAL DESIGN SKILLS OF THE FUTURE

What then does this multi-faceted approach to design demand of the computational design community? First and foremost, it demands partnership. We need a diverse set of skills in computational disciplines that span data engineering, machine learning, multi-objective optimization, and Bayesian inference – all of which must complement well-established modeling disciplines! We cannot possibly expect researchers and students to master each of these disciplines, so they must instead demonstrate their ability to master a few areas and partner very effectively with the others as quickly as possible.

Second and just as importantly, we require practitioners of computational design to become obsessed with the decisions they are driving and the problems they are solving rather than the computational capabilities themselves. Practically, this means being keenly aware of the decisions being influenced by any prediction or optimization algorithm. And in an innovation environment, we need computational leaders to recognize that accelerated *learning* is an objective in itself – learning in new spaces, as efficiently as possible, with the possibility of generating new data in areas where we have never explored. This combination of factors paints an exciting picture of the future of computational design of consumer products, and as consumers ourselves, we should all be eager to enjoy the fruits of these superior design capabilities in our daily lives.

## REFERENCES

1. <https://us.pg.com/annualreport2023/superiority-that-wins-with-consumers/>
2. DeVane, R.H., Wagner, M.S., Murch, B.P., Current State and Future Needs in Materials Modeling. In: *Materials Research for Manufacturing*. Springer, Cham (2016)
3. <https://us.pg.com/blogs/investor-day-2022-series-digital-secret-to-solving-consumer-problems/>

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Process Design for the Energy Transition: An Industrial Perspective

Jaffer H. Ghouse\*

Systems Integration and Operations, New Energy Solutions, Woodside Energy, Houston, TX, USA

\* Corresponding Author: [jaffer.ghouse@woodside.com](mailto:jaffer.ghouse@woodside.com).

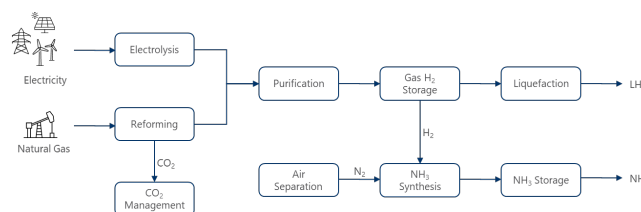
**Keywords:** Energy Systems, Optimization, Process Design, Hydrogen, Ammonia

## EXTENDED ABSTRACT

The United States' Inflation Reduction Act (IRA) of 2022 has established incentives to facilitate the energy transition. While these policies provide economic incentives that encourage investment and may reduce financial risk for the private sector on the supply side, transitioning to a lower carbon or net-zero economy by 2050 presents several challenges. These include designing flexible production systems that can interact with intermittent renewable energy resources, ensure process safety, redesigning existing energy infrastructure to support new energy carriers like hydrogen or ammonia, and making long-term investment decisions in an uncertain and evolving market. Addressing these challenges presents significant opportunities for the computer-aided design community to leverage techniques at the intersection of process design, AI/ML, data science, and optimization, which will play a key role in designing robust systems that can thrive during the energy transition.

This presentation aims to provide an overview of different technologies that are being considered and associated challenges faced by the industry at the ground level, particularly around hydrogen and ammonia production routes, as shown in Figure 1, in a post-IRA era. Specifically, the presentation will highlight the policy details in sections 45V and 45Q of the IRA and how these influence problem formulations and economic calculations. It will also demonstrate the importance of considering process design for hydrogen and ammonia production from a systems-level perspective that incorporates the interaction of tightly coupled, time-varying systems, and how conventional techno-economic designs need to consider emissions at the design stage given the incentive structures in the IRA are related to carbon intensity. In addition, this presentation will discuss how optimization-based end-to-end pipelines are deployed in production within a business environment while addressing

challenges around comparison to pre-existing methods, data uncertainty, problem infeasibilities, and sustaining engagement with end users.



**Figure 1:** Potential production pathways from electricity and natural gas for H<sub>2</sub> and NH<sub>3</sub>

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Designing for the Future: The Role of Process Design in Decarbonization and Energy Transition

M. M. Faruque Hasan<sup>a,b,\*</sup>

<sup>a</sup> Artie McFerrin Department of Chemical Engineering, Texas A&M University, 3122 TAMU, College Station, TX 77845, USA

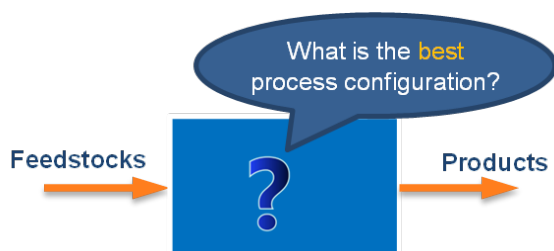
<sup>b</sup> Texas A&M Energy Institute, Texas A&M University, 1617 Research Pkwy, College Station, TX 77845, USA

\* Corresponding Author: [hasan@tamu.edu](mailto:hasan@tamu.edu).

**Keywords:** Process Design, Energy, Energy Efficiency, Carbon Capture, Process Synthesis, Decarbonization, Energy Transition

## INTRODUCTION

The overarching goal of process design (Figure 1) is to find technologically feasible, operable, economically attractive, safe and sustainable processing pathways and process configurations with specifications for the connectivity and design of unit operations that perform a set of tasks using selected functional materials (e.g., catalysts, solvents, sorbents, etc.) to convert a set of feedstocks or raw materials into a set of products with desired quality at a scale that satisfies the demand. Process synthesis and integration can further screen, optimize and improve these pathways for given techno-econo-environmental targets or objectives. These objectives may include, but are not limited to, minimizing the overall investment and processing costs, minimizing the energy consumption, minimizing the emissions or wastes, maximizing the profit, and enhancing the safety, operability, controllability, flexibility, circularity, and sustainability, among others.



**Figure 1.** Overarching goal of optimal process design.

For a long time, fossil fuels (petroleum, natural gas, and coal) have been predominantly used as primary feedstocks as well as primary energy providers for the chemical process industry (CPI). The chemicals and refining is by far the largest contributor of industrial direct

and indirect carbon emissions [1]. In recent times, decarbonization and transition to renewable energy have emerged as the major pathways for reducing anthropogenic greenhouse gas emissions that contribute to global warming and climate change. Like many other sectors, CPI is considering them to reduce its overall carbon footprint.

Industrial decarbonization primarily refers to reduction or elimination of CO<sub>2</sub> emissions from a manufacturing process. For CPI, it would include the following, among others [1]: (i) improving energy efficiency through novel process design, optimization and integration, (ii) incorporating bridging reduced/net-zero/negative emission technologies such as carbon capture, utilization and storage (CCUS) [2], direct air capture (DAC), hydrogen and/or biomass based energy & fuels, (iii) electrification of process heating and cooling, and (iv) substitution of fuels, feedstocks, and energy sources.

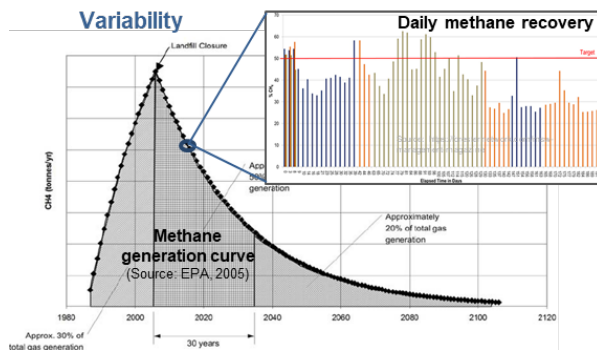
Energy transition is another broad concept that encapsulates a significant shift in the way we produce, distribute, and use our primary energy sources. Typically, it refers to move from conventional fossil fuels to more sustainable and environmentally friendly alternatives, such as renewables (solar, wind, hydro, geothermal), energy storage technologies, and increased energy efficiency measures.

This work explores the pivotal role of process design and optimization as the CPI navigates through different challenges and opportunities towards mitigating carbon emissions and facilitating the shift towards renewable energy.

## CHALLENGES AND OPPORTUNITIES

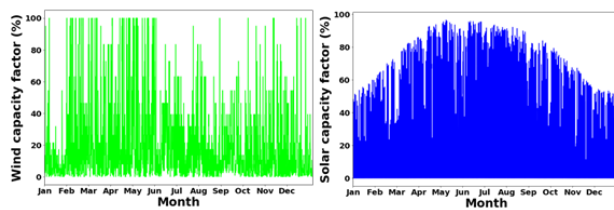
While these measures show great promise for reducing carbon emissions, an effective integration and deployment of these measures pose are significant and,

in some cases, unique challenges for the CPI. Chemical and refining processes are historically developed considering a steady supply of conventional feedstocks and energy. High efficiency and low capital intensity, both of which contributed to high economic gain, were achieved through economies-of-scale of large, centralized processing facilities. However, as we consider shifting towards utilizing unconventional feedstocks and renewable energy sources, we lose the economies-of-scale. Many unconventional feedstocks and most renewable energy sources are distributed, only intermittently available, and uncertain. One such example is shown in Figure 2, where we observe significant variability in the long-term, seasonal, and daily availability of methane from landfill gas, which is an unconventional feedstock.



**Figure 2.** Long-term, seasonal and daily variability in landfill-based methane feedstock availability.

Similar to the unconventional feedstocks, renewable energy sources such as solar and wind are distributed and intermittent. The future price of renewable electricity could decrease substantially to compete with traditional fossil fuels [3]. However, intermittency, spatio-temporal variability and non-dispatchability of renewables pose considerable challenges for systems integration. Figure 3 shows an example of variability of solar and wind energy availability.

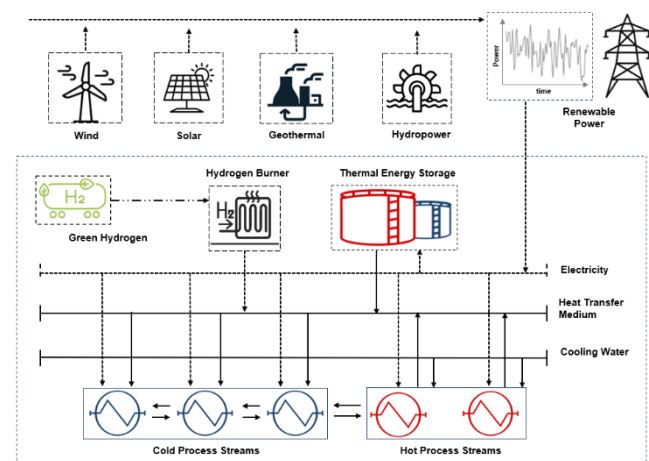


**Figure 3.** Temporal variability in solar and wind energy.

The intermittency of these variable renewable sources and the high energy requirement of carbon capture restrict their widespread deployment in the CPI. This often leads to simultaneous design and scheduling prob-

lems [4]. These challenges are traditionally addressed independently at the grid-level, leading to conservative costs and limited operational flexibility for both systems. However, opportunities exist to examine the synergistic integration of renewables and flexible carbon capture with individual chemical plants. Renewables can provide clean energy for carbon capture, while energy storage can be incorporated to counter renewable intermittency. To assess whether the benefits obtained from integration outweigh the capital cost under spatiotemporal variability of electricity markets and renewable energy, we can develop and use mathematical programming-based optimization frameworks for process design and optimization. We can decouple the design and operational decisions in a two-stage optimization strategy to efficiently solve the large-scale problem.

Electrification of the chemical industry reduces the reliance on fossil fuels across the economy and enables plants to integrate cleaner energy sources such as renewables. However, the variation and intermittency in renewable energy supply is a key challenge in electrifying industrial processes. Plant-wide energy storage will most likely play an important role in electrification by leveling off the change in energy supply [5]. Thermal energy storage (TES) is particularly suitable for chemical process plants due to the use of low-cost storage mediums. With the support from TES, renewable powered electric heaters can be introduced. However, there is a need for systematic methods that simultaneously consider the time-varying utility supply from renewables, the design and integration of single/multistage TES process configurations with heat exchanger networks, the scheduling of optimal charging/discharging operations, and the selection of energy storage mediums. One such concept of renewable-powered electrification is illustrated in Figure 4.



**Figure 4.** Renewable powered chemical process heat integration with TES and backup green hydrogen.

By integrating advanced modeling techniques, simulation tools, and artificial intelligence, engineers can further optimize the performance of existing processes and develop novel, sustainable solutions under uncertainty.

## REFERENCES

1. Industrial Decarbonization Roadmap, US DOE, Sep. 2022
2. MMF Hasan, MS Zantye, MK Kazi. Challenges and Opportunities in Carbon Capture, Utilization and Storage: A Process Systems Engineering Perspective. *Computers & Chemical Engineering*, 107925
3. A. Hasanbeigi et al, 2021, Electrifying U.S. Industry: A Technology- and Process-Based Approach to Decarbonization
4. Zantye, MS, Arora, A, Hasan, MMF. Renewable-integrated flexible carbon capture: a synergistic path forward to clean energy future. *Energy & Environmental Science* 14 (7), 3986-4008
5. MS Zantye, A Gandhi, M Li, A Arora, PS Sengalani, Y Wang, SP Vudata, Bhattachariyya, D, Hasan, MMF. THESEUS: A techno-economic design, integration and downselection framework for energy storage. *Energy Conversion and Management* 284, 116976

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Design for Flexibility: A Robust Optimization Approach

Jnana Sai Jagana<sup>a</sup>, Congqin Ge<sup>b</sup>, Zhihong Yuan<sup>b</sup>, Satyajith Amaran<sup>c</sup>, and Qi Zhang<sup>a\*</sup>

<sup>a</sup> Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, MN 55455, USA

<sup>b</sup> Department of Chemical Engineering, Tsinghua University, Beijing 100084, China

<sup>c</sup> The Dow Chemical Company, Lake Jackson, TX 77566, USA

\* Corresponding Author: [qizh@umn.edu](mailto:qizh@umn.edu).

**Keywords:** Design Under Uncertainty, Optimization

Flexibility is a critical feature of any industrial system as it tells us about the range of conditions under which the system can effectively and safely operate. It is becoming increasingly important as we face greater volatilities in market conditions, diverse customer needs, more stringent safety and environmental regulations, the growing use of resources with varying availability such as renewable energy, and an increased likelihood of disruptions caused by, for example, extreme weather.

In process systems engineering, there is a long history of applying *flexibility analysis* [1] to evaluate and design flexible process systems. In particular, Swaney and Grossmann [2] introduced the concept of flexibility index to quantify the level of operational flexibility in a given design. Here, one assumes that the system's input parameters can take values within certain ranges defined by a so-called uncertainty set, which typically takes the form of a hyperrectangle. The size of the uncertainty set is given by a scalar parameter, and for a given design, the flexibility index is the largest possible value of that size parameter for which the operation of the system is guaranteed to be feasible with proper adjustment of the control variables. In addition, we often want to optimize the design of a system while satisfying given flexibility requirements. In that case, the uncertainty set is fixed, and the design problem can be formulated as a two-stage robust optimization problem [3].

In this work, we propose a new design approach that can directly determine the design that maximizes the flexibility of the system. It also provides a rigorous way of investigating the trade-off between cost and flexibility, and it can be especially useful when the uncertainty set has a more complex structure with multiple parameters defining its shape and size. We find that the problem can be formulated as a two-stage robust optimization problem with endogenous uncertainty [4]. Its trilevel nature and the decision-dependent uncertainty set give rise to a computationally highly complex optimization problem. For its solution, we apply variants of the column-and-

constraint generation algorithm proposed by Zeng and Wang [5]. We demonstrate the key features and efficacy of this new design approach using several examples, including a computational case study considering the design of a supply chain involving mobile modular plants under uncertainty.

## REFERENCES

1. Grossmann IE, Calfa BA, Garcia-Herreros P. Evolution of concepts and models for quantifying resiliency and flexibility of chemical processes. *Computers & Chemical Engineering* 70:22-34 (2014).
2. Swaney RE, Grossmann, IE. An index for operational flexibility in chemical process design. Part I: Formulation and theory. *AIChE Journal* 31:621-630 (1985).
3. Zhang Q, Grossmann IE, Lima RM. On the relation between flexibility analysis and robust optimization for linear systems. *AIChE Journal* 62:3109-3123 (2016).
4. Zhang Q, Feng W. A unified framework for adjustable robust optimization with endogenous uncertainty. *AIChE Journal* 66:e17047 (2020).
5. Zeng B, Wang W. Two-stage robust optimization with decision dependent uncertainty. *arXiv preprint arXiv:2203.16484* (2022).

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# The Current Status of Process and Product Design – What and How to Teach – and a Vision for the Future

Warren D. Seider<sup>a</sup>, and Daniel R. Lewin<sup>b\*</sup>

<sup>a</sup> Chemical and Biomolecular Engineering, University of Pennsylvania, Philadelphia, USA

<sup>b</sup> Chemical Engineering, Technion, I. I. T., Haifa, Israel

\* Corresponding Author: [dlewin@technion.ac.il](mailto:dlewin@technion.ac.il).

**Keywords:** Process Design, Product Design, Education

## EXTENDED ABSTRACT

Process design is a core component of chemical engineering education and either involves or is followed by an extensive design project in most schools. The design project is often considered a core activity in the education of future chemical engineers because it develops their skills in creative and critical thinking beyond the boundaries of their acquired knowledge, as well as training them in teamwork. Such skills are likely to be crucial to empower students to develop process technologies that respond to the relevant future challenges in process design. These future challenges include accommodating alternative raw materials and energy resources, addressing sustainability concerns, and arranging production schedules that are more flexible. Some schools already integrate certain of these challenges in process design courses and design projects, e.g., water and energy conservation as well as CO<sub>2</sub> capture, storage and utilization, and biochemical manufacturing. At the same time, process design tools are also evolving, including, for example, increased emphasis on combination with data-based methods, digital twin concepts, or integration with virtual and augmented reality tools.

On the other hand, only a small fraction of chemical engineering departments teach product design. With industrial diversification, it is of growing importance to teach techniques for selecting products that satisfy consumer needs while incorporating the latest new R&D technologies. Our presentation will describe methods for helping undergraduate students learn the technologies sufficiently well to incorporate them in designing new products. Here expert faculty and industrial persons, along with doctoral researchers, can formulate timely product design problems, while helping students obtain solutions. Clearly, new products should be accompanied by process designs, but product-design strategies are emphasized. Also, initially, methods for selecting new

products using well-known technologies are emphasized. Furthermore, product-designs that involve AI and account for the environment and sustainability can be considered. Finally, reasons many departments teach only process design are covered. For FOCAPD 2024, an extensive report, prepared by faculty and industrial persons, describing the latest approaches, will be summarized. The report, Teaching Chemical Product Design, will have been circulated worldwide with a survey concerning product design – the results of which will be described at the conference.

To summarize our combined position, we propose to present the state of the art in teaching process and product design, and to begin a discussion on how to achieve these objectives within the timeframe that is allocated to chemical process/product design instruction.

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Forces Shaping the Future of Design and Design Education

Jeffrey J Sirola

Davidson School of Chemical Engineering, Purdue University, West Lafayette, Indiana, USA

**Keywords:** Process Design, Process Synthesis, Parameter Optimization, Structural Optimization, Carbon Dioxide Capture, Process Electrification, Hydrogen

## CHEMICAL PROCESS DESIGN

All ABET-accredited engineering programs mandate a culminating major design experience based on knowledge and skills acquired in earlier course work and incorporating realistic appropriate engineering standards and multiple realistic constraints. Some chemical companies organize their Manufacturing Innovation Process into a sequence of stages which typically include Need Identification, Product Design, Basic and Detailed Chemistry, Process Design, Equipment Design, Plant Design, Detailed Engineering and Vendor Specifications, Component Acquisition, Plant Construction Planning and Execution, Operating Procedure Development, Plant Commissioning and Start-up, and Production Planning, Scheduling, and Operation. Each of these stages involve the solution of many "design" problems that could be the subject of the culminating undergraduate chemical engineering design experience.

Most chemical engineering programs historically have chosen Process Design and Equipment Design for the capstone design experience as these aspects incorporate more of the science and engineering principles (mass and energy conservation, classical, solution, and reaction thermodynamics, transport phenomena, kinetics, separations, unit operations, control, safety, etc.) and computational tools (equation solvers, physical property prediction, process and fluid dynamics simulators, etc.) taught within undergraduate chemical engineering curricula. However, some programs have chosen instead a Product Design experience, especially if the product is itself a process (like an analytical instrument, oxygen concentrator, etc.) or if product fitness-for-use depends on its performance in a process (a solvent, mass separating agent, refrigerant, membrane, catalyst, etc.) or user experience (microstructured consumer products, etc.). Other aspects of the innovation process (e.g., Needs Identification, Basic and Detailed Chemistry, Plant De-

sign, Vendor Specification, Construction Planning, Operating Procedure Development, Production Planning and Scheduling, etc.), while certainly interesting, typically have not been used as the capstone design experience either because they require more specialized mathematics, chemistry, biology, materials science, etc. knowledge than most undergraduates are exposed to, or because they require prerequisite Process and Equipment Design.

## Process Design at Purdue

At Purdue our chemical engineering senior design experience has focused on Conceptual Process Design and Economic Evaluation with some aspects of Plant Design, Plant Wide Control, and Process Safety. This includes Equipment Design at different levels of detail appropriate for conceptual economic evaluation and sometimes for vendor specification. Our subject processes may be of any scale, usually continuous but can be batch, and of a complexity that can be reasonably solved by teams of four or five students within one semester. We expect this course scope to continue into the future.

We teach a design paradigm (using terminology appropriate for Process Design, but applicable in general to many design problems within the innovation process) involving Formulation, Synthesis, Analysis, Evaluation, and Optimization.

As might be expected, we have a somewhat greater emphasis on Process Synthesis (generation of process flowsheet alternatives by identifying tasks required to achieve objectives, phenomena to be exploited to accomplish these tasks, equipment and utilities required to implement these tasks, and interconnections among those equipment) than most process design courses or textbooks. We describe three basic approaches to Process Synthesis: Evolutionary Modification (alteration of existing flowsheets), Superstructure Optimization (simultaneous selection of equipment, equipment interconnections, and operating conditions with a purported optimizing simulator), and Systematic Generation

(hierarchical specification of property difference resolving tasks between raw materials and desired products). We make a specific distinction between the Tasks that must be completed in a process and the Equipment in which these tasks will be accomplished. We teach both goal-oriented and constraint-oriented (largely the result of nonideal solution thermodynamics) strategies for identifying process tasks as well as patterns of interconnected tasks that have previously found applicability in specific situations, for example for breaking homogeneous and heterogeneous azeotropes. Our process design approach is specifically task-oriented rather than equipment-oriented. We further take the position that process designs have Primary Objectives (make the right stuff, in the right amount, fit for use - not too difficult; even a computer can do it) while at the same time optimizing as best as possible Secondary Objectives (economics, environmental impact, safety, controllability, sustainability, societal impact, etc. - the real challenge). We note that the relative importance of some of these secondary objectives are currently changing, and new emphasis especially on the environment is the principal challenge impacting the future of chemical process design.

Our approach to Process Analysis and Evaluation is fairly conventional, involving computer-simulation-assisted heat and material balancing about each proposed task in the synthesized flowsheet, choosing operating parameters and finding annual utility requirements and costs for each task, and flows and other properties of the interconnecting streams among the tasks to meet the Primary Design Objectives. Then an initial assumption is made that each task can be implemented in a single piece of equipment after which that equipment is designed at an appropriate conceptual level of detail using (sometimes computer-assisted) methods learned throughout the curriculum just sufficient to estimate a preliminary acquisition and installation cost. Then these equipment capital costs are combined with utility and other annual operating costs in an appropriate way to produce an economic Figure of Merit, initially simple Total Annual Cost (or Levelized Cost) and later Net Present Cost better accounting for variable cash flow timing and financing, depreciation, and income tax implications,

At this point we turn our attention to Optimization of the Secondary Design Objectives. Students first start with Design Heuristics (published results of previous equipment selection and parameter optimization experience) for example, best equipment type selection in particular situations, optimum approach temperatures, operating number of stages and reflux ratios, multistage compression pressure ratios, economic pipe velocities, etc. These will be starting points for more detailed model-based parameter optimizations with realistic objective functions typically economic considering both capital and annual operating costs, but other objectives including

improved safety or environmental impact minimization may also be appropriate. Preliminary safety studies such as Failure Modes and Effects Analysis are also performed early, especially since proposed response to identified critical issues may involve changes to the process structure or operating conditions.

In addition to Design Parameter Optimizations, we also consider what we call Structural Parameter Optimizations. These are based on consideration of some sixty thought-provoking ideas we have accumulated related to reaction path, input-output-recycle structure, heat and power management and thermodynamic efficiency (including heat and power integration algorithms and process condition adjustment to increase the opportunities for such integration), alteration of phase, solution, or reaction equilibria or rate- (and size-) determining mechanisms (classical process intensification), changing the order of tasks and the relationship between tasks and the equipment used to execute them (another form of process intensification), advantaged scale-up or scale-down, cost reduction, and improved safety, operability, controllability, and maintainability. This broader consideration of structural parameter optimization all involves synthesizing local alternatives (for example, a proposed heat integration, or compressor staging with intercoolers, or enhanced heat transfer device, or alternative solvent, or combined reaction and separation, etc.) with selection then based on capital and operating cost economic evaluation or other appropriate criteria. Finally, development of equipment layout, plot plans, and a plant-wide control strategy complete our process design approach. Although not required for conceptual capital cost estimation and more appropriate for vendor specification, sometimes more detailed design of selected equipment such as shell-and-tube heat exchangers or distillation columns using more of the methods described in textbooks and handbooks is performed for the learning experience.

## WHAT IS CHANGING

Not too long ago, after not discovering any new large oil or gas fields for a few decades, there was prediction of "peak oil" and a time when we would run out of finite fossil reserves not only for fuels but also as organic chemical industry feedstocks exacerbated by higher prices as worldwide demand approached known productive capacity. While some suggested a return to coal (whose reserves appeared to be an order of magnitude greater than those of oil and gas), others advocated a switch to "renewable" (recently alive biological) feedstocks ultimately derived from continuing photosynthesis. As these renewable materials, for example cellulose, starch, lignin, triglycerides, etc., are compositionally different from traditional coal, petroleum, or natural gas, it was thought that entirely new processes would need to

be developed, or entirely new products more appropriate to renewable raw materials would need to be developed to meet society's needs (especially for fuels, but other products as well). As these renewable raw materials were created by biological reactions, it was assumed that biological reactions may also be important in their transformation to products. As a result, biochemistry was added to the scientific foundation in many chemical engineering curricula and fermentation was added to the list of unit operations studied. While the process design paradigm remained applicable essentially unchanged, the chemistries and products to which it could be applied in practice and in the classroom were significantly expanded.

Then, with the successful development of directional drilling, logging-while-drilling, slickwater hydraulic fracturing, and microseismic monitoring technologies to extract fossil hydrocarbons from impermeable shale formations, shale resources became included within exploitable reserves expanding them to more than a century of projected demand. Peak oil never happened, and prices fell (and in the case of natural gas, to half its historical ratio compared to oil). Access to hydrocarbon feedstocks for both fuels and chemicals no longer appears to be an immediate problem.

## Carbon Dioxide Emissions

However, continued emission of carbon dioxide to the atmosphere is. These emissions principally come from burning carbonaceous fuels for heating, power (electricity production), transportation, and from certain chemistries (notably the smelting of ores, limestone decomposition, and the production of hydrogen). These emissions are measurably increasing the concentration of CO<sub>2</sub> in the atmosphere which in turn is causing deleterious climate change. Carbon dioxide emissions, and not feedstock availability, is the real sustainability issue facing society in general and the processing industries in particular.

If the carbon source resulting in CO<sub>2</sub> emissions were biological ("renewable"), then some would consider such a process to be "carbon neutral" just recycling CO<sub>2</sub> back to where it came, although others would argue that such a process should not be considered carbon neutral until the biological feedstock is actually regrown.

If CO<sub>2</sub> emissions are from a stationary source like an electricity-producing power plant boiler, utility steam boiler, furnace, fired heater, lime kiln, etc., in principle they could be captured, separated from other things with which they may be admixed, and permanently sequestered from the atmosphere, for example by injection into a suitable capped porous geological formation, dissolution in a deep saline aquifer, or reaction with an appropriate alkaline geologic material such as basalt. This is not widely done at present, but to do so does not involve any new process design paradigm.

If CO<sub>2</sub> emissions are from a mobile source, carbon capture may not be practical. Either some other non-CO<sub>2</sub>-emitting process must be used, or the emitted amount of CO<sub>2</sub> must be captured from the atmosphere (true for any CO<sub>2</sub>-emitting process). CO<sub>2</sub> must also be captured from the atmosphere if it is desired to reduce atmospheric concentrations from existing levels.

## Process Electrification

Since most CO<sub>2</sub> emissions are currently from the combustion of carbonaceous fuels for heat and power (space heating, industrial heating, electricity generation, and transportation), there is interest in substituting carbon-emissions-free electric power in as many of these applications as possible. Carbon-emissions-free electricity does exist from nuclear, hydroelectric, wind, photovoltaic, and geothermal sources (and stationary fossil-fired power plants might also be included if the resulting CO<sub>2</sub> were captured and sequestered). Electricity is already widely used for space heating and also increasingly for short distance transportation. Until now because of cost, electricity is not much used for heating in the process industries, with the exception of electric arc furnaces in steel production. That is about to change, and that will be one of the biggest factors affecting the future of the process industries.

## Hydrogen

Hydrogen is a major chemical mostly used for desulfurization in petroleum refineries and for ammonia and methanol production. It is now also being considered as a carbon-emissions-free fuel for all sorts of heating applications in furnaces, for power production, and for transportation. Currently, hydrogen is usually made from high temperature endothermic steam reforming of a carbonaceous feedstock (methane, light naphtha, coal, etc.) resulting in significant CO<sub>2</sub> coproduction. Although not yet much practiced, this CO<sub>2</sub> could be captured and sequestered ("blue hydrogen"), or hydrogen could be made by water electrolysis with carbon-emissions-free electricity using a number of different electrolyzer technologies ("green hydrogen").

## Direct Air Capture

If to offset CO<sub>2</sub> emissions to the atmosphere an equal amount of CO<sub>2</sub> is removed from the atmosphere (Direct Air Capture), then how should this CO<sub>2</sub> be captured, and what should be done with this captured atmospheric CO<sub>2</sub>? Both biological (agricultural, silvicultural, aquicultural, etc.) and technical (absorption, adsorption, etc.) carbon capture processes are possible. Then after capture, this CO<sub>2</sub> could be sequestered or it could be converted into the very products that resulted in CO<sub>2</sub> emissions to the atmosphere, although this latter choice would generally involve chemical reduction, likely with carbon-emissions-free hydrogen (a new demand for

hydrogen).

## THE FUTURE OF PROCESS DESIGN

What does all this likely mean for the future of process design?

First, many new processes, using new chemistries, and involving new raw materials (possibly biological where that may make sense) will be developed. But not necessarily any new theory of process design.

Second, energy will become more expensive (because of carbon capture and sequestration or CO<sub>2</sub> emissions avoidance). That may lead to even more consideration of energy-saving structural optimizations involving heat integration and process intensification. Even so, it is not necessarily certain that there will be a fundamental shift in the ratio of operating costs to capital costs, since historically as process energy costs have risen (for whatever reasons), so too soon thereafter have process capital costs (because energy is the major economic input in producing the material for process equipment).

Third, carbon-free electricity will replace fossil fuel combustion for space heating and short distance transportation throughout society. This will approximately double total electricity demand requiring a corresponding increase in total electrical generation and distribution capacity (with any fossil-fired capacity either retired or retrofitted with carbon capture and sequestration).

Fourth, the process industries will become electrified for heat as they already have increasingly become for work. Electric heating will be implemented in numerous ways including resistive heating, inductive heating, dielectric (microwave) heating, plasma heating, arc heating, etc. Much of this detail will be new to chemical process and equipment designers. Electricity may be used to boil steam (so that chemical processes heated by steam condensation remain unchanged), or to directly heat equipment (perhaps more precisely) including replacement of fired heaters and furnaces for temperatures greater than steam.

Fifth, chemical processes are unique in that many kinds of operations like distillation and absorption both require heat at higher temperature and reject nearly the same quantity of heat at lower temperature. These situations may be appropriate for electricity-powered heat pumps moving heat from the lower temperature source to the higher temperature requirement (generally more efficient than straight electrical heat itself) which heretofore have been practically limited to the distillation of close-boiling mixtures but whose applicability may be much expanded given CO<sub>2</sub> emission elimination considerations.

Sixth, hydrogen will be more in demand both as a chemical reducing agent and as a carbon-emissions-free fuel for both mobile and stationary applications. It remains to be seen whether that hydrogen is best

produced by carbon-free electrolysis or carbon-free electrically-powered carbon-captured fossil reforming.

Seventh, powerful oxidants oxygen and chlorine are already produced by electrically powered air separation and chloride electrooxidation. Reductant hydrogen will either be made by water electrolysis or electrically powered fossil reformation (with carbon capture). Given these reagents, it is not yet clear if chemical electrosynthesis will find other major advantaged applications.

While process design theory may not change, the future overarching emphasis on carbon dioxide emissions elimination, atmospheric CO<sub>2</sub> concentration reduction, and process electrification will have a significant impact on Structural Optimization objectives in addition to traditional economics, safety, waste minimization, and controllability. This inevitably will result in new and different process flowsheets, product costs different from current experience, and many new problems and case studies for undergraduate design courses. However, our students will be prepared and will be up to the challenge.

## ACKNOWLEDGEMENT

This work was supported by the National Science Foundation under Cooperative Agreement No. EEC-1647722, an Engineering Research Center for Innovative and Strategic Transformation

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# Decarbonization of Oil Refineries through Electrification and Low-Carbon Feedstocks

Ana I. Torres<sup>a\*</sup>

<sup>a</sup> Carnegie Mellon University, Department of Chemical Engineering, Pittsburgh, PA, USA

\* Corresponding Author: [aitorres@cmu.edu](mailto:aitorres@cmu.edu)

**Keywords:** Decarbonization, Green Fuels, Process Design, Oil refineries, Sustainability

Chemical Process Industries must navigate a series of changes in their operations to comply with increasing sustainability targets. These changes may involve the use of electricity-based operations, the implementation of carbon capture strategies, and the use of biomass or end-of-life carbon-containing waste as feedstocks. Decarbonizing oil refineries is particularly challenging as they possess highly valuable infrastructure. Discarding this infrastructure before the end of its life to build entirely new electric and biomass-based operations does not seem to be an economical or even a sustainable solution. This presentation will cover recent work in my group related to the decarbonization of oil refineries, focusing on proposing solutions that could be integrated with existing plants.

In the first part of the talk, we will present a decision-making tool to assist chemical industries in their transition to decarbonization. The proposed tool considers a given process flowsheet, representing current operations in the refinery, a list of decarbonization initiatives based on the electrification of steam and hydrogen production, and carbon capture initiatives, along with costs, expenditure limits, and carbon emission reduction goals. Based on the available decarbonization options and existing plant needs, we propose a superstructure and model it as a multiperiod MILP. The tool returns an optimal technology switch strategy and the timeline for its implementation, where optimal means the one with the lowest net present cost. We will present the results of an oil refinery in the United States that intends to reduce its carbon emissions by half by 2030 and become carbon-free by 2050. We then explore the trends across various scenarios of electricity costs and carbon tax levels, among others, to find those that will result in a more extensive adoption of electrified technologies.

The second part of the talk discusses several works on the use of biomass in oil refineries. We start with a short overview of a project whose objective was to

design a sustainable aviation fuel (SAF) process based on the hydrotreatment of oil seeds that could be adapted to an existing refinery infrastructure. Hydrotreatment of oil seeds results in hydrocarbons that produce green naphtha and green diesel that can be seamlessly blended with oil-based fuels (i.e. drop-in fuels).

Next, we discuss a second project where we expand these ideas to consider other possible biomass and technologies. In this case, we assume six well-established biomass to fuel technologies and identify the oil refinery operations that can be repurposed for biomass processing. A retrofitting problem is formulated to find the optimal facility repurpose strategy that meets a given fuel demand.

Finally, we explore possible synergies between the forest supply chain and existing fossil-based facilities to produce fuels from biomass residues from the first one. We propose a decision-making problem to define the location and possible integration of new facilities to produce intermediates, assuming that traditional pulp-producing and sawmill facilities, as well as refineries and biorefineries, are already operating.

## Acknowledgments

Funding from the following sources is gratefully acknowledged:

- This work was financed [in part] by a grant from the Commonwealth of Pennsylvania, Department of Community and Economic Development, in collaboration with Shell Global
- Carnegie Mellon University, Ana I. Torres start-up package.
- ANII (Uruguay) – FSE 1 2018 1 152900

Academic and Industrial Collaborators:

- Prof. I. Grossmann (Carnegie Mellon University), B. Chen and Z. Yuan, (Tsinghua University, China), S.



Gutiérrez, L. Pittaluga, (Udelar, Uruguay), MA  
Rodríguez, (Universidad Nacional de Córdoba and  
CONICET, Argentina)

- Dr. R. Gandhi, Shell Global, USA; Ing. Quím. A.  
Pedezert, ANCAP, Uruguay
- Students: S. Chattopadhyay, V. González, F.  
Piedra-Jiménez, L. Zhang

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE  
Press. This is an open access article under the creative com-  
mons CC-BY-SA licensing terms. Credit must be given to creator  
and adaptations must be shared under the same terms. See  
<https://creativecommons.org/licenses/by-sa/4.0/>



# Challenges in Design for Sustainability

Gavin Towler\*

Honeywell, Sustainable Technologies

\* Corresponding Author: [gavin.towler@honeywell.com](mailto:gavin.towler@honeywell.com)

**Keywords:** Process Design, Renewable and Sustainable Energy

## EXTENDED ABSTRACT

There is a global consensus that steps must be taken to mitigate the impact of anthropogenic climate change. The Paris Agreement on climate change has been ratified by 192 countries and the signatories have pledged to make changes to their patterns of energy and land use that achieve “carbon neutrality” or net-zero emissions of greenhouse gases (GHG) by approximately mid-century. In these countries, energy ministries, energy companies and utilities are evaluating alternative fuels and power sources that can deliver the heat and power required for a modern economy with reduced GHG emissions. While technically proven low-emissions alternatives exist for almost every application, most of these alternatives cost substantially more than the fuels or energy sources they replace. Consequently, most countries will use a combination of regulations, taxes and subsidies to distort the energy market in favor of the lower-emissions alternatives.

Achieving a net-zero GHG economy while meeting the energy demand and economic growth needs of the world’s population requires a substantial transformation of every industry. No human activity is strictly carbon neutral unless it removes sufficient carbon dioxide from the atmosphere to offset the emissions from all the manufacturing, operation and decommissioning steps associated with the activity. Electric power can be produced with relatively low GHG footprint by wind turbines, solar cells, hydroelectric plants, geothermal plants and nuclear power, enabling electrification as a decarbonization strategy in many applications.

While electrification will pay a large role in reducing GHG emissions, there are several industrial and transportation applications that are difficult to electrify. These applications either require high temperature heat that is best delivered by burning a fuel, or else require rapid refueling to maximize capital efficiency of vehicle utilization. In the future, these applications will be served by a mixture of hydrogen, renewable fuels derived from

biological sources, and fossil fuels whose GHG impact has been offset by activities such as land use changes that sequester an amount of carbon dioxide equivalent to the GHG emissions.

The pace of the energy transition will be constrained by the speed at which new vehicle, power generation and fuel production technologies can be deployed. The transition to lower GHG emission transportation modes (battery electric vehicles and hydrogen- or electric-powered trucks) is constrained by both the rate of global fleet turnover and the rate at which the global vehicle manufacturing industry transitions fully to low emissions vehicles. Similarly, the transition to lower GHG forms of industrial heat will be constrained both by the supply of renewable electric power and the availability of low-GHG fuels such as hydrogen.

There are multiple implications for design practitioners, particularly those who are developing large capital projects that must remain economically viable throughout the energy transition period. Some recent case studies from industrial practice will be presented to illustrate how companies are responding to these challenges.

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



# New Design Paradigm for Integrated Energy Systems Needed for Sustainability

David Tucker<sup>a\*</sup>, N. Farida Harun<sup>a,b</sup>, Biao Zhang<sup>a,b</sup>, Nana Zhou<sup>a,b</sup>, and Samuel Bayham<sup>a</sup>

<sup>a</sup> National Energy Technology Laboratory, 3610 Collins Ferry Road, Morgantown, WV 26505, USA

<sup>b</sup> NETL Support Contractor, 3610 Collins Ferry Road, Morgantown, WV 26505, USA

\* Corresponding Author: [david.tucker@netl.doe.gov](mailto:david.tucker@netl.doe.gov).

**Keywords:** Intelligent Systems, System Identification, Energy Systems, Process Control, Process Design

## INTRODUCTION

### Achieving Sustainability

Achieving sustainability in the energy sector requires an economically viable path with a balanced transition that does not aggravate environmental and sociological problems associated with current fossil-based power production. Increasing the grid penetration of intermittent renewables to realize a sustainable energy future without consideration of the balanced transition may result in devastating economic and societal impacts [1]. As we press for the minimization of renewable power curtailment, current fossil-based technology struggles to meet demand under extreme transient and part-load conditions. This results in dramatic reduction of efficiency and a corresponding increase in emissions of not only carbon, but far more devastating pollutants.

Without a doubt, the needs of a sustainable future require new technologies, not only for renewable power like wind and solar, but also for dispatchable power [2]. Ideally, technology to serve the transition would provide a significant increase in the efficiency of fossil fuels conversion under transient and part-load conditions as well as an opportunity to use renewable fuels and integrate with existing wind and solar. To meet the aggressive timelines for decarbonization and transition to a renewable energy sector posed by many world governments, these new technologies would need to be developed with an unprecedented haste not supported by our current technology development paradigm [3].

### Integrated Energy Systems

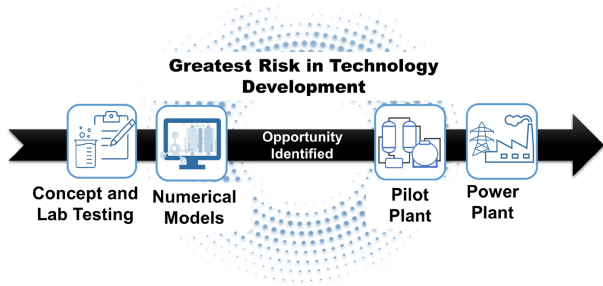
Integrated energy systems provide hope for high efficiency conversion of fuels to electricity, low emissions, load following flexibility, and low-cost carbon management. By integrating a variety of power generation assets, developing technologies can be matched with

mature technologies, mitigating risk and facilitating early adoption of novel component concepts.

As an example, the synergies of a solid oxide fuel cell (SOFC) – gas turbine system through thermal integration result in efficiencies that exceed the sum of the individual parts. Through optimizing the integration of fuel reforming in the cycle, theoretical efficiencies over 70% low heating value (LHV) natural gas can be realized at SOFC fuel utilizations below 45% where the fuel cell represents less than half the power of the cycle [4]. Merging developing (SOFC) and mature (gas turbine) technologies mitigates SOFC early adoption risk by lowering the fraction of power generation and associated costs. Even if the SOFC does not function to its fullest potential, higher efficiencies and turndown capabilities can be realized through the thermal energy storage capacity of the SOFC materials, facilitating the transition to a sustainable future. Similarly, concentrated solar power and thermal energy storage can be coupled to the mature gas turbine technology.

### Co-Design

Although many benefits of integrated energy systems have been extolled in the literature and even in distributed scale demonstrations, system integration and dynamic control issues inhibit commercial adoption of these technologies. In the current technology development paradigm, integration and control issues are not often identified until the completion of a pilot scale system [3,5], as shown in the abbreviated technology development timeline in Figure 1. Failure at the pilot scale stems further investment and shatters all hope for commercialization efforts. As such, the bridge between numeric models and the hardware of a pilot plant represents the greatest risk in technology development and the opportunity for the greatest gain in accelerating the path to commercialization of new concepts.

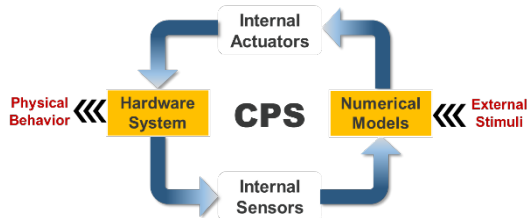


**Figure 1.** Current paradigm of technology development

The simultaneous co-design of system components, system integration, and dynamic control would be required to minimize development risks. If a pilot plant could be built with some confidence of the operability and potential for automation, the chances for success would be greatly improved. This could be best accomplished if component design could be changed based on identified needs for system integration and controls. To realize effective co-design, the use of new tool would be required [3,6,7].

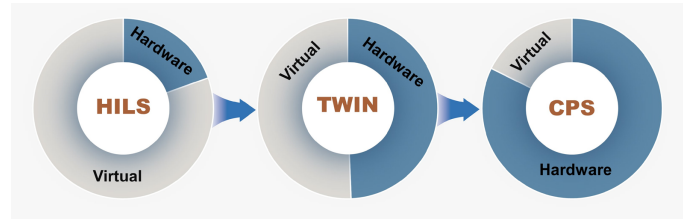
### Cyber-Physical Systems

When considering the design tools needed in the transition from numeric models to pilot-scale hardware, cyber-physical systems (CPS) come to the forefront as a method to model complex integrated energy systems. As shown in Figure 2, CPS interact with a physical environment through a seamless combination of numeric models and hardware through a series of sensors and actuators.



**Figure 2.** Interaction between physical environment and real-time numerical models in cyber-physical systems

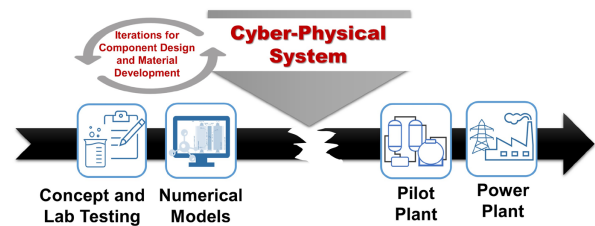
Simulations using cyber-physical models are distinguished from hardware-in-the-loop simulations (HILS) that couple a physical component to a virtual environment for testing. Although commercial digital twin (Twin) products also couple numeric models to hardware, CPS requires the models to converge within the sampling frequency of the dynamic controller. In power systems, Twins generally require convergence within the optimization time of a supervisory controller, or about 5 to 15 minutes. The CPS facility at NETL requires convergence in 5 milliseconds. Concepts of HILS, Twin and CPS are compared in Figure 3.



**Figure 3.** Illustration of the concepts of hardware-in-the-loop, digital twins, and cyber-physical system

### Changing the Paradigm

To minimize the highest risk or “valley of death” in the technology development, the current paradigm must be modified to accommodate co-design by including a feedback loop in the highest risk part of the design process [3,5]. As shown in Figure 4, if CPS techniques could be employed before investment in a full pilot scale power system, opportunities to modify materials and bench scale systems would be available to maximize success at the pilot scale.



**Figure 4.** De-risk technology development through cyber-physical systems

The new design paradigm for integrated energy systems of the future must include pathways that make use of dynamic models, hardware-in-the-loop simulation, digital twins, cyber-physical models, and ultimately, cyber-physical systems to support intelligent power generation. As shown in Figure 4, filling the “valley of death” with a continuum of simulation methods between numeric models and hardware models can provide an opportunity to achieve commercialization of new energy system technologies within the time scale needed to realize economically viable energy transitions on an international scale.

### ACKNOWLEDGEMENTS

This work was done with support and funding from the Advanced Sensors, Controls, and Novel Concept Program and the Solid Oxide Fuel Cell Research Program, Office of Fossil Energy and Carbon Management, U.S. Department of Energy.

## DISCLAIMER

This project was funded by the United States Department of Energy, National Energy Technology Laboratory, in part, through a site support contract. Neither the United States Government nor any agency thereof, nor any of their employees, nor the support contractor, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## REFERENCES

1. Moore, F. C. *et al.* Determinants of emissions pathways in the coupled climate–social system. *Nature* **603**, 103–111 (2022).
2. McDonell, V. *et al.* Hydrogen Based Energy Storage System for Integration with Dispatchable Power Generator (Phase I Feasibility Study). <https://www.osti.gov/biblio/1874681> (2022).
3. Tekaat, J. L. *et al.* The Paradigm of Design Thinking and Systems Engineering in the Design of Cyber-Physical Systems: A Systematic Literature Review. 2021 IEEE International Symposium on Systems Engineering (ISSE), Vienna, Austria, pp. 1–8, doi: 10.1109/ISSE51541.2021.9582548 (2021).
4. Chen, H. *et al.* High efficiencies with low fuel utilization and thermally integrated fuel reforming in a hybrid solid oxide fuel cell gas turbine system. *Appl. Energy* **272**, 115160 (2020).
5. Garcia-Sanz, M. Control Co-Design: An engineering game changer. *Adv Control Appl.*; 1:e18. <https://doi.org/10.1002/adc2.18> (2019).
6. Ruchkin, I. Integration Beyond Components and Models: Research Challenges and Directions. *In Proceedings of the 3th Architecture Centric Virtual Integration Workshop (ACVI)* (in conjunction with WICSA/CompArch). Venice, Italy (2016).
7. M. Torngren, D. *et al.* Tool supporting the co-design of control systems and their real-time implementation: Current status and future directions. 2006 IEEE Conference on Computer Aided Control System Design, 2006 IEEE International Conference on Control Applications, 2006 IEEE International Symposium on Intelligent Control, Munich, Germany, 2006, pp. 1173–1180, doi: 10.1109/CACSD-CCA-ISIC.2006.4776809 (2006).

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>





# Design Education Across the Curriculum for the Future of Design

Ashlee N. Ford Versypt\*

University at Buffalo, The State University of New York, Department of Chemical and Biological Engineering, Buffalo, NY, USA

\* Corresponding Author: [ashleefv@buffalo.edu](mailto:ashleefv@buffalo.edu)

**Keywords:** Education, Modelling and Simulations

## EXTENDED ABSTRACT

The future of computer-aided process design hinges on continued recruitment, training, and retention of the next generations of engineers. Many elementary and secondary school programs focused on engineering have made substantial impacts in informing children about careers in science, technology, engineering, and mathematics (STEM). A report by the National Academies established three general principles for pre-college engineering education, the first of which is that elementary and secondary engineering education should emphasize engineering design<sup>1</sup>. Curricula focused on teaching the engineering design process have been developed for K-12 students and educators<sup>2,3</sup>. These materials often emphasize engineering design as an engaging process with iterative prototype testing and a supportive process where learning from failure is encouraged in a positive way. However, these engineering design process concepts have not necessarily been communicated to university-level engineering faculty (in contrast to the scientific method). This is a problem when college students only see design as a high-stakes, time-intensive senior capstone project or as highly specialized content siloed in upper-division courses with design explicitly in the name such as Plant Design, Process Design, or Product Design.

Despite upward trends for exposure to engineering design in pre-college settings and for students majoring in engineering more broadly, particularly computer science, a concerning trend is the decline in chemical engineering enrollments<sup>4</sup>. Computer-aided process design presents the opportunity to integrate design into engineering education at the pre-college and undergraduate levels while also leveraging strong interests in computational sciences and data-oriented careers. This presentation will focus on surveying several design project activities from the author and others<sup>5-9</sup> aimed at high

school and early college students to teach chemical engineering principles and attract them to and retain them in chemical engineering. Recommendations for integrating design across the curriculum through hands-on and computational activities will be also discussed to support the future of computer-aided process design.

## REFERENCES

1. National Academy of Engineering and National Research Council. Engineering in K-12 Education: Understanding the Status and Improving the Prospects. The National Academies Press (2009).
2. TeachEngineering.org
3. Bayles TM, Enszer JA, Ross JM. Incorporating engineering design into high school STEM initiatives. *ASEE Annu Conf Expo* (2012).
4. Rhinehart RR. How to prepare for the ChE undergraduate enrollment trend. *Chem Eng Progress* 118:50-55 (2022).
5. Eastep CV, Harrell GK, McPeak AN, Ford Versypt AN. A MATLAB app to introduce chemical engineering design concepts to engineering freshmen through a pharmaceutical dosing case study. *Chem Eng Ed* 53:85-90 (2019).
6. Ford Versypt AN. Choose your own kinetics adventure: student-designed case studies for chemical reaction engineering course projects. *Trans Techniques STEM Ed* 3:48-56 (2017).
7. Butterfield AE, Branch K, Trujillo E. First-year hands-on design course: implementation and reception. *Chem Eng Ed* 49:19-26 (2015).
8. Tirapelle M, Chia DN, Duanmu F, Katsoulas K, Marchetto A, Sorensen E. Practical learning activities to increase the interest of university applicants in STEM careers in the era of industry 4.0. In: Kokossis AC, Goergiadis MC, Pistokopoulos E (Eds). 33rd European Symposium on Computer

Aided Process Engineering. (pp. 3465-3470)  
Elsevier (2023).

9. Bayles TM, Work in progress: twenty year evolution of an outreach program. *ASEE Annu Conf Expo* (2019).

---

© 2024 by the authors. Licensed to PSEcommunity.org and PSE Press. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>



---

# Index

Author Index

---

# Author Index

- Aboagye, Emmanuel A., **621**  
Adams, Thomas A., II, **2, 484**  
Adeyemo, Samuel, **99**  
Adjiman, Claire S., **145**  
Adnan, Muhammad, **186**  
Agi, Damian T., **395**  
Agrawal, Rakesh, **527, 883**  
Aguirre-Villegas, Horacio, **732**  
Ahmed, Abdulhakeem, **629**  
Akintola, Ayooluwa, **387**  
Akkor, Ilayda, **635**  
Akula, Paul, **222**  
Ali, Mohamed, **245**  
Allan, Douglas A., **448, 504, 950**  
Allen, R. Cory, **237**  
Allman, Andrew, **920**  
Almajed, Hussain M., **641**  
Al-Mohannadi, Dhabia M., **726, 783, 792, 844**  
Al-Rawashdeh, Ma'moun, **167**  
Al-Sakkari, Eslam G., **245**  
Amaran, Satyajith, **997**  
Amazouz, Mouloud, **245**  
Amusat, Oluwamayowa O., **253, 267**  
Anantharaman, Rahul, **215**  
Anglou, Elisavet, **652**  
Araia, Alazar, **316**  
Aras, Chinmay M., **105**  
Arnold, Travis, **941**  
Arwa, Erick O., **403**  
Assen, Niklas von der, **585, 604**  
Atia, Adam A., **253, 490**  
Avendano, Marco E., **660**  
Avraamidou, Styliani, **598, 820**  
Bachawala, Sunidhi, **194**  
Bakshi, Bhavik R., **10, 373**  
Baldea, Michael, **570**  
Ball, Madelyn R., **959**  
Bao, Jie, **387**  
Bardow, André, **585, 876**  
Barhate, Yash, **261**  
Barochia, Dev, **705**  
Bartholomew, Timothy, **253**  
Baumgärtner, Nils, **585**  
Bavarian, Mona, **367**  
Bayham, Samuel, **1006**  
Beacham, Paloma, **186**  
Benavides, Pahola Thathiana, **691**  
Beykal, Burcu, **16, 705**  
Bhattacharya, Riddhi, **652**  
Bhattacharyya, Debangsu, **99, 222, 330, 448, 504**  
Biegler, Lorenz T., **448, 504, 546, 829, 860**  
Bjarnason, Aaron, **308**  
Boffito, Daria C., **245**  
Bojovic, Vesna, **570**  
Bollas, George M., **409, 771**  
Borisut, Prapatsorn, **668**  
Boukouvalla, Fani, **351, 652**  
Boush, David, **380**  
Brown, Cameron J., **308**  
Bugosen, Sergio I., **113**  
Burrows, Laron D., **409**  
Cabeza, Andres F., **121**  
Campos, Gustavo, **417**  
Cansino-Loeza, Brenda, **674**  
Capellades, Gerard M., **283**  
Carrillo, E.J., **681**  
Casas-Orozco, Daniel, **261**  
Chattopadhyay, Sampriti, **426**  
Chea, John D., **60, 691**  
Chen, Xinhe, **434, 490**  
Chen, Yunxiang, **387**  
Chrisandina, Natasha J., **698**  
Chung, Hegwon, **442, 541**  
Conway, Matthew, **691**  
Cortés-Peña, Yoel R., **129**  
Cremaschi, Selen, **344, 668, 750, 892**  
Cresko, Joe, **533**  
Cui, Chengtian, **612**  
Curtis, Sean, **137**  
Dabadghao, Vibhav, **504**  
Dagdougui, Hanane, **245**  
Daoutidis, Prodromos, **807**  
Deshpande, Anuja, **950**  
Diaz, M. Soledad, **546, 868**  
Ding, Chaoying, **67**  
Dinh, San, **504**  
Donnelly, Kevin B., **373**  
Dosso, Cheick, **860**  
Dougher, Molly, **711**  
Dowdle, John, **635**  
Dowling, Alexander W., **22, 308, 395, 434, 490, 711, 927, 966**  
Drouven, Markus, **829, 941**  
Dudchenko, Alexander V., **253, 267**  
Duewall, Clinton M., **719**  
Durand, Helen, **186**  
Elfaki, Hesam, **726**  
El-Halwagi, Mahmoud M., **698, 719**  
Endler, Elizabeth E., **570**  
Erickson, Evan D., **732**  
Esche, Erik, **153, 179**  
Estrada, Vanina, **868**  
Fallanza, Marcos, **904**  
Farlessyost, William, **275**  
Federici, Justin A., **32**  
Feng, Xiao, **562**  
Flores, Betsie Montano, **461**  
Foster, Charles, **933**  
Fox, James, **44**  
Fracchiolla, Michael, **137**  
Fu, Qiang, **660**  
Galdos, Francisco, **344**  
Gandhi, Akhilesh, **469**  
Gandhi, Rahul, **426**  
Gao, Barnabas, **137**  
Gao, Peiyuan, **387**  
Gazzani, Matteo, **215**  
Ge, Congqin, **997**  
Geissler, Caleb H., **738**  
Gettelfinger, Brian T., **991**  
Ghanem, Fred, **283**  
Ghouse, Jaffer H., **993**  
Giridhar, Nishant V., **448, 504**  
González, Leonardo D., **744**  
Gonzalez, Marcial, **194**  
Gooty, Radhakrishna Tumbalam, **490**  
Gopinath, Smitha, **145**  
Gounaris, Chrysanthos E., **201, 635**  
Grossmann, Ignacio E., **426, 853, 904**  
Guerra, Omar J., **641**  
Guillén-Gosálbez, Gonzalo, **39**  
Guittet, Darice, **417, 434**  
Gundersen, Truls, **562**  
Gunter, Dan, **966**  
Gupta, Shaswat, **338**  
Hagedorn, Dörthe Franzisca, **604**  
Harjunkoski, Iiro, **455**  
Hartmann, Jan, **585**  
Harun, N. Farida, **1006**  
Harvell, Justin, **344**  
Hasan, M. M. Faruque, **105, 292, 469, 994**  
Hawa, Hani A. E., **395**  
Herdiac, Monica A., **570**  
Hernandez-Betancur, Jose D., **60**  
Hernandez-Torres, Eduardo, **380**

Hesketh, Robert P., **137, 621**  
 Hightower, Ryan, **186**  
 Hildebrandt, Diane, **44**  
 Hodge, Bri-Mathias, **641**  
 Hoffmann, Christian, **153**  
 Holly, Marcus, **490**  
 Holm, Sigmund Eggen, **554**  
 Hopkinson, David, **860**  
 Hosseini, Erfan, **750**  
 Hu, Jianli, **316**  
 Huber, George W., **820**  
 Huynh, Dat T., **757**  
 Iakovou, Eleftherios, **698**  
 Ibáñez, Raquel, **904**  
 Ierapetritou, Marianthi, **67, 229, 757, 933**  
 Iftakher, Ashfaq, **105, 292**  
 Ikegwu, Ugochukwu M., **763, 814**  
 Irhamna, Adrian R., **771**  
 Isenberg, Natalie M., **201**  
 Iyengar, Arun, **387**  
 Iyer, Shachit S., **635**  
 Jaffe, Tyler, **490**  
 Jagadeesan, Prem, **778**  
 Jagana, Jnana Sai, **997**  
 Jamdade, Shubham, **173**  
 Jan, Sherzoy, **208**  
 Jerpoth, Swapana, **137**  
 Jiang, Yazhou, **387**  
 Jiang, Zheyu, **322**  
 Jiang, Zhihua, **750**  
 Jogwar, Sujit, **912**  
 Jung, Howoun, **160**  
 Kafaro, V., **681**  
 Kakodkar, Rahul, **461**  
 Karimi, Kian, **215**  
 Katebah, Mary A., **167**  
 Kazi, Monzure-Khoda, **469**  
 Kim, Jinsu, **173**  
 Kim, Jiyong, **442, 541**  
 Kim, Sunwoo, **477**  
 Kim, Youngdae, **237**  
 Knezovic, Katarina, **455**  
 Knueven, Bernard, **434**  
 Kodate, Purnima M., **283**  
 Koepfel, Brian, **387**  
 Kontoravdi, Cleo, **75**  
 Krishnadoss, Rajalakshmi, **484**  
 Krone, David, **179**  
 Kudva, Akshay, **373**  
 Kusuma, Victor, **860**  
 Lair, Laurianne, **711**  
 Laird, Carl D., **113, 208, 853**  
 Laky, Daniel J., **261, 308, 490**  
 Laliwala, Chris, **496**  
 Lameh, Mohammad, **783**  
 Lao, Jianpei, **660**  
 Lee, Andrew, **950, 966**  
 Lee, Carissa Anne, **344**  
 Lee, Jay H., **160, 477**  
 Lehn, Reid C. Van, **763, 814**  
 Lehr, Austin L., **621**  
 Leonard, Akkarakaran Francis, **186**  
 Lewin, Daniel R., **998**  
 Li, Jianping, **300**  
 Li, Mingrui, **448, 504**  
 Lin, Yilun, **461**  
 Linke, Patrick, **167, 783, 799**  
 Lipke, Elizabeth A., **344**  
 Liu, Andrew, **387**  
 Lizcano-Prada, J., **681**  
 Longo, Jared, **621**  
 Lopez, Blake R., **975**  
 Luo, Yuqing, **229**  
 Lynch, Hailey G., **308**  
 Ma, Jiase, **732**  
 Macdonald, Ruaridh, **511**  
 Machado, Pedro G., **927**  
 Majzoub, Wafaa N., **792**  
 Malinov, Nikola, **67**  
 Mallapragada, Dharik S., **511**  
 Mann, Margaret, **417**  
 Maravelias, Christos T., **519, 738**  
 Marina, Olga, **387**  
 Martín, Mariano, **52**  
 Martino, Marcello Di, **799**  
 Masud, Md Abdullah Al, **316**  
 Masuku, Cornelius. M., **612**  
 Matuszewski, Michael, **222**  
 McDonald, Molly A., **519**  
 McIntosh, Vernon, **674**  
 Melnick, Jason, **380**  
 Messina, Dominic, **186**  
 Mills, Celeste, **744**  
 Miraliakbar, Alireza, **322**  
 Mitrai, Ilias, **807**  
 Morgan, Joshua C., **208, 950**  
 Mota-Aguilar, David, **380**  
 Mukherjee, Angan, **330**  
 Munguía-López, Aurora del C., **744, 814, 820**  
 Muñoz, Salvador García, **338, 380**  
 Munoz-Briones, Paola A., **820**  
 Nagy, Zoltan K., **194, 261, 359**  
 Naik, Sakshi, **829**  
 Nair, Sankar, **660**  
 Naraharisetti, Pavan Kumar, **836, 983**  
 Nascimento, Cláudio A.O., **927**  
 Ncube, Selusiwe, **578**  
 Neira, David E. Bernal, **121, 904**  
 Nicholson, Bethany, **966**  
 Nielsen, Robert D., **32**  
 Nieman, Kip, **186**  
 Nikkhah, Hasan, **705**  
 Nimbalkar, Sachin U., **533**  
 Nkoutche, Catherine, **698**  
 Nogaja, Akash, **527**  
 Noring, Alexander, **387**  
 Nuchitprasittichai, Aroonsri, **668**  
 O'Neill, Renee, **186**  
 Okeke, Ikenna J., **533**  
 Omell, Benjamin P., **222, 950**  
 Oqbi, Manar Y., **844**  
 Orjuela, Alvaro, **121**  
 Ostace, Anca, **950**  
 Oudalov, Alexandre, **455**  
 Ouimet, Jonathan Aubuchon, **711**  
 Ovalle, Daniel, **853**  
 Pajand, Pejman, **570**  
 Palys, Matthew J., **807**  
 Panagakos, Grigorios, **222, 860**  
 Papageorgiou, Dimitri J., **32, 237**  
 Papathanasiou, Maria M., **75**  
 Park, JoungHo, **477**  
 Park, Minseong, **442, 541**  
 Park, NohJin, **160**  
 Parker, Robert B., **113, 966**  
 Paul, Brandon, **387, 950**  
 Paulson, Joel A., **373**  
 Pazik, John, **621**  
 Pedrozo, Hector A., **546, 860**  
 Pellegrini, Laura A., **591**  
 Perez, David, **60**  
 Phillip, William A., **711**  
 Pistikopoulos, Efstratios N., **461, 698, 799**  
 Poon, Sarah, **966**  
 Qiu, Kang, **554**  
 Ragab, Ahmed, **245**  
 Rajendiran, Shenbageshwaran, **344**  
 Ramos, Fernando D., **868**  
 Ramos, Matías H., **868**  
 Raudenbush, Katherine, **67**  
 Ravutla, Suryateja, **351**  
 Realff, Matthew J., **173, 660**  
 Reddy, Jayanth V., **67**  
 Rehner, Philipp, **876**  
 Reinert, Christiane, **604**  
 Reklaitis, Gintaras V., **194, 261**  
 Ren, Siyue, **562**  
 Repke, Jens-Uwe, **153, 179**  
 Rhim, Jungsoo, **359**  
 Rho, Jong Hyun, **570**  
 Ribeiro, Celma O., **927**



Rodriguez-Gil, Edwin A., **883**  
Rodriguez-Vallejo, D., **681**  
Roth, Steven, **137**  
Roussanaly, Simon, **554**  
Ruff, Ronald, **338**  
Ruiz-Mercado, Gerardo J., **60, 691**  
Sachio, Steven, **75**  
Safari, Habibollah, **367**  
Sánchez-Rivera, Kevin L., **820**  
Santhamoorthy, Pooja Zen, **892**  
Sanyal, Oishi, **959**  
Sarkis, Miriam, **75**  
Schell, Kristen R., **403**  
Schilling, Johannes, **876**  
Schweidtmann, Artur M., **84**  
Seider, Warren D., **998**  
Sempuga, Baraka C., **44, 578**  
Sen, Maitraye, **338**  
Shah, Nilay, **75**  
Shah, Utkarsh, **373**  
Shahab, Mohammad, **194**  
Shamlou, Elmira, **941**  
Shellman, Melody H., **941**  
Sherman, Jason A. F., **201**  
Shu, David Y., **585**  
Shumaker, Ethan, **621**  
Sibychan, Jerrin Job, **380**  
Siirola, Jeffrey J, **999**  
Siirola, John D., **201, 434**  
Silva, Aline V. da, **927**  
Singh, Shireen, **344**  
Singh, Shweta, **275, 778**  
Skiborowski, Mirko, **179**  
Slater, C. Stewart, **137**  
Smith, Raymond L., **899**  
Smith, Wilson A., **641**  
Somoza-Tornos, Ana, **641**  
Sorace, Nicola, **380**  
Sousa, Marco De, **461**  
Spatolisano, Elvira, **591**  
Stacey, Neil, **44**  
Stathatou, Patricia, **652**  
Stinchfield, Georgia, **208**  
Straus, Julian, **554**  
Suárez, JP Osorio, **546**  
Subraveti, Sai Gokul, **215**  
Summits, Stephen, **222**  
Tang, Wei-Ting, **373**  
Tarka, Thomas J., **711**  
Tawarmalani, Mohit, **527**  
Ternus, Krista, **674**  
Theuma, David, **137**  
Thirumaran, Kiran, **533**  
Tian, Huayu, **67**  
Tian, Yuhe, **316, 387, 959**

Tominac, Philip, **732, 941**  
Torres, Ana I., **426, 496, 629, 1003**  
Towler, Gavin, **1005**  
Tran, Norman, **853**  
Tristán, Carolina, **904**  
Tucker, David, **1006**  
Tumbalam-Gooty, Radhakrishna, **434**  
Tyrrell, Katie, **186**  
Uday, Vikram, **912**  
Uribe-Rodríguez, A., **546, 681**  
Vedant, Shivam, **698**  
Vercellino, Roberto, **417**  
Vergara-Zambrano, Javiera, **598**  
Versypt, Ashlee N. Ford, **1009**  
Volkmer, Marvin, **604**  
Vyas, Javal, **941**  
Wang, Dewei, **387**  
Wang, Hongxuan, **920**  
Wang, Le, **635**  
Wang, Yifan, **604**  
Wang, Yuxin, **316**  
Watson, Madelynn J., **927**  
Wen, Ching-Mei, **933**  
Williams, Bianca, **668**  
Wu, Sean M., **344**  
Xu, Sidra, **344**  
Xu, Zhijie, **387**  
Yenkie, Kirti M., **137, 283, 621, 691**  
Yuan, Yanhui, **173**  
Yuan, Zhihong, **997**  
Yuliu, Zhifei, **229**  
Zamarripa, Miguel A., **208, 387, 546, 829, 941, 950**  
Zavala, Victor M., **129, 674, 732, 744, 763, 814, 820, 975**  
Zhang, Biao, **1006**  
Zhang, Qi, **997**  
Zhao, Yufei, **612**  
Zhou, Nana, **1006**  
Zhou, Panzheng, **814**  
Zhu, Lingxiang, **860**  
Zibunas, Christian, **585**  
Zitney, Stephen E., **448, 504**

---

# Index

## Keyword Index

---

# Keyword Index

- absolute zero, **2**
- absorption, **145**
- acceleration methods, **215**
- adsorption, **660**
- algorithms, **99, 105, 145, 201**
- ammonia, **993**
  - production, **316**
- anaerobic digestion, **750**
- antibody extraction, **283**
- artificial intelligence, **16, 22, 84, 245, 554**
- artificial neural network, **367**
- Aspen Plus, **409, 641, 705**
- atmospheric water extraction awe, **173**
- batch crystallization, **308**
- batch process, **275**
- battery energy storage, **417**
- Bayesian optimization, **373**
- biofuels, **660, 738, 836, 927**
- biomass, **738, 757, 836, 868**
- biomethane, **750**
- biopharmaceutical manufacturing, **67**
- biosolids, **726**
- biosystems, **75, 344, 442**
- blockchain, **844**
- blue hydrogen and nitrogen production, **771**
- calcium citrate, **121**
- CAMPD, **292**
- capacity expansion, **807**
- carbon accounting, **461**
- carbon dioxide, **44, 533, 578, 668, 783, 844**
  - capture, **32, 145, 160, 208, 222, 245, 546, 554, 585, 641, 681, 719, 836, 844, 860, 950, 959, 994, 999**
  - sequestration, **32, 738, 836, 844**
  - utilization and storage, **844**
- carbon reduction policies, **844**
- carbon tax, **844**
- cement, **533**
- CFD, **860, 883**
- chemical engineering, **92**
- chemical heat pump, **484**
- chemical process
  - design, **113**
  - optimization, **113**
- chemical process (continued)
  - (see also process design)
- chemical recycling, **792**
- chemical-looping reforming, **771**
- circular economy, **629, 792**
- circularity, **674, 792**
- climate change, **403**
- clinker, **533**
- computational fluid dynamics, **860, 883**
- computer vision, **84**
- computer-aided process design, **591**
- consumer products, **991**
- COSMO-RS, **763**
- cracking, **883**
- critical minerals, **853**
- curriculum, **975**
- cybersecurity, **186**
- cyclic adsorption processes, **215**
- data mining, **84**
- data-driven, **105**
  - analysis, **16**
  - model, -modelling, **351, 442**
  - optimization, **705**
- decarbonization, **533, 612, 883, 994, 1003**
- decomposition, **477, 604**
- defects, **380**
- derivative-free optimization, **237, 261, 344, 417, 719**
- desalination, **705**
- design, **92, 991**
  - sustainability, **807**
  - under uncertainty, **99, 201, 229, 519, 799, 927, 997**
  - (see also process design)
- design space, **194**
- diafiltration cascade, **853**
- digital twins, **67, 308**
- digitalization, **844**
- direct air capture, **403**
- direct compression, **194**
- discrete element method, **380**
- distillation, **129, 179, 300, 527, 660**
- distributed generation, **417**
- dynamic degradation modelling, **448**
- dynamic modelling, **75, 186, 275, 283, 778**
- ecosystem services, **10**
- EDEM, **380**
- education, **92, 959, 966, 975, 983, 998, 1009**
- electric vehicle fast charging, **417**
- electricity & electrical devices, **426, 455, 511**
- electricity markets, **490**
- empty fruit bunch, **681**
- energy, **32, 44, 578, 726, 783, 799, 927, 994**
  - and environment, **504**
  - conservation, **330**
  - conversion, **455, 876**
  - efficiency, **484, 527, 562, 994**
  - management, **469**
  - refueling, **541**
  - storage, **52**
  - systems, **208, 455, 511, 519, 554, 570, 585, 598, 698, 738, 993, 1006**
  - transition, **461, 994**
- entropy analysis, **44, 578**
- environment, **10, 39, 442, 541, 585, 799, 836, 868, 899, 927**
- equality constraints, **330**
- equivalent system mass, **674**
- ethylene, **883**
- exergy, **2**
  - efficiency, **484, 876**
- extraction, **300**
- fault detection and diagnosis, **322**
- flowsheet convergence, **129**
- flushing, **137**
- food & agricultural processes, **409, 799**
- forward problem, **330**
- fuel cells, **448**
- gasification, **681**
- generalized disjunctive programming, **853**
- global optimization, **719**
- graph attention network, **367**
- graph theory, **129, 975**
- green
  - ammonia, **807**
  - ammonia supply chain, **477**
  - fuels, **1003**
  - solvents, **763**
- grid refinement, **153**
- heat integration, **484, 912**

HIPSC cardiac differentiation, **344**  
 historical view, **16**  
 hybrid modelling, **351**  
 hydrogen, **395, 403, 442, 448, 469, 511, 533, 541, 993, 999**  
   -carriers, **591**  
   -production, **612**  
   -transport, **591**  
 identifiability, **778**  
 image-based control, **186**  
 implementation, **504**  
 industry 4.0, **75, 186, 261, 359**  
 instrumentation, **186**  
 integrated energy systems, **434, 490**  
 integrated temporal approach, **477**  
 integration of design and operation, **153**  
 intelligent systems, **308, 1006**  
 interdisciplinary, **10, 22, 732**  
 inverse problem, **330**  
 Jacobian, **950**  
 kinetic analysis, **173**  
 knowledge graph, **84**  
 life cycle assessment, -analysis, **10, 60, 585, 621, 660, 744, 757, 763, 783, 814, 820, 899, 904, 933**  
 linearization, **604**  
 liquid air energy storage, **562**  
 liquid extraction, **129**  
 lithium recovery, **853**  
 low temperature, **2**  
 machine learning, **22, 60, 75, 99, 113, 237, 245, 267, 275, 300, 316, 387, 434, 621, 711, 983**  
 mass conservation, **330**  
 material transition, **461**  
 membranes, **395, 711, 860**  
 metal-organic framework, **173**  
 meteorological analysis, **173**  
 methane reforming, **44, 578**  
 methanol, **668**  
 mixed-integer  
   -linear programming, **941**  
   -nonlinear programming, **477, 604, 853, 941**  
   -programming, **461**  
 minimizing carbon dioxide emissions, **44, 578**  
 model  
   -identification, **351**  
   -initialization, **395**  
   -predictive control, **373, 417**  
   -reduction, **283**  
   model-based design, **308**  
   modelling, **52, 60, 137, 215, 395, 585, 621, 691, 732, 820, 829, 868, 950, 983, 991** (see next)  
   modelling and simulations, **105, 121, 160, 261, 359, 409, 562, 641, 814, 904, 959, 966, 1009**  
   molecular design, **876**  
   monoethanolamine, **222**  
   multicomponent analysis, **705**  
   multifidelity, **351**  
   multilayer perceptron, **367**  
   multi-objective optimization, **920**  
   multiscale modelling, **22, 32, 292, 461, 511, 612, 698, 711**  
   multi-timescale decision-making, **477**  
   natural language processing, **84**  
   negative emission technologies, **403**  
   neon, **2**  
   network  
     -design, **469**  
     -optimization, **941**  
     -theory, **920**  
   net-zero, **10, 883**  
   neural networks, **316**  
   NMPC, **504**  
   noisy data, **330**  
   nonlinear  
     -model predictive control, **186**  
     -optimization, **113**  
     -programming, **153, 635**  
   nuclear, **511**  
   numerical methods, **983**  
   oil and gas, **892**  
   oil refineries, **1003**  
   optimal control problem, **771**  
   optimization, **22, 44, 99, 137, 145, 179, 194, 201, 208, 215, 222, 229, 237, 245, 253, 261, 283, 292, 359, 387, 426, 434, 442, 448, 469, 490, 527, 546, 562, 578, 585, 598, 635, 652, 668, 732, 738, 783, 820, 829, 844, 860, 868, 876, 892, 899, 904, 912, 927, 950, 966, 983, 993, 997**  
   (see also:  
     Bayesian optimiz.  
     chemical process optimiz.  
     data-driven optimiz.  
     derivative-free optimiz.  
     global optimiz.  
   optimization (continued)  
     multi-objective optimiz.  
     nonlinear optimiz.  
     process design optimiz.  
     parameter optimiz.  
     simulation-based optimiz.  
     quantum optimiz.  
     stochastic optimiz.  
     structural optimiz.  
     superstructure optimiz.  
     supply chain optimiz.)  
   parallelization, **237**  
   parameter optimization, **999**  
   PEM, **403**  
   pharmaceutical  
     -manufacturing, **67**  
     -process, **194**  
   pharmaceutics, **380**  
   phase equilibria, **179**  
   phenomena building block, **179**  
   planning, **892**  
   planning & scheduling, **698**  
   plastic(s), **652**  
     -recycling, **629, 763**  
     -waste, **229, 792**  
   polymers, polymerization, **229, 691, 732, 763, 814, 367**  
   post-combustion carbon capture, **635**  
   power grid, **455**  
   pretreatment, **253, 681**  
   process  
     -control, **983, 1006**  
     -design, **10, 52, 75, 137, 153, 160, 201, 208, 215, 229, 245, 292, 316, 344, 359, 387, 395, 409, 426, 490, 527, 546, 570, 621, 668, 674, 744, 757, 763, 814, 829, 860, 868, 876, 899, 904, 912, 941, 959, 993, 994, 998, 999, 1003, 1005, 1006**  
     -design and optimization, **496**  
     -electrification, **570, 999**  
     -integration, **121**  
     -intensification, **121, 160, 222, 300, 316, 711, 959**  
     -material balance, **44**  
     -modelling, **67**  
     -monitoring, **322**  
     -operations, **338, 426, 490**  
     -optimization, **883**  
     -optimization & control, **504**  
     -simulation, **129, 719**  
       (see also simulation)  
     -synthesis, **10, 16, 44, 52, 145,**

process synthesis (cont.)  
   **179, 261, 387, 541, 578, 621, 668, 994, 999**  
 produced water management, **941**  
 product design, **998**  
 propane, **167**  
 pulp and paper sludge, **750**  
 pybound, **668**  
 Pyomo, **201, 267, 308, 546, 829, 904, 950, 966**  
 quantum optimization, **292**  
 rare earth elements, **496**  
 rate-based model, **635**  
 reaction, reactor **167**  
   -engineering, **167, 367**  
   -design, **612, 883**  
 recycling, **496, 652**  
 reinforcement learning, **245, 387**  
 relaxation, **604**  
 renewable and sustainable energy, **39, 52, 426, 455, 469, 598, 698, 711, 904, 1005**  
 renewable electricity, **912**  
 reverse osmosis, **253**  
 Riemannian manifold, **322**  
 round concave tablet, **380**  
 sand, **892**  
 sector-coupling, **585**  
 sensitivity analysis, **635**  
 separator, **892**  
 shift reactor, **771**  
 simulation, **186, 668, 726, 966**  
 simulation-based optimization, **245**  
 sloppiness, **778**  
 social equity, **10**  
 software design, **490**  
 solar energy, **562**  
 solid oral dosage forms, **380**  
 solid oxide  
   -cells, **448, 504**  
   -electrolysis cells, **504**  
   -fuel cells, **504**  
 space manufacturing, **674**  
 sparse regression, **351**  
 stability, **778**  
 statistical process control, **322**  
 steam methane reforming, **612**  
 Stirling engine, **562**  
 stochastic optimization, **283, 519, 807**  
 structural optimization, **999**  
 superstructure optimization, **853**  
 supply chain, **60, 75, 469, 554, 652, 691, 698, 711, 732, 783, 820, 844, 899**  
 supply chain optimization, **629, 807**  
 support vector machine, **322**  
 surrogate modelling, **16, 105, 113, 237, 253, 275, 434, 546, 719, 799**  
 sustainability, **75, 504, 629, 674, 899, 920, 1003**  
 sustainable design, **373**  
 sustainable energy, **591**  
 syngas, **641**  
 system identification, **99, 275, 778, 1006**  
 systems engineering, **92**  
 tablet coating, **380**  
 tablet press, **194**  
 target material balance, **44, 578**  
 technoeconomic analysis, **160, 167, 229, 253, 267, 359, 409, 442, 541, 554, 641, 660, 744, 750, 757, 757, 763, 814, 820, 933** (see next)  
 technoeconomic assessment, **591**  
   (see previous)  
 thermo-mechanical exergy, **2**  
 time series aggregation, **604**  
 time series clustering, **434**  
 torrefaction, **681**  
 tributyl citrate, **121**  
 turbines, **598**  
 two-stage stochastic programming, **173**  
 uncertain systems, **373**  
 utilization, **726**  
 valorization, **750**  
 waste management, **652**  
 wastewater, **60, 267**  
 water, **253, 267, 799**  
 water networks, **829**  
 water-gas shift, **395**  
 wind, **598**  
 work analysis, **44, 578**  
 zero liquid discharge, **705**





[PSEcommunity.org](http://PSEcommunity.org)

PSE Press • Hamilton • Notre Dame