Optimal Process Design of Hydrogen Production Using Response Surface Methodology

Boeun Kim* Seongmin Heo** Kyung Hwan Ryu***

**Andlinger Center for Energy and the Environmental, Princeton University, Princeton, NJ 08544, USA
(e-mail: bk3460@princeton.edu).*

***Department of Chemical Engineering, Dankook University, Yongin 16890, Korea
(e-mail: smheo@dankook.ac.kr)*

**** Department of Chemical Engineering, Sunchon National University, Suncheon 57922, Korea
(e-mail: khryu@scnu.ac.kr)*

Abstract: Conceptual process design is a key step to the commercialization of novel technologies, where a network of reaction-separation-recycle processes is synthesized to produce target products from given raw materials. Such design is often realized using commercial process simulators (e.g., Aspen Plus, gPROMS) to calculate mass and energy balances, which can then be used as a basis for the evaluation through, for example, techno-economic analysis and life cycle assessment. In such evaluation, the reference design is very important as it forms the foundation for further analysis such as sensitivity analysis. However, it is typically determined heuristically, leading to a design with suboptimal (or even poor) evaluation metrics. To address this problem, in this work, model-free design of experiment methods are implemented to optimize process designs developed using Aspen Plus. Performances of different methods are compared using representative chemical process examples.

Keywords: Process design framework; Model-free optimization; Response surface method; Central composite design; Techno-economic analysis.

1. INTRODUCTION

As a paradigm shift to sustainable production, many researchers have focused on developing technologies associated with greenhouse gas emission reduction and hydrogen-based (or other renewable) energy systems. However, most new technologies are developed at a lab-scale without considering upstream and downstream processes. Thus, it is not easy to evaluate the actual potential of emerging technology for its commercialization. From this point of view, process simulation-based assessment methods, i.e., techno-economic analysis (TEA) and life cycle assessment (LCA) are widely used to assess both economic and environmental sustainability of the entire process with new technology, including its essential upstream/downstream processes. TEA aims at analyzing the economic performance of the target process. Depending on the process's technical maturity, it considers operating expenditure (OPEX), capital expenditure (CAPEX), or revenue as a performance index of the process. (Zimmermann 2020) Since economics is the most crucial factor from the perspective of industrial companies, many studies on newly developed technologies/processes have presented TEA results together to verify their economic viabilities. (Davis 2011, Heo 2019, Lee 2021, Swanson 2010) In a similar vein, the LCA has become an essential assessment as the interest in reducing greenhouse gases increases. (Artz 2018, Corti 2004, Feiz 2015, Lee 2021) While it is important to ensure that TEA and LCA are performed fairly and accurately, there is no globally certified procedure for them. In the case of LCA, the International Organization for Standardization (ISO) has announced the standard steps that can be employed to assess environmental performance. However, it does not provide an accurate evaluation procedure and greenhouse gas (GHG) emission parameters. Many researchers have proposed standardized procedures or guidelines for evaluating economic and environmental performances based on the steps published in the ISO standards. (Müller 2020, von der Assen 2014)

The evaluation results of both assessment methods are based on the process inventory that contains the information on the input and output streams, consumed energy, equipment size, etc. The process inventory can be established through the literature survey and/or process simulation. In general, it is difficult to gather all the process information at a commercial scale via actual experiments during the process development. For addressing this problem, chemical process simulation software (e.g., ASPEN PLUS, ASPEN HYSYS, and PROII) has been typically used for the conceptual process design and analysis in both industrial and academic areas. Such process simulators play an important role in the TEA and LCA studies as well as the process development and improvement. Their highly predictive capabilities come from their solid property database and process models. However, given the inherent complexity of chemical processes due to chemical reactions, recycling steams, and many constraints, an optimal process design through a commercial process simulator remains challenging. One of the systematic approaches is that the process is mathematically modeled, and its design problem is formulated as an optimization problem that can be solved by
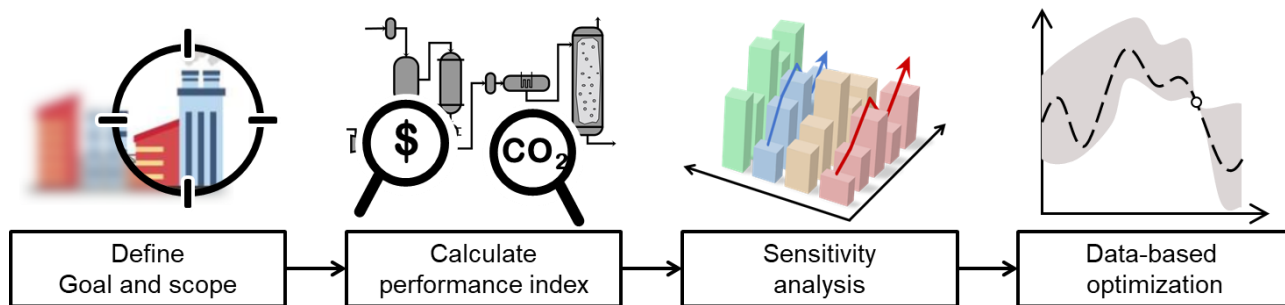
Figure 1. Flowchart of the proposed framework for optimal process design

the well-established optimization solvers (CONOPT, IPOPT, SNOPT, and BARON).

Unfortunately, detailed mathematical models and derivatives of objective function are not directly available from the chemical process simulators. Despite some simulators (ex. ASPEN PLUS) provide equation-oriented simulations, most of the optimization studies based on the process simulators have been carried out in a heuristic way because of the difficulty in the converge. In ASPEN PLUS, the IPOPT-based optimization method (in the absence of an algebraic model) is embedded, however, as simulation becomes more complex, the reliability of optimization result often degrades. (Cozad 2014)

This study proposes a model-free optimal process design framework by utilizing a process simulator. First, the goal and scope of process design are defined, and a suitable process model and performance index are designed. After that, influential design variables are selected through a sensitivity analysis, and they are optimized using the response surface methodology (RSM) with the design of experiment (DoE). It enables us to efficiently find the optimal design, avoiding the excessive runs of experiments. (Yolmeh 2017) Furthermore, all the type of user-defined objective function can be readily applied without going through additional modeling procedures. A conventional process of hydrogen production from natural gas is employed to introduce how the proposed framework can be adopted and to compare the performances of optimized designs and its nominal design obtained from previous work (Boyano 2011). We consider three different design objectives, i.e., maximizing productivity and minimizing minimum selling price and CO2 emission. Also, two experimental design methods, i.e., central composite design (CCD) and Latin hypercube design (LHD) are tested in the case study.

2. AUTOMATED OPTIMAL PROCESS DESIGN FRAMEWORK

This section introduces the proposed process design framework integrated with the design of experiment method. The model-free design of experiment method allows us to systematically decide the optimal process design with less effort and resources.

2.1 The proposed framework for optimal process design

Figure 1 shows the flowchart of the proposed optimal process design framework using a process simulator. In this framework,

we first define the goal and scope of our process design problem and develop the process model using a suitable commercial process simulator, which can meet the predetermined goal and scope. Suppose a conversion process is newly developed and just replaces the existing process without any modification. In that case, we may consider the conversion step only and compare their economic and/or environmental performances. Otherwise, we are required to extend the system boundary, including upstream/downstream processes.

Also, depending on the goal, we choose an appropriate performance index such as productivity, operating cost, and CO2 emission, generally calculated from the mass and energy balances of the developed process model. Next, the sensitivity analysis is performed by changing the values of process variables to identify the ones that have high impacts on the corresponding performance index. Most chemical processes have numerous design and operating variables that possibly influence the process's performance. However, optimization of all process variables is not efficient. As the number of decision variables increases, the computational burden increases exponentially, but not all of them significantly affect the performance index.

Finally, the data-based optimization is executed to decide the optimal process design, providing the best performance index. The selected process variables from the sensitivity analysis are only considered in this step. And the design of experiments (DoE) is adopted to design the most informative points (of virtual experiments in this study) to modeling and optimization
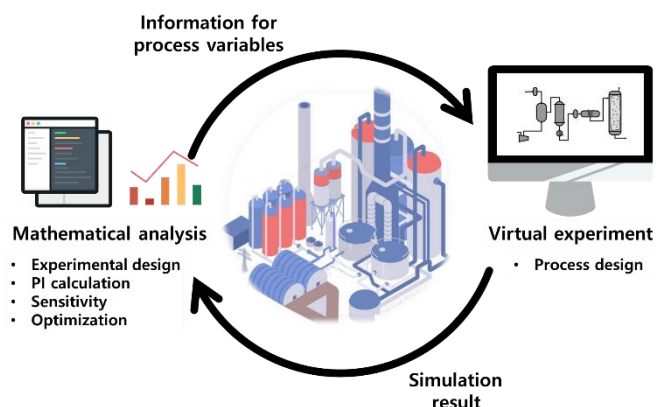

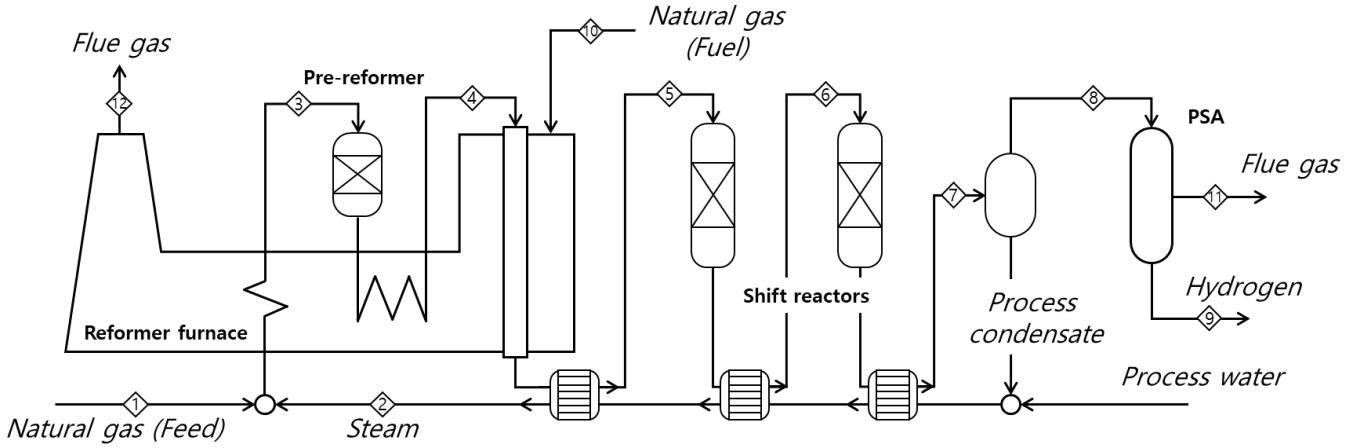Figure 2. Automated process design scheme connecting with experiment

Figure 4. The process flow diagram of hydrogen production process built in ASPEN PLUS

of the process and reduce the number of trial runs in the simulator. Details of the data-based optimization are described in the next subsections.

For the automation of the developed framework, ASPEN PLUS (virtual experiment) has interacted with MATLAB (mathematical analysis), as illustrated in Figure 2. ASPEN PLUS carries out the process simulations for given the designed experiments by MATLAB. MATLAB evaluates performance index calculation with the provided stream information, required energy, equipment size, etc., and performs the data-based optimization.

2.2 Design of experiments and response surface methodology

Design of experiments (DoE) is a widely used statistical method to measure the effects of one or more factors (input variables) on response or responses (output variables) with minimal effort. With the purpose of the process design/optimization, DoE is typically incorporated into response surface methodology (RSM) to express output variable(s) as a function of multiple input variables and find the optimal input variables giving the best output (Sanchez 2005). In general, a second-order polynomial form is used:

$$y = p_0 + \sum_{i=1}^{n} p_i x_i + \sum_{i=1}^{n}\sum_{j=i}^{n} p_{ij} x_i x_j + \sum_{i=1}^{n} p_{2,i} x_{i\,i}^2 \qquad (1)$$

where, y is the output variable, xi is the i-th input variables, n is the number of input variables, p's are the coefficients whose values are determined using the process simulation results, and $p0$ is the model intercept. Besides $p0$, three terms of right had side in Eq. 1 represent the input variables' main, interaction, and curvature effects on the output variable, respectively. If a certain term is not statistically significant, it would be removed from the regression model.

To gather the data for model fitting, we employ two different DoE methods, i.e., central composite design CCD and Latin hypercube design (LHD), and compare their performances in the later case study. One of the important tasks in the DoE and model fitting is the codification of the levels of input variables since they have different units and ranges (experimental

domain). The coding approach makes them over the same range without dimensions; for example, two levels (low and high) of each input variable are coded at -1 and +1. The regression model is developed for the coded input variables in the case study, but we will represent the optimized designs using their original values for convenience.

2.2.1 Central composite design (CCD)

A full factorial design (FFD), which investigates all possible combinations of the levels of the input variables (e.g., 2-level, 3-level), is simple and sufficient to give most of the important effects of the input variables. However, FFD requires a large number of experimental runs (likely to be impractical) when the number of input variables becomes larger, and their quadratic effects are modeled. To overcome this limitation of the FFD, the central composite design (CCD) adds a central point and star-points to the 2-level FFD. The star points are at a distance α from the center, providing sufficient information for the second-order regression model with the smaller number of runs compared to 3-level FFD. There are three types of CCD depending on the level of α (ex. $\alpha = (2^n)^{1/4}$ or 1) and rotatability: circumscribed, inscribed, and faced. Their
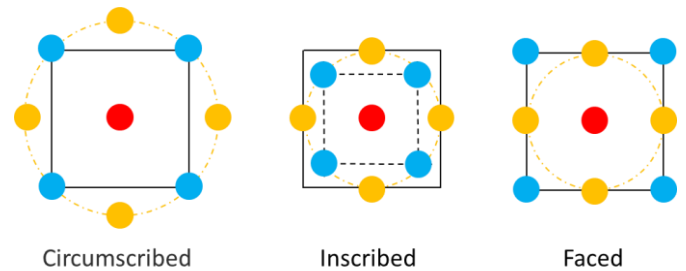

Figure 3. Three types of central composite design.

graphical representations (when n=2) are shown in Figure 3.

2.2.1 Latin hypercube design (LHD)

Latin hypercube design (LHD) has the space-filling properties of factorial designs using fine gratings and can provide effective design for quantitative factors. (Sanchez 2005) In LHD, the design space is divided into equally sized bins (their number equals the number of runs), and one point is randomly

selected for each bin. Since the LHD chooses data arbitrarily, the number of design points must be greater than the number of input variables. LHD shows good orthogonality when the number of design points is sufficient. However, since it is difficult to collect a large number of design points for every experiment, this approach is often to choose the best design among many randomly generated LHD.

3. Case study

The proposed framework was tested on an illustrative example of THE hydrogen production process. Figure 4 depicts the process flow diagram of the conventional hydrogen production from natural gas. (Boyano 2011) The system converts natural gas (CH_4 98.6%, C_2H_6 1.18%, C_3H_8 0.1%, n-C_4H_{10} 0.01%, i-C_4H_{10} 0.01%, N_2 0.1%, mol basis) into high purity hydrogen (99.99%) via conventional steam methane reforming process. The hydrogen production process model was built in ASPEN PLUS as a virtual experimental environment. In this case study, COM technology was used to connect MATLAB with ASPEN PLUS for executing the automated optimal process design. The proposed framework proceeds for the hydrogen production process in the steps explained in section 2.1.

As a first step, the goal of this process design was to improve the gate-to-gate conventional hydrogen production according to the given objective. In this process, high-pressure steam is optional produced or consumed depending on the operating conditions. The economic and environmental impact of HP steam is also considered in the system boundary. A furnace-type reformer is used to supply heat, and natural gas is also used as a fuel to determine the practical design.

In order to show the applicability of the proposed framework performance in TEA and/or LCA, we account for various performance indices, widely used economical and environmental sustainability indices: productivity, operating expenditure (OPEX), minimum selling price, and CO2 emission. In the present study, four different performance indices are used for evaluation—productivity, operating expenditure (OPEX), minimum selling price, and CO2 emission. Productivity is directly obtained from stream results. OPEX and minimum selling price can be estimated using techno-economic analysis, and CO2 emission can be estimated via life cycle assessment. Operating expenditure is simply calculated via stream results and raw material and utility cost for the operating process. Capital expenditure is estimated through inside battery limits (ISBL) and outside battery limits (OSBL). (Cetinkaya 2012) The cost of the freight on board (FOB) for each equipment is estimated from simulation results and published correlation equations (Woods 2007). The key economic parameters used in this study are summarized as follows:

Table 1. Economic parameters

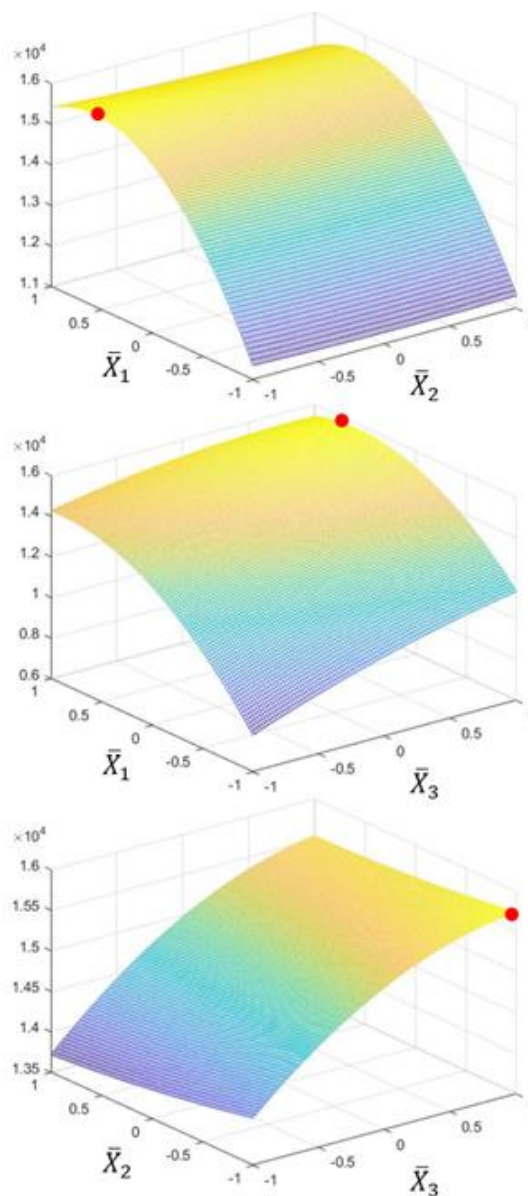| Price	
Name	Value
Process water	0.15 USD/ton
Natural gas	2.89 USD/MMBTU
Electricity	16.42 cent/kWh
Cooling water	0.02 USD/ton

Figure 5. Response surface plots of coded decision variables on hydrogen productivity

HP steam	6.08 USD/GJ

In order to analyze the environmental impact of the process, the life cycle assessment (LCA) method used in the previous study is taken into account (Lee 2021). Direct and indirect emission of the process is calculated via stream information and GWP values.

Table 2. Life cycle inventory parameters

Name	Value
Natural gas (NG)	0.15 kgCO$_{2-eq}$/kgNG
Electricity	108.9 kgCO$_{2-eq}$/kWh

Next, we screened the variables having large effects on the predefined performance index through the sensitivity analysis.

The considered hydrogen production process has six operating and design variables: reformer temperature (X1), number of tubular reactors (X2), high temperature water-gas shift (WGS) reactor temperature (X3), low temperature WGS reactor temperature (X4), carbon-to-hydrogen ratio (X5), and pre-reformer temperature (X6). Their nominal values are 950℃, 200, 400℃, 300℃, 3, and 500℃, respectively. Figure 5 shows the changes in each objective function values when the nominal value of each variable was perturbed by ± 10 %. As a result, the reformer temperature (X1), low temperature WGS reactor temperature (X4), and carbon-to-hydrogen ratio (X5) could influence all the performance indices, they were selected as the decision variables for the next optimal process design.

Finally, optimal decision variables were determined via data-based optimization. We employed three types of CCD and LHD to gather the simulation results for fitting the second-order polynomials in Eq. 1. 15 design sets for 3 decision variables (not 15 points for each variables) were generated via each DoE method. The upper and lower values of decision variables are 800~1100℃ (X1), 250~350℃ (X4), and 1.5~4.5 (X5), respectively. The response surface plots obtained from the fitted model with the DoE can provide insights into the effect of each decision variable on the performance index, and we can easily find the optimum denoted as red points in Figure 6. To compare the performances of the design methods fairly, LHD also generated the same number of design points for model fitting and optimization. However, the number of sampling points is not sufficient, and thus LHD cannot guarantee a good orthogonality property. 100 different cases of LHD were used to evaluate the average performance of data-based optimization with LHD.

The optimal designs obtained from the proposed framework are summarized in Tables 3 – 6. For all cases, the resulting optimal designs from the proposed framework showed better performance indices than their nominal values despite using very limited number of process simulation data. The best design could be found with LHD, however, performance indices obtained from other design methods are placed in the range of 95% confidence interval of those from case 1 (i.e., hydrogen productivity: 1.517~1.520×10^4, OPEX: 0.779~0.786, minimum selling price: 1.036~1.039, CO_2 emission: 12.85~12.94). Since there is no significant difference between the performances of the proposed framework with CCD and LHD, it is more important to determine which DoE method is suitable for our target process to design an optimal process with minimum effort.

Table 3. Optimal design for maximizing H_2 productivity, nominal result: 1.47×10^4 kgH$_2$/hr

	X1	X4	X5	Optimized
Circumscribed	1052	282.4	4.5	1.529×10^4
Inscribed	1026	272.9	3.92	1.518×10^4
Faced	1035	250	4.5	1.531×10^4
LHD (best)	1039	250	4.5	1.531×10^4
LHD (mean)	-	-	-	1.518×10^4

Table 4. Optimal design for minimizing operating cost, nominal result: 0.91 USD/kgH$_2$

	X1	X4	X5	Optimized
Circumscribed	1096	250	1.5	0.775
Inscribed	1054	250	1.5	0.780
Faced	1100	281.4	1.5	0.776
LHD (best)	1100	250	1.5	0.775
LHD (mean)	-	-	-	0.783

Table 5. Optimal design for minimizing selling price, nominal result: 1.19 USD/kgH$_2$

	X1	X4	X5	Optimized
Circumscribed	1084	250	1.5	1.029
Inscribed	1046	250	1.5	1.033
Faced	1089	263.7	1.5	1.031
LHD (best)	1100	250	1.5	1.028
LHD (mean)	-	-	-	1.037

Table 6. Optimal design for minimizing CO_2 emission, nominal result: 13.2 kgCO$_2$/kgH$_2$

	X1	X4	X5	Optimized
Circumscribed	1073	250	1.5	12.87
Inscribed	995	250	1.5	12.75
Faced	1088	289	1.5	12.90
LHD (best)	950	250	1.5	12.73
LHD (mean)	-	-	-	12.90

4. CONCLUSIONS

In the present study, we utilize the concept of model-free design of experiments to identify optimal process designs with a minimal number of process experiments/simulations. Specifically, two techniques are used for the model-free design of experiments: response surface method with central composite design and response surface method with Latin hypercube design. The conventional hydrogen production process is used as an illustrative example. The proposed framework with CCD and LHD determines the optimal designs for four objectives. All the optimization results show better performance compared with the nominal case. For example, the optimal design shows 4.1% increased productivity, 14.8% decreased OPEX, 13.4% decreased selling price, and 2.3% decrease in CO2 emission, respectively. The proposed method can determine the optimal design despite using a minimal number of data. Further study will consider two possible extensions; enlarging objective function that can handle a multi-objective design problem and implementing Bayesian optimization (BO) that can use less data than the used DoE.

REFERENCES

Artz, J., Müller, T. E., Thenert, K., Kleinekorte, J., Meys, R., Sternberg, A., ... & Leitner, W. (2018). Sustainable conversion of carbon dioxide: an integrated review of catalysis and life cycle assessment. Chemical reviews, 118(2), 434-504.

Boyano, A., Blanco-Marigorta, A. M., Morosuk, T., & Tsatsaronis, G. (2011). Exergoenvironmental analysis of a steam methane reforming process for hydrogen production. Energy, 36(4), 2202-2214.

Cetinkaya, E., Dincer, I., & Naterer, G. F. (2012). Life cycle assessment of various hydrogen production methods. International journal of hydrogen energy, 37(3), 2071-2080.

Corti, A., & Lombardi, L. (2004). Biomass integrated gasification combined cycle with reduced CO_2 emissions: Performance analysis and life cycle assessment (LCA). Energy, 29(12-15), 2109-2124.

Cozad, A., Sahinidis, N. V., & Miller, D. C. (2014). Learning surrogate models for simulation-based optimization. AIChE Journal, 60(6), 2211-2227.

Davis, R., Aden, A., & Pienkos, P. T. (2011). Techno-economic analysis of autotrophic microalgae for fuel production. Applied Energy, 88(10), 3524-3531.

Feiz, R., Ammenberg, J., Baas, L., Eklund, M., Helgstrand, A., & Marshall, R. (2015). Improving the CO_2 performance of cement, part I: utilizing life-cycle assessment and key performance indicators to assess development within the cement industry. Journal of Cleaner Production, 98, 272-281.

Heo, H. Y., Heo, S., & Lee, J. H. (2019). Comparative techno-economic analysis of transesterification technologies for microalgal biodiesel production. Industrial & Engineering Chemistry Research, 58(40), 18772-18779.

Lee, J., Lee, W., Ryu, K. H., Park, J., Lee, H., Lee, J. H., & Park, K. T. (2021). Catholyte-free electroreduction of CO_2 for sustainable production of CO: concept, process development, techno-economic analysis, and CO_2 reduction assessment. Green Chemistry, 23(6), 2397-2410.

Müller, L. J., Kätelhön, A., Bachmann, M., Zimmermann, A., Sternberg, A., & Bardow, A. (2020). A guideline for life cycle assessment of carbon capture and utilization. Frontiers in Energy Research, 8, 15.

Sanchez, S. M. (2005, December). Work smarter, not harder: guidelines for designing simulation experiments. In Proceedings of the Winter Simulation Conference, 2005. (pp. 14-pp). IEEE.

Swanson, R. M., Platon, A., Satrio, J. A., & Brown, R. C. (2010). Techno-economic analysis of biomass-to-liquids production based on gasification. Fuel, 89, S11-S19.

von der Assen, N., Voll, P., Peters, M., & Bardow, A. (2014). Life cycle assessment of CO_2 capture and utilization: a tutorial review. Chemical Society Reviews, 43(23), 7982-7994.

Woods, D. R. (2007). Rules of thumb in engineering practice. John Wiley & Sons.

Yolmeh, M., & Jafari, S. M. (2017). Applications of response surface methodology in the food industry processes. Food and Bioprocess Technology, 10(3), 413-433.

Zimmermann, A., Müller, L., Wang, Y., Langhorst, T., Wunderlich, J., Marxen, A., ... & Sick, V. (2020). Techno-Economic Assessment & Life Cycle Assessment Guidelines for CO_2 Utilization (Version 1.1).