Proceedings of the

17th Nordic Process Control Workshop

Technical University of Denmark

Kgs Lyngby, Denmark

January 25-27, 2012

25th January **Tutorial: Industrial Model Predictive Control**

26th - 27th January **17th Nordic Process Control Workshop**

Editors | **John Bagterp Jørgensen**
Jakob Kjøbsted Huusom
Gürkan Sin

Technical University of Denmark
DTU Informatics
DK-2800 Kgs Lyngby
Denmark

ISBN 978-87-643-0946-1

Printed in Denmark,
Technical University of Denmark, Lyngby 2012

17th Nordic Process Control Workshop

Welcome to the 17th Nordic Process Control Workshop (NPCW-17) and to the Technical University of Denmark. The NPCW-17 is held this year at the Technical University of Denmark (DTU) and jointly organized by DTU Informatics, Center for Energy Resource Engineering (CERE) and CAPEC at DTU Chemical Engineering.

The Nordic Process Control Workshop has been and continues to be an informal gathering that brings together academics and researchers from Nordic universities as well as professionals from industry in the process control area to present and discuss recent advances, share experiences and expand and consolidate networking in the process control field. The workshops are organized once every one and half year and the venue alternates between Denmark, Finland, Norway and Sweden.

The workshop is organized by the Nordic Working Group on Process Control currently consisting of the following members: Dr. John Bagterp Jørgensen, DTU, Denmark (Chair); Dr. Jeno Kovacs, Foster Wheeler Co., Finland (co-chair); Prof. Sirkka-Liisa Jamsa-Jounela, Helsinki Univ. Tech., Finland; Dr. Elling W. Jacobsen, KTH, Sweden; Prof. Sigurd Skogestad, NTNU, Norway; Dr. Kaj Juslin, VTT, Finland; Dr. Jan Peter Axelsson, Pfizer, Sweden; Prof. Claes Breitholtz, CTH, Sweden; Prof. Kurt Erik Häggblom, Åbo Akademi, Finland; Prof. Bjarne Foss, NTNU, Norway; Dr. Annika Leonard, Vattenfall, Sweden; Dr. Alf Isaksson, ABB, Sweden; Prof. Bernt Lie, Telemark Univ. College, Norway; Dr. Hans Aalto, Neste Jacobs, Finland; Dr. Bjørn Glemmestad, Telemark Univ. College, Norway; Prof. Tore Hägglund, Sweden; Dr. Gürkan Sin, DTU, Denmark; Dr. Tommy Mølbak, Dong Energy, Denmark.

Another important feature of the workshop is the Nordic Process Control Award appointed by the Nordic Working Group on Process Control. The Award is given to outstanding process control professionals, who have made "lasting and significant contributions to the field of process control". **The 2012 Nordic Process Control Award is given to Prof. Lorenz T Biegler from Department of Chemical Engineering, Carnegie Mellon University, USA.** We warmly congratulate Prof. Biegler and look forward to his award lecture entitled "A Nonlinear Programming Path to NMPC and Real-Time Optimization".

On behalf of the local organizing committee, we wish to thank many of our PhD and postdoctoral students for their help during the organization. In particular we thank Mrs. Anne Mette Eltzholtz Larsen for her excellent secretarial support, as well as Laura Standardi and Jason Anthony Price for their valuable help in compiling the preprints of the workshop.

Finally we would like to wish you a very productive and constructive workshop rich with scholarly discussions, increased networking and stimulating ideas for further research and cross collaboration across the board in the field of process control.

Lyngby, January 2012

John Bagterp Jørgensen
DTU Informatics & CERE

Jakob Kjøbsted Huusom
CAPEC, DTU Chemical Engineering

Gürkan Sin
CAPEC, DTU Chemical Engineering

17th Nordic Process Control Workshop

List of registrants

Name	Email	Institution	Country
Dr. AALTO, Hans	hans.aalto@nestejacobs.com	Neste Jacobs Oy	FINLAND
Mr. AALTONEN, Harri	harri.aaltonen@oulu.fi	University of Oulu	FINLAND
Mr. AGERBæK, Kasper	K.Agerbaek@Beckhoff.com	Beckhoff Automation	DENMARK
Mr. AIRIKKA, Pasi	pasi.airikka@metso.com	Metso Corporation	FINLAND
Mr. ALBAEK, Mads	maoa@novozymes.com	Novozymes / Technical University of Denmark	DENMARK
Dr. ALSOP, Nicholas	nicholas.alsop@borealisgroup.com	Borealis	SWEDEN
Mr. ANDERSEN, Torben Ravn	traan@statoil.com	Statoil Refining Denmark A/S	DENMARK
Mr. ANDERSSON, Jonas	jae@novozymes.com	Novozymes A/S	DENMARK
Dr. AXELSSON, Jan Peter	jan.peter.axelsson@vascaia.se	Vascaia AB	SWEDEN
Mr. BACKI, Christoph	backi@itk.ntnu.no	Department of Engineering Cybernetics	NORWAY
Prof. BIEGLER, Lorenz T.	lb01@andrew.cmu.edu	Carnegie Mellon University	UNITED STATES OF AMERICA
Mrs. BIEGLER, Lynne	xxx	Carnegie Mellon University	UNITED STATES OF AMERICA
Mr. BOIROUX, Dimitri	dibo@imm.dtu.dk	DTU Informatics	DENMARK
Mr. BORIOUCHKINE, Alexandre	alexandre.boriouchkine@aalto.fi	Aalto University	FINLAND
Prof. BREITHOLTZ, Claes	claesbr@chalmers.se	Chalmers	SWEDEN
Mr. CAPOLEI, Andrea	acap@imm.dtu.dk	DTU Informatics, CERE	DENMARK
Mr. DE OLIVEIRA, Vinicius	vinicius.deoliveira@ntnu.no	NTNU	NORWAY
Mr. DWIVEDI, Deeptanshu	dwivedi@nt.ntnu.no	Norwegian University of Science and Technology	NORWAY
Mr. ENGOULEVENT, Franck	franckengoulevent@gmail.com	IMM	DENMARK
Mr. GADE-NIELSEN, Nicolai Fog	nfga@imm.dtu.dk	Technical University of Denmark	DENMARK
Ms. GHADRDAN, Maryam	ghadrdan@nt.ntnu.no	NTNU	NORWAY
Prof. GLEMMESTAD, Bjørn	Bjorn.Glemmestad@hit.no	Telemark University College	NORWAY

Mr. GRIMHOLT, Chriss	chriss.grimholt@ntnu.no	NTNU	NORWAY
Mr. HALVGAARD, Rasmus	rhal@imm.dtu.dk	DTU	DENMARK
Dr. HALVORSEN, Ivar J.	ivar.j.halvorsen@sintef.no	SINTEF	NORWAY
Mr. HANSEN, Ole Fink	ofh@flsmidth.com	FLSmidth	DENMARK
Mr. HANSEN, Søren Skov	s072277@student.dtu.dk	DTU Informatics	DENMARK
Mr. HAST, Martin	martin.hast@control.lth.se	Dept. of Automatic Control, Lund Univeristy	SWEDEN
Mr. HAUGEN, Finn	finn.haugen@hit.no	Telemark University College	NORWAY
Mr. HEIRUNG, Tor Aksel N.	heirung@itk.ntnu.no	NTNU Norwegian University of Science and Technology	NORWAY
Prof. HOVD, Morten	morten.hovd@itk.ntnu.no	NTNU	NORWAY
Mr. HULTGREN, Matias	matias.hultgren@oulu.fi	University of Oulu	FINLAND
Dr. HUUSOM, Jakob Kjøbsted	jkh@kt.dtu.dk	DTU Chemical Engineering	DENMARK
Prof. HäGGBLOM, Kurt-Erik	khaggblo@abo.fi	Åbo Akademi	FINLAND
Prof. HäGGLUND, *Tore	tore.hagglund@control.lth.se	Department of Automatic Control, Lund University	SWEDEN
Mrs. IPSEN, Anne-Katrine	d01ai@statoil.com	Statoil Refinery	DENMARK
Dr. ISAKSSON, Alf	alf.isaksson@se.abb.com	ABB Corporate Research	SWEDEN
Mr. JACQUES, RICHALET	jacques.richalet@wanadoo.fr	consultant	FRANCE
Mr. JAHANSHAHI, Esmaeil	jahansha@nt.ntnu.no	Norwegian University of Science and Technology	NORWAY
Mr. JOHANSEN, Daniel Witt	s072093@student.dtu.dk	DTU Informatics	DENMARK
Mr. JOSEFSSON, Fredrik	fredrik.josefsson@borealisgroup.com	Borealis AB	SWEDEN
Prof. JäMSä-JOUNELA, Sirkka-Liisa	sirkka-l@tkk.fi	Aalto University	FINLAND
Dr. JäSCHKE, Johannes	jaschke@chemeng.ntnu.no	Department of Chemical Engineering, Norwegian University of Science and Technology (NTNU)	NORWAY
Prof. JøRGENSEN, John Bagterp	jbj@imm.dtu.dk	DTU Informatics	DENMARK
Prof. JøRGENSEN, Sten Bay	sbj@kt.dtu.dk	Dept. Chemical and Biocehmical EngineeeringDTU	DENMARK

Dr. KNUDSEN, Jørgen K. H.	JorgenKHKnudsen@gmail.com	2-control Aps.	DENMARK
Mr. KORTELA, Jukka	jukka.kortela@aalto.fi	Aalto University School of Chemical Technology	FINLAND
Dr. KOVáCS, Jenö	jeno.kovacs@oulu.fi	University of Oulu	FINLAND
Ms. KUULUVAINEN, Venla	venla.kuuluvainen@aalto.fi	Aalto University School of Chemical Technology	FINLAND
Dr. LARSSON, Per-Ola	perola@control.lth.se	Department of Automatic Control, Lund University	SWEDEN
Ms. LEONARD, Annika	annika.leonard@vattenfall.com	Vattenfall Research and Development AB	SWEDEN
Mr. LERCH, Dariusz	dal@kt.dtu.dk	IMM DTU	DENMARK
Mr. LIN, Choy-Hsien	choy-hsien.lin@storaenso.com	Stora Enso Hylte	SWEDEN
Mrs. LINDHOLM, Anna	anna.lindholm@control.lth.se	Department of Automatic Control, Lund University	SWEDEN
Ms. LOHINIVA, Laura	laura.lohiniva@oulu.fi	University of Oulu	FINLAND
Mr. MAGBOOL JAN, Nabil	nabil123@gmail.com	Norwegian University of Science and Technology	NORWAY
Mr. MAREE, Johannes P	phillip.maree@itk.ntnu.no	NTNU	NORWAY
Dr. MAURICIO IGLESIAS, Miguel	mim@kt.dtu.dk	DTU Kemiteknik	DENMARK
Mr. MAYER, Martin	mayer@innoversum.com	innoversum GmbH	AUSTRIA
Mr. MINASIDIS, Vladimiros	vladimiros.minasidis@chemeng.ntnu.no	NTNU	NORWAY
Mr. MIRZAEI, Mahmood	mhmd.mirzaei@gmail.com	Technical University of Denmark	DENMARK
Ms. MOLLERUP, Ane	molle@kt.dtu.dk	DTU, Chemical Department	DENMARK
Dr. MøLBAK, Tommy	tommo@dongenergy.dk	DONG Energy	DENMARK
Mr. NIELSEN, Morten Steen	s072115@student.dtu.dk	DTU Informatics	DENMARK
Mr. OLESEN, Daniel Haugård	s100094@student.dtu.dk	DTU Informatics	DENMARK
Mr. PETERSEN, Lars Norbert	lnpe@imm.dtu.dk	DTU Informatics	DENMARK
Mr. PETERSSON, Conny	conny.petersson@borealisgroup.com	Borealis AB	SWEDEN
Mr. POULSEN, Kristian Bøcher Poulsen	krbop@dongenergy.dk	DONG Energy	DENMARK
Mr. POULSEN, Niels Kjølstad	nkp@imm.dtu.dk	DTU Informatics	DENMARK

Mr. POZO GARCIA, Octavio	octavioado@gmail.com	Aalto University	FINLAND
Dr. PRADO-RUBIO, Oscar Andres	oap@kt.dtu.dk	DTU Chemical Engineering	DENMARK
Prof. PREISIG, Heinz A	heinz.preisig@chemeng.ntnu.no	NTNU	NORWAY
Mr. PRICE, Jason	japr@kt.dtu.dk	Technical University of Denmark	DENMARK
Mr. PRUNESCU, Remus Mihail	rempr@dongenergy.dk	Dong Energy and DTU	DENMARK
Dr. RECKE, Bodil	bre@flsmidth.com	FLSmidth A/S	DENMARK
Mrs. ROMERO SEGOVIA, Vanessa	vanessa.romero@control.lth.se	Department of Automatic Control, Lund University	SWEDEN
Mr. SHARMA, Roshan	roshan.sharma@hit.no	Telemark University College	NORWAY
Mr. SHIRDEL, Amirhossein	amir.shirdel@abo.fi	Åbo Akademi	FINLAND
Prof. SIN, Gurkan	gsi@kt.dtu.dk	DTU Chemical Engineering	DENMARK
Prof. SKOGESTAD, Sigurd	skoge@ntnu.no	NTNU	NORWAY
Mr. SMITH, Kevin	kevs@byg.dtu.dk	Byg @ DTU	DENMARK
Mr. SOKOLER, Leo Emil	emil.sokoler@gmail.com	DTU Informatics	DENMARK
Dr. SOLBERG, Brian Willum	brsol@dongenergy.dk	DONG Energy Power	DENMARK
Mr. SOLTESZ, Krisitian	kristian@control.lth.se	Lund University	SWEDEN
Ms. STANDARDI, Laura	laus@imm.dtu.dk	DTU Informatic	DENMARK
Dr. STEINKE, Florian	florian.steinke@siemens.com	Siemens AG	GERMANY
Dr. STOCKS, Stuart	STUS@novozymes.com	Novozymes A/S	DENMARK
Dr. STOICAN, Florin	florin.stoican@itk.ntnu.no	NTNU	NORWAY
Mr. THOMSEN, Søren Nymann	snt@volund.dk	B&W Vølund	DENMARK
Mr. TIKKALA, Vesa-Matti	vesa-matti.tikkala@aalto.fi	Aalto University	FINLAND
Ms. TOTU, Luminita Cristiana	lct@es.aau.dk	Aalborg University	DENMARK
Mr. UTZEN, Christer	christer.utzen@geagroup.com	GEA Niro	DENMARK
Mr. VALVERDE PéREZ, Borja	bvape@student.dtu.dk	CAPEC	DENMARK
Mr. VöLCKER, Carsten	cv@imm.dtu.dk	IMM-DTU	DENMARK
Mr. WAGNER, Falko Jens	falko.wagner@krones.dk	Krones	DENMARK

Dr. WIK, Torsten	tw@chalmers.se	Department of Signals and Systems	SWEDEN
Mr. YAZDI, Hassan	hy@flsmidth.com	FLSmidth	DENMARK
Mr. YELCHURU, Ramprasad	ramprasad.yelchuru@gmail.com	Norwegian University of Science & Technology	NORWAY
Dr. ZAKHAROV, Alexey	alexey.zakharov@aalto.fi	Aalto University	FINLAND
Mr. ZHOU, Guofeng	zhouguo.fer@gmail.com	Technical University of Denmark	DENMARK
Mr. ÖZDENKCI, Karhan	karhan.ozdenkci@aalto.fi	Aalto University	FINLAND

DTU

Legend

- Departments
- Oticon Hall
- Glass Hall
- Administration
- Campus Service
- Residential halls and guest houses
- SCION DTU
- Instructional buildings
- Bus stop
- Canteen
- DTU Library
- DTU Meeting Center

N

NPCW meeting room
Building 308
Auditorium 12

Lunch area
Canteen
Building 101

How to get to DTU

From Kastrup Airport directly to DTU (through *Nørreport station*):

Step A- from the airport to Nørreport station:

- Take the Metro (M2) from the airport in the direction of Vanløse station.
- Get off at Nørreport station.

or

- Take the bus 5A from the airport in the direction of Husum Torv.
- Get off at Nørreport station.

or

- Take the Øresund train in the direction of Helsingør station (Elsinore st.) or Nivå station.
- Get off at Nørreport station.

Step B- from Nørreport station to Lyngby station (and Rævehøjvej, DTU w/Bus 150S or 173E):

- From Nørreport station there are two S-trains and three buses to choose between:
  - S-train B in the direction of Holte station
  - S-train E in the direction of Hillerød station.
- Get off at Lyngby station.

or

- Bus 184 in the direction of Lyngby station.
- Get off at Lyngby station.

or

- Bus 150S in the direction of Kokkedal station.
- Get off at Rævehøjvej, DTU.

or

- Bus 173E in the direction of Fredensborg station.
- Get off at Rævehøjvej, DTU.

Step C - from Lyngby station to DTU:

- Bus 190 in the direction of Holte station.
- Get off at your destination at DTU Campus.

or

- Bus 300S in the direction of Nærum station.
- Get off at your destination at DTU Campus.

or

- Bus 353 in the direction of Helsingør station (Elsinore st.).
- Get off at your destination at DTU Campus.

or

- Bus 591P in the direction of Lyngby station.
- Get off at your destination at DTU Campus.

or

- Bus 590E in the direction of Lyngby station.
- Get off at your destination at DTU Campus.

Below are basic information about the mentioned busses and trains:

The Metro (M2)	The metro comes approximately every 3 minutes from 5 a.m. to 12 p.m. on friday - sunday nights it comes every 15 minute.
The Øresund Train	The Øresund train comes approximately every 20 minutes from 4 a.m. to 12 p.m. In the night it comes once an hour.
The S-trains	S-trains comes approximately every 6 minutes from 5 a.m. to 12 p.m. In the night it comes approximately every 30 minutes.
Bus 5A	Bus 5A comes approximately every 8 - 15 minutes from 4.55 a.m. to 12.10 a.m.
Bus 184	Bus 184 comes approximately every 20 minutes from approximately 5.30 a.m. to 12 p.m. Weekend and holidays not included.
Bus 190	Bus 190 comes every 30 minutes from approximately 6 a.m. to 6 p.m. Weekend and holidays not included.
Bus 300S	Bus 300S comes every 15 minutes from approximately 5 a.m. to 12 p.m. Weekend and holidays not included.

Bus 353	Bus 353 comes approximately every 30 minutes from 6.40 a.m. to 8 p.m. Weekend and holidays not included.
Bus 591P	Bus 591P comes approximately every 20 minutes from approximately 11 a.m. to 5.30 p.m. Weekend and holidays not included.
Bus 590E	Bus 590E comes approximately every 7 minutes from 7.32 a.m. to 9.45 a.m. Weekend and holidays not included.
Bus 150S	Bus 150S comes approximately every 10-15 minutes from 5.55 a.m. to 00.13 a.m. Weekend and holidays not included.
Bus 173E	Bus 173E comes approximately every 15 minutes from 6.43 a.m. to 8.47 a.m. Thereafter from 1.51 p.m. to 5.31 p.m. Weekend and holidays not included.

From central stations in Copenhagen to DTU

It is the same S-trains (E, B) that stops at the four main stations (Copenhagen Central station, Vesterport-, Nørreport- and Østerport station) and continues to DTU.

Here there is a guide to how to get to DTU from the four main stations:

First:

- S-train E in the direction of Hillerød station;
- Get off at Lyngby station.

or

- S-train B in the direction of Holte station;
- Get off at Lyngby station.

As you are at Lyngby st., you should take one of these busses to reach DTU:

- Bus 190 in the direction of Holte station;
- Bus 300S in the direction of Nærum station;
- Bus 353 in the direction of Helsingør station (Elsinore st.);
- Bus 591P in the direction of Lyngby station;
- Bus 590E in the direction of Lyngby station.

Good advice:

- DTU is a big campus therefore the mentioned busses stops at several places at campus. Make sure beforehand to know precisely at which bus stop to get off.

- Find out where to go by checking on a map of DTU Campus.
- When taking the bus, it is a good idea to ask the bus driver if it goes to DTU, and ask if the bus driver can announce in the speakerphone, when you have to get off.

Below are basic information about the mentioned busses and trains:

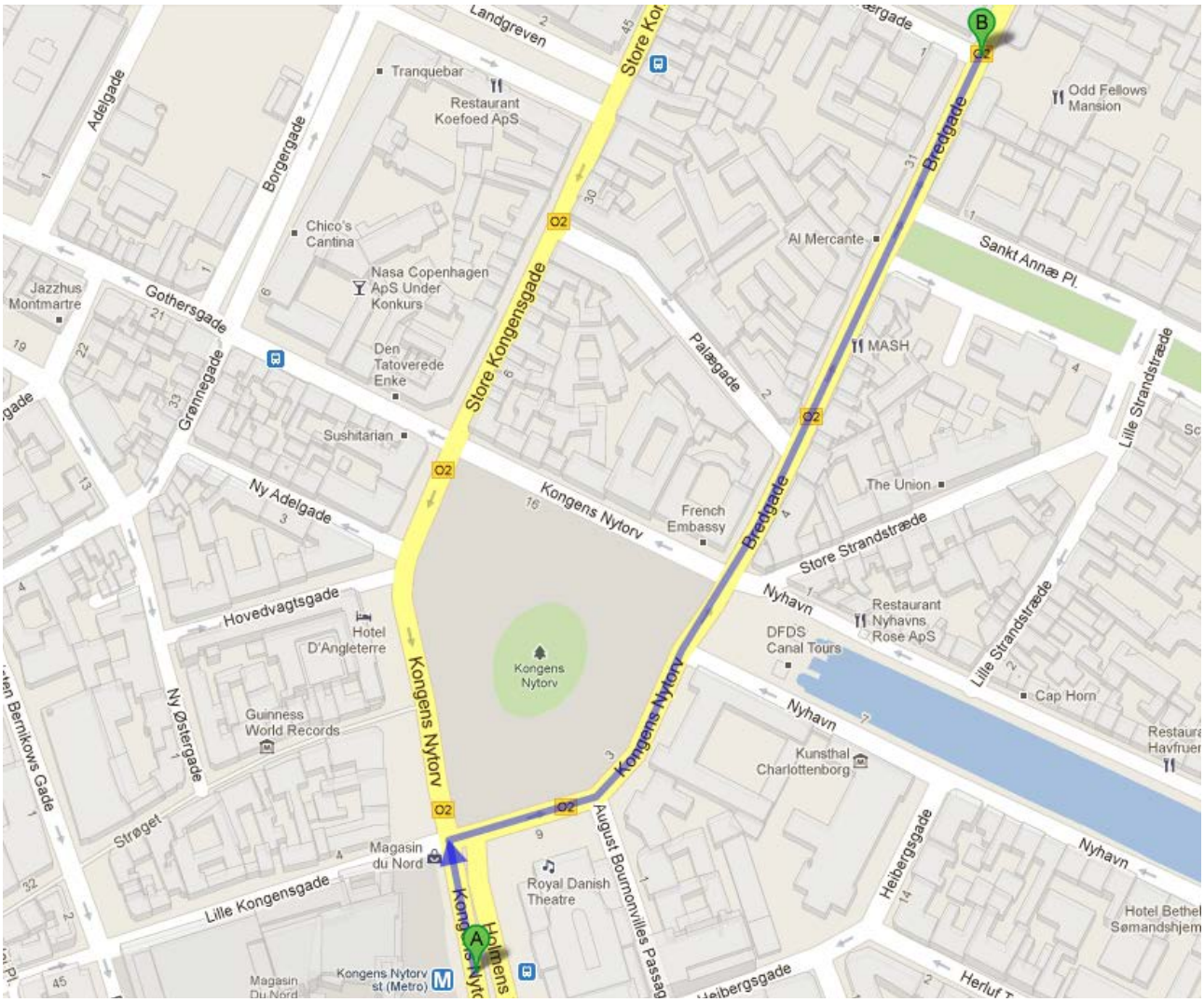
The S-trains	S-trains come approximately every 6 minutes from 5 a.m. to 12 p.m. In the night it comes approximately every 30 minutes.
Bus 184	Bus 184 comes approximately every 20 minutes from approximately 5.30 a.m. to 12 p.m. Weekend and holidays not included.
Bus 190	Bus 190 comes every 30 minutes from approximately 6 a.m. to 6 p.m. Weekend and holidays not included.
Bus 300S	Bus 300S comes every 15 minutes from approximately 5 a.m. to 12 p.m. Weekend and holidays not included.
Bus 353	Bus 353 comes approximately every 30 minutes from 6.40 a.m. to 8 p.m. Weekend and holidays not included.
Bus 591P	Bus 591P comes approximately every 20 minutes from approximately 11 a.m. to 5.30 p.m. Weekend and holidays not included.
Bus 590E	Bus 590E comes approximately every 7 minutes from 7.32 a.m. to 9.45 a.m. Weekend and holiday not included.

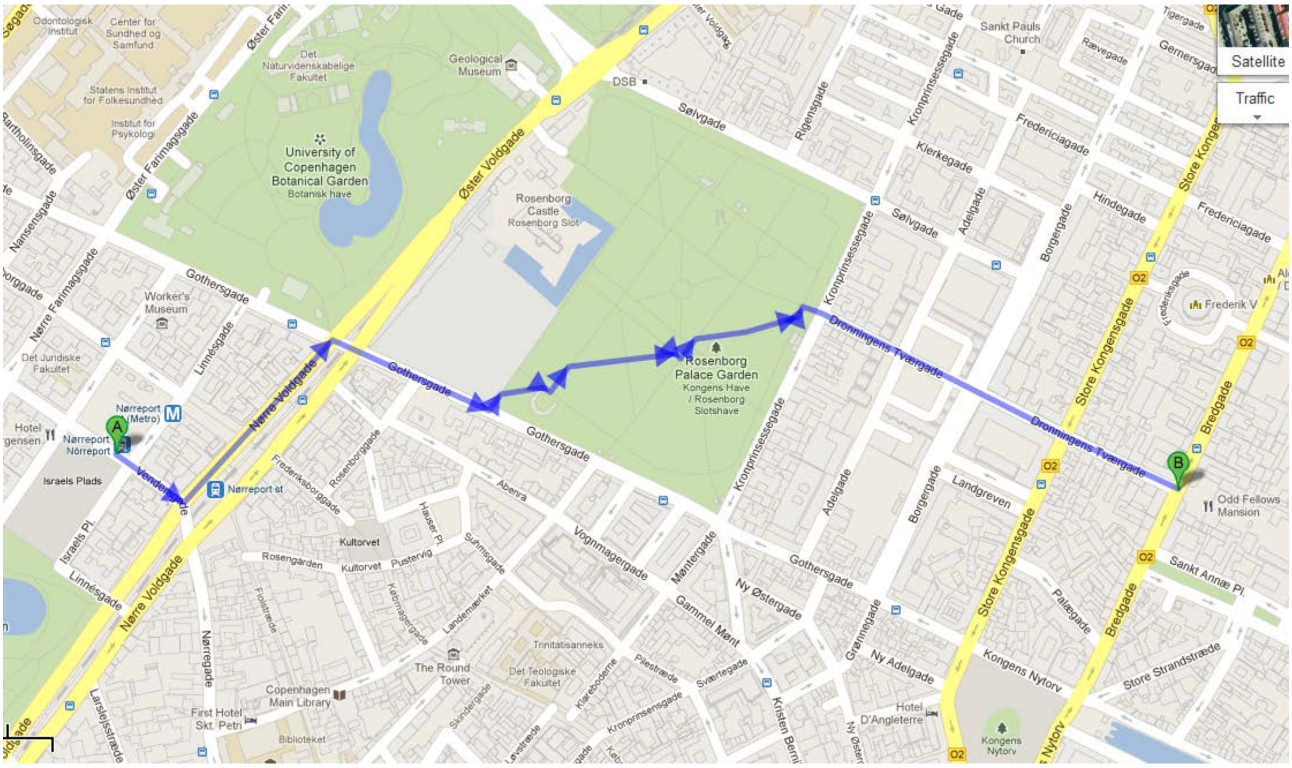
How to get to Old Fellow Palæet

More information are available on the web-site http://www.oddfellowpalaeet.dk/.

The restaurant is place in *Bredgade 28, DK-1260 København K*.

It's 1.3 km from Nørreport St. and 500 m. from the metro station Kongens Nytorv, which can be reached from Nørreport St. trough metro M1 to Vestamager or M2 to Lufthavnen (please see maps below).


B
A
Landgreven
Store Kongensgade
Bredgade
Tranquebar
Restaurant Koefoed ApS
Odd Fellows Mansion
Adelgade
Borgergade
Chico's Cantina
Al Mercante
Sankt Annæ Pl.
Jazzhus Montmartre
Gothersgade
Nasa Copenhagen ApS Under Konkurs
Palægade
MASH
Grønnegade
Den Tatoverede Enke
Lille Strandstræde
Sushitarian
Ny Adelgade
Kongens Nytorv
French Embassy
The Union
Store Strandstræde
Nyhavn
Hovedvagtsgade
Hotel D'Angleterre
Restaurant Nyhavns Rose ApS
DFDS Canal Tours
Kongens Nytorv
Cap Horn
Ny Østergade
Guinness World Records
Kunsthal Charlottenborg
Strøget
Magasin du Nord
Royal Danish Theatre
August Bournonvilles Passage
Heibergsgade
Lille Kongensgade
Hotel Bethel Sømandshjem
Kongens Nytorv st (Metro)
Herluf



University of Copenhagen Botanical Garden
Botanisk have
Geological Museum
Rosenborg Castle
Rosenborg Slot
Rosenborg Palace Garden
Kongens Have / Rosenborg Slotshave
Worker's Museum
Nørreport (Metro)
Nørreport st
Israels Plads
Kultorvet
The Round Tower
Copenhagen Main Library
Kongens Nytorv
Odd Fellows Mansion
Sankt Pauls Church
Gothersgade
Dronningens Tværgade
Nørre Voldgade
Øster Voldgade
Sølvgade
Kronprinsessegade
Store Kongensgade
Bredgade
Borgergade
Adelgade
Nørregade
Satellite
Traffic

Program

17th Nordic Process Control Workshop (NPCW-17)

Technical University of Denmark, Kgs Lyngby, Denmark

Thursday, January 26, 2012

08:30 Registration

Session 1: Award Session (Chair: John Bagterp Jørgensen)

09:00 Welcome and opening

09:10 Award ceremony

09:20 Award lecture - A Nonlinear Programming Path to NMPC and Real-Time Optimization

Lorenz T. Biegler. Carnegie Mellon University

Session 2: Identification and Monitoring (Chair: Sten Bay Jørgensen)

10:30 Fuel Moisture Soft-sensor and its Validation for the Industrial BioGrate Boiler
Jukka Kortela & Sirkka-Liisa Jämsä-Jounela. Aalto University 1

10:50 Temperature Modelling of the Biomass Pretreatment Process
Remus M. Prunescu, Mogens Blanke, Jakob M. Jensen & Gürkan Sin.
DONG Energy & DTU 8

11:10 Identification of Switching System
Amir H Shirdel & Hannu Toivonen. Åbo Akademi University 18

11:30 ModelID, an Interactive Program for Identification of MPC Relevant State-Space Models
Jørgen K.H. Knudsen. 2-control.....22

11:50 Data Mining for Process Identification
D. Peretzki, A. Isaksson, A.C. Bittencourt & K. Forsman.
Linköping University, ABB & Perstorp 29

Lunch

Session 3: Control Applications in Industry – Invited Session (Chair: Gürkan Sin)

13:10 Intelligent Control in the Future Energy System. Challenges, Solutions, and Business Potential
Tommy Mølbak. DONG Energy .....N/A

13:35 MPC at Statoil's Kalundborg Refinery
Anne-Katrine Ipsen. Statoil Refinery, Kalundborg.....30

14:00 Industrial challenges from the Cement and Mining Industry
Bodil Recke & Hasan Yazdi. FLSmidth 31

14:25 Application and Development of APC Schemes in Novozymes Fermentation Pilot Plant
Jonas Andersson & Stuart M. Stocks. Novozymes 32

14:50 Using computer models to save energy: An early warning model for tunnel pasteurizer energy consumption
Falko Jens Wagner. Sander Hansen Competence Center, Krones Nordic 33

Break

Session 4: Optimization (Chair: Alf Isaksson)

15:30 Optimal Controlled Variables for Parallel Process Units
Johannes Jäschke & Sigurd Skogestad. NTNU 38

15:50 On Performance for Tracking MPC
J.P. Maree, J. B. Rawlings & L. S. Imsland.
NTNU & University of Wisconsin-Madison.....39

16:10 Single Shooting and ESDIRK Methods for adjoint-based optimization of an oil reservoir
Andrea Capolei, Carsten Völcker, Jan Frydendall & John Bagterp Jørgensen. DTU
.....40

16:30 Optimization of lift gas allocation in a gas lifted oil field as non-linear optimization problem
Roshan Sharma, Kjetil Fjalestad & Bjørn Glemmestad. Telemark University College & Statoil 47

16:50 Model-Based Optimization of Economical Grade Changes for Borstar Polyethylene Plant.
Per-Ola Larsson, Johan Åkesson, Niclas Carlsson, Niklas Andersson. Lund University & Borealis 58

End of the scientific program, day #1

17:10-18:00 NPCW Steering Group Meeting, Building 305 Room 053

19:30 Dinner at restaurant Odd Fellow Palæet in downtown Copenhagen

Friday, January 27, 2012

08:30 Coffee & refreshments

Session 5: Perspectives of Control in Industry (Chair: Sigurd Skogestad)

09:00 Modeling Operating Modes during Plant Life Cycle
Sten Bay Jørgensen &Morten Lind. DTU 60

09:20 Teaching Control Principles to Industry Practitioners
Kristian Soltesz , Charlotta Johnsson & Tore Hägglund. Lund University 61

09:40 Industrial Application of Predictive Functional Control
Jacques Richalet.....68

Session 6: Poster session with coffee & refreshments (Chair: Jakob Kjøbsted Huusom)

10:00 Poster session / see the poster program below

Session 7: Control Theory (Chair: Sirkka-Liisa Jämsä-Jounela)

11:00 The SIMC method for smooth PID controller tuning
Sigurd Skogestad & Chriss Grimholt. NTNU. 71

11:20 Wannabe-MPC for Large Systems based on Multiple Iterative PI Controllers
Pasi Airikka & Mats Friman. Metso Corporation 100

11:40 Regulatory Layer Selection through Partial Control
Ramprasad Yelchuru & Sigurd Skogestad. NTNU 105

Lunch

Session 8: Monitoring and Fault Diagnosis (Chair: Jenö Kovacs)

13:00 Fault detection and diagnosis system for the drying section of a board machine
A. Zakharov & Sirkka-Liisa Jämsä- Jounela. Aalto University.....106

13:20 MPC Techniques in Fault Tolerant Control Design
Florin Stoican, Sorin Olaru &Morten Hovd. NTNU 107

13:40 Event-driven Flow Control of Incompressible Fluid
Sebastian Roll & Heinz A. Preisig. NTNU 114

14:00 Process Monitoring of Three Tank System
Harri Aaltonen. University of Oulu 119

Break

Session 9: Control Applications (Chair: Tore Hägglund)

14:30 Model-based analysis of control performance in sewer systems
A. H. Mollerup, M. Mauricio-Iglesias, NB. Johansen, D. Thornberg, PS. Mikkelsen, G. Sin. Copenhagen Wastewater Innovation & DTU 123

14:50 Dynamic Simulation of Oxy Combustion in a Pilot Scale CFB Boiler
M. Hultgren, L. Lohiniva, J. Kovacs, J. Ritvanen & A. Tourunen. Univ of Oulu, Foster Wheeler, LUT & VTT 128
Experimental Testing of Oxy Combustion in a Pilot Scale CFB boiler
H. Mikkonen, L. Lohiniva,A. Tourunen, M. Jegoroff, M. Hultgren, J. Kovács. VTT, University of Oulu & Foster Wheeler Energia 129

15:20 Greenhouse Artificial Lighting Control
Torsten Wik, Anna-Maria Carstensen, Tessa Pocock. Chalmers University of Technology & Heliospectra 130

15:40 Control of Blood Glucose for People with Type 1 Diabetes: an in Vivo Study
D. Boiroux, S. Schmidt, L. Frøssing, K. Nørgaard, S.Madsbad, O. Skyggebjerg, A-K. Duun-Henriksen,NK. Poulsen, H. Madsen, JB. Jørgensen.
Hvidovre Hospital and DTU. 133

16:00 Closing Remarks and introducing NPCW-18 **(Chair: Jakob Kjøbsted Huusom)**

Thank you all folks & bon voyage home!

List of Poster Presentations – NPCW-17

A Dantzig-Wolfe Decomposition Algorithm for Linear Economic MPC of a Power Plant Portfolio.
Laura Standardi, Niels Kjølstad Poulsen & John Bagterp Jørgensen. DTU Informatics. Kristian Edlund DONG Energy 141

A model-based FDD method of interacting control loops and its application to a mixing tank process.
Karhan Özdenkçi, Alexey Zakharov & Sirkka-Liisa Jämsä-Jounela. Aalto University 142

A New Static Estimator for Estimation of Primary Variables from Combination of Secondary Measurements.
Maryam Ghadrdan & Sigurd Skogestad. NTNU. Ivar J. Halvorsen. SINTEF ICT 143

Design of Optimal Low-Order Feedforward Controllers for Disturbance Rejection.
Martin Hast & Tore Hägglund. Lund University 144

Educational toolkit for teaching FDD methods
Venla Kuuluvainen, Vesa-Matti Tikkala, Esa Kiiski, Sirkka-Liisa Jämsä- Jounela. Aalto University 145

Experimental study of anti-slug control on small-scale test rig.
Esmaeil Jahanshahi & Sigurd Skogestad. NTNU 146

Fault Analysis as a Tool for Fault Detection and Diagnosis Development in Large- Scale Systems.
Vesa-Matti Tikkala, Helena Laavi & Sirkka-Liisa Jämsä-Jounela. Aalto University 147

Application of Shape-based Stiction Detection Methods to the Critical Valves of a Board Machine.
Octavio Pozo Garcia & Sirkka-Liisa Jämsä-Jounela. Alto University 153

Heuristic Evolutionary Random Optimizer for Non-linear Optimization Including Control and Identification Problems.
Pasi Airikka.Metso Corporation 154

Industrial validation of the dynamic BioGrate boiler model.
Alexandre Boriouchkine & Sirkka-Liisa Jämsä-Jounela. Aalto University. 160

Iterative Methods for MPC on Graphical Processing Units.
Nicolai Fog Gade-Nielsen, John Bagterp Jørgensen & Bernd Damman. DTU Informatics 161

Minimum Backed-off Operating Point Selection.
M. Nabil & Sridharakumar Narasimhan. IIT Madras. Sigurd Skogestad, NTNU. 162

Model Predictive Control of a Nonlinear Systems with Known Scheduling Variable.
Mahmood Mirzaei & Niels Kjølstad Poulsen. DTU Informatics. Hans Henrik Niemann. DTU Electrical Engineering 163

Modelica library for simulation of bioprocesses.
Jan Peter Axelsson. Vascaia AB 169

Modeling and simulating an electrical grid subsystem for power balance analysis
Luminita C. Totu, John J. Leth & Rafael Wisniewski. Aalborg University 170

Modeling of the freezing process for fish in vertical plate freezers.
Christoph Backi & Jan Tommy Gravdahl. NTNU..... 174

Modeling Smart Energy Systems for Model Predictive Control.
Rasmus Halvgaard, Niels Kjølstad Poulsen, Henrik Madsen & John Bagterp Jørgensen. DTU Informatics 175

Modelling and control design for SHARON/Anammox reactor sequence.
B. Valverde-Pérez, Miguel Mauricio-Iglesias & Gürkan Sin. DTU Chemical Engineering 176

Modelling Fungal Fermentations For Enzyme Production.
Mads O. Albaek & Krist V. Gernaey. DTU Chemical Engineering. Morten S. Hansen & Stuart M. Stocks. Novozymes A/S 185

Operation and Control of Enzymatic Biodiesel Production.
Jason A. Price, Jakob Kjøbsted Huusom, Mathias Nordblad & John Woodley. DTU Chemical Engineering 186

Optimal Control of a Batch Reactor Using the Linearized Hamilton-Jacobi-Bellman Equation.
Per Rutquist. Tomlab Optimization AB. Torsten Wik & Claes Breitholtz. Chalmers University of Technology 187

Optimal Input Design for Parameter Identification in Dynamic Systems Using Nonlinear Programming.
Tor Aksel N. Heirung & Bjarne Foss. NTNU. B. Erik Ydstie. Carnegie Mellon University ... 189

Optimisation of Oil Production in Two – Phase Flow Reservoir Using Simultaneous Method and Interior Point Optimiser.
Dariusz Lechr, Andrea Capolei, Carsten Völcker & John Bagterp Jørgensen. DTU Informatics
Erling Halfdan Stenby. DTU Chemistry 191

Data Analysis and Monitoring of Thickness Sensor Fouling Using Self-Organizing Maps.
Vesa-Matti Tikkala & Sirkka-Liisa Jämsä-Jounela. Aalto University 192

Production Optimization for Two-Phase Flow in an Oil Reservoir.
Carsten Völcker, John Bagterp Jørgensen & Per Grove Thomsen. DTU Informatics..... 198

Reducing Revenue Loss due to Utility Disturbances using Buffer Tanks - A Case Study at Perstorp.
Anna Lindholm, Charlotta Johnsson & Tore Hägglund. Lund University. Hampus Carlsson. Perstorp AB 199

Regulatory Control of 4-product Kaibel column.
Deeptanshu Dwivedi. Applied Cypernetics. Ivar J. Halvorsen & Sigurd Skogestad. NTNU ...201

Robust implementation of optimal control policies for transient processes.
Vinicius de Oliveira, Johannes Jäschke & Sigurd Skogestad. NTNU 202

Simulation, Control and Optimization of Single Cell Protein Production in a U-Loop Reactor.
Franck Guillaume Engoulevent & John Bagterp Jørgensen. DTU Informatics 203

State Estimation for the Automotive SCR Process.
Guofeng Zhou & Jakob Kjøbsted Huusom. DTU Chemical Engineering. John Bagterp Jørgensen. DTU Informatics. Christophe Duwig, Haldor Topsøe A/S 204

State Estimators for a Pilot Anaerobic Digestion Reactor.
Finn Haugen, Rune Bakke & Bernt Lie. Telemark University College 205

Stochastic Model Predictive Control with Applications in Smart Energy Systems.
Leo Emil Sokoler, Niels Kjølstad Poulsen, Henrik Madsen & John Bagterp Jørgensen. DTU Informatics Kristian Edlund, Tommy Mølbak, DONG Energy.....206

On controllability of an integrated bioreactor and periodically operated membrane separation process.
Prado-Rubio O.A. , Jørgensen, S.B. and Jonsson, G., DTU Department of Chemical and Biochemical Engineering……...207

Model Predictive Control for an Industrial SAG Mill.
Valeriu Ohan & John Bagterp Jørgensen DTU Informatics, Florian Steinke, Michael Metzger & Thomas Runkler, Siemens AG……...……208

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Fuel moisture soft-sensor and its validation for the industrial BioGrate boiler

Jukka Kortela * **Sirkka-Liisa Jämsä-Jounela** **

* *Aalto University School of Chemical Technology, PL 16100, FI-00076 Aalto (e-mail: jukka.kortela@aalto.fi).*
** *(e-mail:sirkka-l@tkk.fi)*

Abstract: This paper presents a soft-sensor for on-line monitoring of fuel moisture in a furnace. The method utilizes combustion power estimation and the dynamic model of the secondary superheater. In addition, the time delay for detecting a change in the moisture content of the fuel is small enough for the method to be used for controlling air and fuel feed preventing any steam and pressure oscillations. To verify the fuel moisture soft-sensor, experiments were performed at a BioPower 5 CHP plant, which utilizes BioGrate combustion technology for very wet biomass fuels with a moisture content as high as 65%. Finally, the results are analyzed and discussed.

Keywords: biomass, combustion, control, fuel moisture, industrial application

1. INTRODUCTION

Usage of biomass fuel for heat and power production is growing due to increasing demand for replacing fossil energy sources with renewable energy. The fuel is usually a blend of different batches, for example spruce bark and dry woodchips and with the moisture content between 30% and 55% (Yin et al. (2008)). The varying moisture content of the fuel results in uncertainty in the energy content of the fuel and complicates operation of the combustors. The typical procedure to determine moisture content of the fuel in small or medium-scale grate furnaces is to analyze manually collected samples of each fuel batch delivered to the plant. This method, however, is not accurate enough to predict moisture content of the fuel mix that enters the furnace. A change in moisture content of the fuel has to be detected at a resolution of seconds that the control system is able to make a correct response to the combustion air and the fuel feed system. There is, thus, a special need for a control system or for an operator to have information about moisture content for necessary adjustments of the combustors to be made.

On-line measurement of the fuel moisture content can be carried out with direct measurements or indirect methods. Direct measurements include dual X-ray, near infrared spectroscopy (NIR), radio frequent (RF), microwave and nuclear magnetic resonance (NMR) (Nyström and Dahlquist (2004)). Nordell and Vikterlöf (2000) performed series of experiments to verify if dual energy x-ray can be utilized as a technique to measure the moisture content in biofuels. They used standard medical X-ray equipment in the tests including ten different kind of wood fuels and grouped the results into three main groups depending on fuel type, wood-based, peat-based, willow and mixtures. Standard error of performance (SEP) 2% was achieved for individual fuel type. However, when the results were combined, correction coefficients were needed to take into account different carbon content of fuel. Therefore, more research in this area is needed for taking into account varying fuel compositions.

Several studies have been made to investigate NIR for different types of peat and wood. Ayalew and Ward (2000) tested NIR on three peat types (low-, medium-, and high-density peats) with moisture content ranging from ca. 33-63%. Based on the computed standard error of estimate, the 95% confidence limits of the moisture meter were 2.75, 1.75, 3.2, respectively, for low-, medium and high-density peats. However, when all peat types were grouped together, the 95% confidence limit increased to a higher error 5.3. Therefore, it is necessary to be able to distinguish between peat types. In another study, Axrup et al. (2000) evaluated NIR measurements with a silicone array spectrometer for the on-line analysis of wood chips and bark. They reported promising results with the root mean square error of prediction (RMSEP) of 2.3% for moisture content of wood chips and 5.1% for moisture content of bark. To conclude, NIR is the most promising method to be used in continous fuel flow, it is very well tested and adjustments to the problems with wood chips have been done. The disturbances such as distance from a sample and varying fuel composition in NIR are compensated for by using multivariate calibration and PLS. However, if some parameters change, for example fuel composition, a new calibration for NIR must be done.

James et al. (1985) tested microwave beam for measuring moisture content. They showed that the attenuation predominantly reflects moisture content, phase change reflects both moisture and density, and details of the depolarization indicate grain angle. The results showed that the microwave system is potentially capable of providing information about moisture. Okamura and Zhang (2000) found out however that the moisture content can be determined without any measurements of weight, thickness and microwave attenuation of the sample by using phase shifts in two microwave frequencies. These findings are in

agreement with James et al. (1985). The standard error of this method is 2.1% in the moisture range from 2 to 30%. Nevertheless, the method is influenced to a minor degree by temperature and grain direction.

Rosenberg et al. (2001) carried out a number of measurements to test NMR for measuring moisture content of spruce and pine. The difference between the NMR instrument and oven-dry measurements varied from -0.7% to +0.4% with an average of -0.13%. The results can be reported to be promising. However, the instrument has to be combined with density measurement, as it measures numbers of detected thermalized neutrons (counts/second).

A typical indirect method for determining fuel moisture content is to determine first moisture content of the flue gas from which the moisture content can be then derived by a mass balance calculation (Kortela and Marttinen (1985)). The only delay of the measurement signal in this setup is the transport time of the gas from the furnace to the measurement position. This time delay can be measured in seconds and it opens up thus possibilities of controlling both combustion air and fuel feed.

The Fourier-transform infrared (FT-IR) technology is one indirect method of determining the gas moisture content (Bak and Clausen (2002)). Jaakkola et al. (1998) investigated the feasibility of a transportable, low resolution FT-IR-gas analyzer for wet extractive stack gas analysis. They reported the relative standard deviation of 4.1% for moisture content. However, the accuracy of the FT-IR has been reported to be sensitive to the absolute temperature level, pressure, temperature gradients and particles carried with the gas, complicating measurements directly in the flue gas duct and this weakens its usability.

Another method for measuring the gas moisture content using a relative-humidity (RH) sensor was developed with the aim of improving the accuracy level of indirect determination of the moisture content of the fuel in a biomass furnace (Hermansson et al. (2011)). The accurate implementation was achieved by cooling of an extracted flue gas stream, elevating the RH of the flue gases, before performing the measurement. The results of the tests showed that the method is able to detect variations in moisture content in seconds. However, in order to use this method, new devices, measurements and calibration are needed.

All above methods have their strengths and weaknesses. There is thus a need for more cost efficient method for measuring fuel moisture, especially for a small scale boiler. This paper presents an indirect method for estimating the fuel moisture utilizing combustion power estimation and the dynamic model of the secondary superheater. The paper is organized as follows. In Section 2, the BioPower 5 CHP plant process is presented. Section 3 presents the fuel moisture soft-sensor. The test results are given in Section 4 and Section 5, followed by the conclusions in Section 6.

2. DESCRIPTION OF THE PROCESS

In the BioPower 5 CHP plant, the heat used for steam generation is obtained by burning solid biomass fuel: bark, sawdust and pellets, which are fed to the steam boiler together with combustion air. As a result combustion heat and flue gases are generated. The heat is then used in the steam-water circulation process.

Fig. 1 shows the boiler part of the BioPower 5 CHP plant. The essential components of the water-steam circuit are an economizer, a drum, an evaporator and superheaters. Feed water is pumped from a feed water tank to the boiler. First the water is led to the economizer (4) that is heated by flue gases. The temperature of flue gases is decreased by the economizer and the efficiency of the boiler is improved.

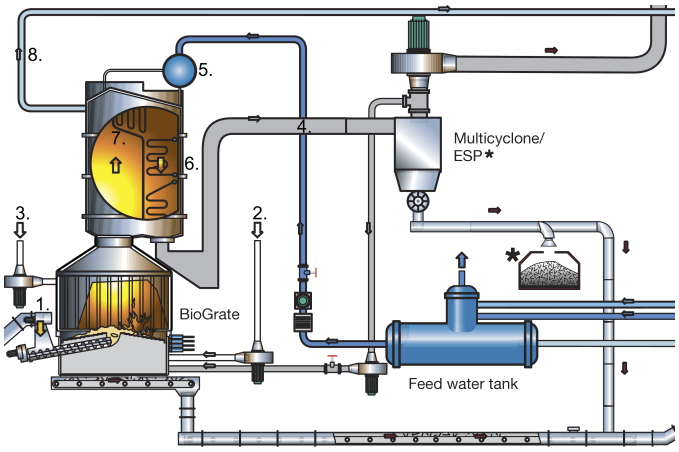


Fig. 1. 1. Fuel, 2. Primary air, 3. Secondary air, 4. Economizer, 5. Drum, 6. Evaporator, 7. Superheaters, 8. Superheated steam

From the economizer, heated feed water is led to the drum (5) and along downcomers into the bottom of the evaporator (6) tubes that surround the boiler. From the evaporator tubes the heated water and steam return back to the steam drum, where steam and water are separated. Steam rises to the top of the steam drum and flows to the superheaters (7). Steam heats up furthermore so it superheats. The superheated high-pressure steam (8) is led to a steam turbine, where electricity is generated.

In the BioGrate system, the fuel is fed onto the center of a grate from below by a stoker screw. The grate is divided into concentric rings with alternate rotating rings and the rings between remaining stationary. Alternate rotating rings are pushed hydraulically clockwise or counterclockwise respectively. This design distributes the fuel evenly over the entire grate with the burning fuel forming an even layer of the required thickness. Wärtsilä (2005)

The water content of the wet fuel in the centre of the grate evaporates rapidly due to the heat of the surrounding burning fuel and thermal radiation from the brick walls. The gasification and visible combustion of the gases and solid carbon take place as the fuel moves to the periphery of the circular grate. Finally, at the edge of the grate ash falls into a water-filled ash basin underneath the grate. Wärtsilä (2005)

The primary air for combustion, and the recirculation flue gas, are fed from underneath the grate and penetrate the fuel through slots in the concentric rings. Secondary air is fed above the grate directly into the flame. Air distribution is controlled by dampers and speed-controlled fans. Wärtsilä (2005)

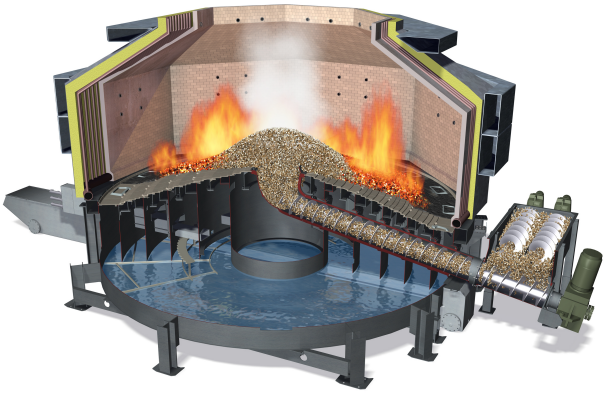


Fig. 2. BioGrate including the stoker screw and a water-filled ash basin underneath the grate

2.1 Fuel composition and fuel quality

The composition and the quality of fuel have big effect on its heat value. Thus fuel quality is playing a key role when designing a control strategy of a biopower plant and guaranteeing its optimal operation. Common elements to all biomass fuels are carbon (C), hydrogen (H), oxygen (O) and nitrogen (N). In addition, biomass fuels contain substances from soil, such as water, minerals, rock materials and sulphur (S). The actual combustible components of fuels are carbon, hydrogen and sulphur. Sulphur is an unwanted component, because it forms harmful sulphur dioxide, when it is burned. Part of nitrogen reacts with oxygen forming harmful nitrogen oxides. Water in fuel requires heat for its evaporation. Because of this, moisture decreases the heat value of fuel. Table 3 lists the elemental composition and typical moisture content of wood fuels burned in the BioPower 5 CHP plant.

Table 1. The composition of wood fuels burned in the Biopower 5 CHP plant

Fuel	Dry content (%)					Moisture (%)
	w_C	w_H	w_O	w_N	Ash	w
Pine	54.5	5.9	37.6	0.3	1.7	60
Spruce	50.6	5.9	40.2	0.5	2.8	60
Wood mix	50.4	6.2	42.5	0.5	0.4	50

The heat value of fuel can be determined by the equation that has been derived from heat values between combustible components and oxygen (Effenberger (2000)). The effective heat value of a dry fuel is

$$q_{wf} = 0.348 \cdot w_C + 0.938 \cdot w_H + 0.105 \cdot w_S + 0.063 \cdot w_N - 0.108 \cdot w_O [\text{MJ/kg}] \quad (1)$$

where w_C is mass fraction of carbon in fuel (%), w_H is mass fraction of hydrogen in fuel (%), w_S is mass fraction of sulphur in fuel (%), w_N is mass fraction of nitrogen in fuel (%), and w_O mass fraction of oxygen in fuel (%). The effective heat value of a wet fuel is obtained from equation

$$q_f = q_{wf} \cdot (1 - w/100) - 0.0244 \cdot w [\text{MJ/kg}] \quad (2)$$

where w is moisture content of the wet fuel (%). In order to use Equation (2), the composition of fuel has to be known.

3. FUEL MOISTURE SOFT-SENSOR

The fuel moisture soft-sensor is based on combustion power estimation and the dynamic model of the secondary superheater. Firstly, fuel flow $\hat{m}_f(w)$ is estimated in the combustion model as shown in Fig. 3. Secondly, the estimated fuel flow $\hat{m}_f(w)$ is used to calculate the flue gas flow m_{fg} and flue gas temperature T_{fg}. Thirdly, the flue gas flow $\hat{m}_{fg}$ and flue gas temperature $\hat{T}_{fg}$ are used together with input enthalpy h_1 and steam flow m_1 for predicting the output enthalpy $\hat{h}_2(i)$ of the secondary superheater model. Finally, the unknown fuel moisture disturbance w can be then estimated with least squares method.

$$\min J(w) = \sum_{i=0}^{N} |h_2(i) - \hat{h}_2(i)|^2 \quad (3)$$

where N is moving window, h_2 is measured output enthalpy of the secondary superheater (MJ/kg), and $\hat{h}_2$ estimated output enthalpy of the secondary superheater (MJ/kg).

The details of the models are presented in the following Sections 3.1 and 3.2.

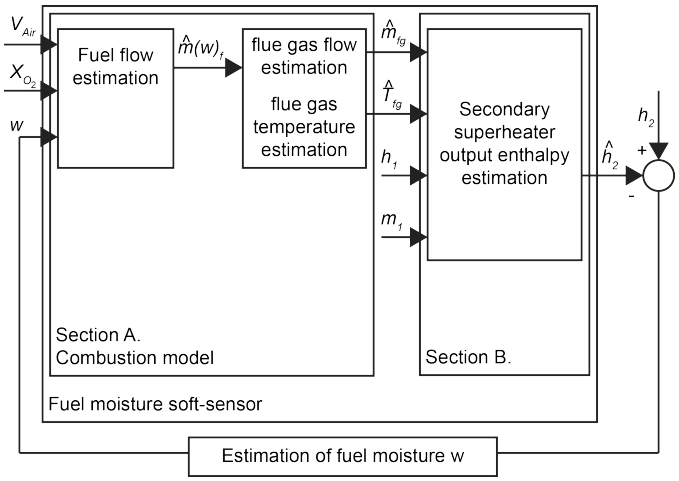


Fig. 3. The fuel moisture soft-sensor

3.1 Estimation of fuel flow, flue gas flow and temperature

The combustion reaction in biopower plants occurs mainly between carbon and oxygen. Therefore, the oxygen consumption and flue gas composition are good measures of heat generation in the furnace (Kortela and Lautala (1981)). The amount of oxygen needed for fuel combustion can be determined from the reaction equations. Table 2 gives moles per unit of fuel from mass fractions of the fuel. Summarizing, the oxygen needed for different components and subtracting the amount of oxygen in the fuel, the theoretical amount of oxygen needed to burn completely one kilogram of the fuel is

Table 2. Moles of the components of the fuel per unit mass

Comp.	Mass fraction (%)	M_i(g/mol)	n_i (mol/kg)
C	$w_c(1-w/100)$	12.011	$w_c(1-w/100)10/M_C$
H	$w_h(1-w/100)$	2.0158	$w_h(1-w/100)10/M_H$
S	$w_s(1-w/100)$	32.06	$w_s(1-w/100)10/M_S$
O	$w_o(1-w/100)$	31.9988	$w_o(1-w/100)10/M_O$
N	$w_n(1-w/100)$	28.01348	$w_n(1-w/100)10/M_N$
Water	w	18.0152	$10/M_W$

$$N^g_{O_2} = n_C + 0.5 \cdot n_{H_2} + n_S - n_{O_2}[\text{mol/kg}] \quad (4)$$

In addition to combustion products, nitrogen N that comes with the air, is included in flue gases. There is thus 3.76 times more nitrogen compared with needed oxygen in flue gas calculations. Incombustible components for example water are included in the equations as such. Therefore, flue gas flow for one kilogram of fuel is

$$N_{fg} = n_C + n_{H_2} + n_S + 3.76 \cdot N^g_{O_2} + n_{N_2} + n_{H_2O}[\text{mol/kg}] \quad (5)$$

The estimated amount of fuel burned is given as follows (Kortela and Jämsä-Jounela (2010))

$$\hat{m}_f(w) = \frac{(0.21 - \frac{X_{O_2}}{100})n_{Air}}{N^g_{O_2} + \frac{X_{O_2}}{100}(N_{fg} - 4.76 \cdot N^g_{O_2})}[\text{kg/s}] \quad (6)$$

where $X_{O_2}(t+\tau)$ is oxygen content of flue gas (%), and n_{Air} total air flow (mol/s). Therefore, flue gas flow is

$$\hat{m}_{fg} = \hat{m}_f(w) \cdot N_{fg}[\text{kg/s}] \quad (7)$$

and flue gas temperature is

$$\begin{aligned}\hat{T}_{fg} = (q_f &+ 0.21(F_{Air}/(22.41 \cdot 10^{-3} \cdot \hat{m}_f(w))C_{O_2} \\ &+ 0.79(F_{Air}/(22.41 \cdot 10^{-3} \cdot \hat{m}_f(w))C_{N_2})/ \\ &(n_C C_{CO_2} + n_S C_{SO_2} + (n_{H_2O} + n_{H_2})C_{H_2O} \\ &+ (3.76 \cdot N^g_{O_2} + n_{N_2})C_{N_2} + 0.21 \cdot N_{ExAir}C_{O_2} \\ &+ 0.79 \cdot N_{ExAir}C_{N_2})[^\circ\text{C}] \end{aligned} \quad (8)$$

where F_{Air} is total air flow (m^3/s), and C_i specific heat capacity i (J/molT).

$$\begin{aligned}N_{ExAir} = (&F_{Air}/(22.41 \cdot 10^{-3} \cdot m_f) \\ &- 4.76 \cdot N^g_{O_2}[\text{mol/kg}]\end{aligned} \quad (9)$$

3.2 Secondary superheater model

The behaviour of the boiler is captured by global mass and energy balances. The heat released by the combustion of fuel is transferred to the water and steam of the boiler where each section can be considered as a thermal system (Åström and Bell (2000)). In this paper fuel quality is estimated using the dynamic model of the secondary superheater. The Energy balance of the boiler section and temperature of metal walls are considered separately for improving the model accuracy.

Heat transfer from flue gas to metal walls for mixed convection and radiation heat transfer is (Lu (1999) and Lu and Hogg (2000))

$$Q_w = \alpha_w \hat{m}^{0.65}_{fg}(\hat{T}_{fg} - T_w) + k_w(\hat{T}^4_{fg} - T^4_w)[\text{MJ/s}] \quad (10)$$

where α_w is convection heat transfer, T_w is temperature of metal walls (°C), and k_w radiation heat transfer coefficient. Energy balance for tube walls is

$$\frac{dT_m}{dt} = \frac{1}{m_t C_p}(Q_w - Q_t)[\text{K/s}] \quad (11)$$

where m_t is mass of the metal tubes (kg), and C_p specific heat of the metal (MJ/kgK). Heat transfer from metal walls to steam/water for convection heat transfer (superheaters) is

$$Q_t = \alpha_c m^{0.8}_2(T_m - T)[\text{MJ/s}] \quad (12)$$

where α_c is convection heat transfer coefficient.

$$T = (T_1 + T_2)/2[^\circ\text{C}] \quad (13)$$

where T_1 is input steam/water temperature (°C), and T_2 output steam/water temperature (°C).

Finally the energy balance for output steam enthalpy is

$$\frac{d\hat{h}_2}{dt} = \frac{1}{\varrho V}(Q_t + m_1 h_1 - m_2 \hat{h}_2)[\text{MJ/(s} \cdot \text{kg)}] \quad (14)$$

where $\hat{h}_2$ is specific output enthalpy of steam/water (MJ/kg), ϱ is specific density of steam/water (kg/m^3), V is volume of steam/water (m^3), m_1 is input steam/water flow (kg/s, h_1 is specific input enthalpy of steam/water (MJ/kg), and m_2 output steam/water flow (kg/s).

4. DESCRIPTION OF THE TESTING ENVIRONMENT

All the experiments were conducted at the BioPower 5 CHP plant that produce 13.5 MW heat and 2.9 MW electricity. The plant utilizes BioGrate combustion technology as presented in Fig. 2. This technology has been developed for very wet biomass fuels with a moisture content as high as 65%.

Fig. 4. Measurement set-up for fuel feed sampling

Fig. 5. Measurement set-up of the FT-IR analyzer in the flue gas duct

Two fuels were used to test moisture soft-sensor: spruce bark with an average moisture content of 57% and a typical composition for bark (carbon 51% , hydrogen 6.2%, nitrogen $< 0.2\%$, sulphur $< 0.2\%$, and ash 0.5%) as shown in Table 3, and dry woodchips (spruce) with the moisture content of 20%.

Table 3. The fuel composition of the tests

Fuel composition	(%)
Carbon	51.0
Hydrogen	6.2
Nitrogen	<0.2
Sulphur	<0.2
Ash	0.5

To ensure the stepwise changes of the moisture content, the automatic feeding system was disabled and dry biomass was fed manually into the screw conveyor between wet fuel through the extra feeding box. The arrangement for sampling of fuel feed is shown in Fig. 4. The samples were taken from fuel dropping from the fuel silo just before the stoker screw every 5 minutes. The measurement arrangement for Servomex 2500 FT-IR analyzer is shown in Fig. 5. The flue gas was extracted from the flue gas duct and led to the analyzer. Samples were recorded every second.

Calculations presented in Section 3 were performed to obtain the current rate of water evaporation (fuel moisture soft-sensor value) and the current rate of thermal decomposition of the biomass based on measurements of current air mass flows, current flue gas oxygen content, current steam temperatures, current steam flow and current steam pressure as well as the results of dry fuel analysis. All values were recorded every second. The following assumptions were made: complete combustion, the constant composition of dry biofuel, and simultaneous devolatilization and char burnout.

5. TEST RESULTS OF THE FUEL MOISTURE SOFT-SENSOR

Experiments were conducted to validate the fuel moisture soft sensor. In the test, 10m^3 dry biomass was fed manually into the screw conveyor at a time step marked by a vertical line in Fig. 6. Temperature after the secondary superheater and drum pressure first increase but decrease rapidly due to control action that decrease primary air. Furthermore, flue gas oxygen content is kept at 4% using secondary air. Temperature increase in grate rings 2, 4 and 10 is due to dry fuel that moves to the periphery of the grate as shown in Fig. 7. Furthermore, temperature after the secondary superheater increases again and thus causing fluctuation in process variables. Regardless of disturbances in process variables, fuel moisture soft-sensor shows a decrease in fuel moisture 20 minutes before drum pressure drop and thus opening new possibilities to control combustion.

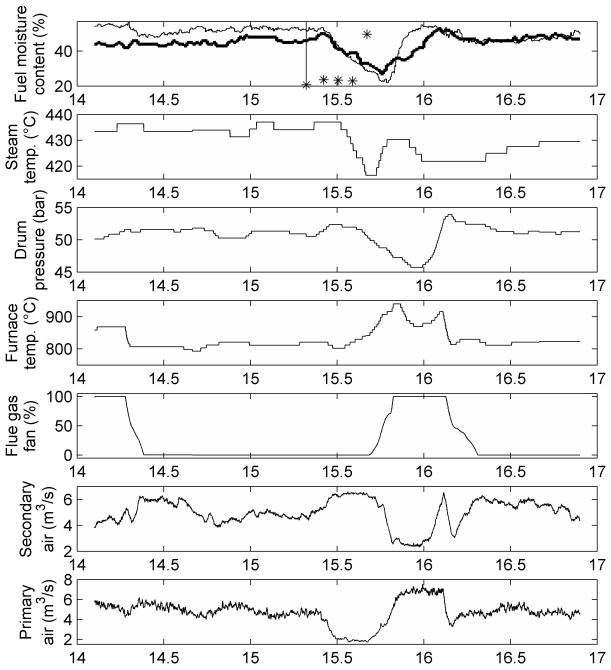


Fig. 6. Boiler measurements during the second test. The top picture includes fuel moisture soft-sensor (thick line), sampled fuel moisture (stars), and fuel moisture calculated from the FT-IR measurement (thin line) for comparison.

The accuracy of the fuel moisture soft-sensor was investigated in the BioPower 5 CHP plant during the tests by sampling fuel feed and with the FT-IR analyzer. According to the fuel sampling, the wet fuel contained 54.4% moisture per kg fuel on average. To compare, the estimated fuel moisture soft-sensor content resulted in an average fuel moisture content of 54.6% per kg fuel. The dry fuel contained 25.9% moisture per kg fuel on average. The estimated fuel moisture soft-sensor content resulted in an average fuel moisture content of 27% per kg fuel. Therefore, the fuel moisture soft-sensor value and the sampled moisture content match really well.

The dynamic behaviour of the fuel moisture soft-sensor was studied in the BioPower CHP plant by producing a step function of moisture in the flue gases by feeding dry biomass manually between wet fuel. The results, presented

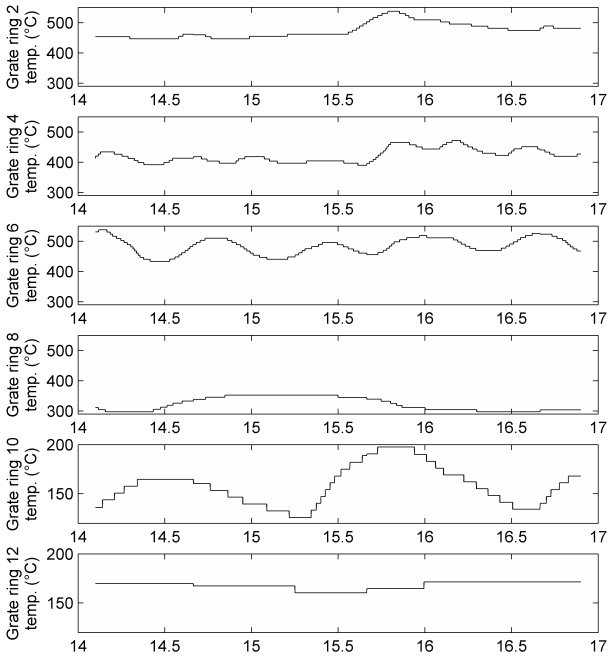


Fig. 7. Grate temperatures during the second test. The grate rings are numbered from the center (Grate ring 2) to the edge of the grate (Grate ring 12)

in Fig. 6, show that the fuel moisture soft-sensor responds to step changes within 1 minute compared with FT-IR in which the 1 minute filter was used. Furthermore, the delay by cross-correlation between combustion power in the oxygen content measurement position and temperature over the secondary superheater is only 45 seconds. Moreover, fuel moisture soft-sensor shows no sign of hysteresis responding equally to both a positive and a negative change in moisture content. This verifies method for detection of varying moisture to be small enough for the method to be used for controlling air and fuel feed.

6. CONCLUSIONS

The fuel moisture soft-sensor based on combustion power estimation and the dynamic model of the secondary superheater was presented in this paper. The experiments were conducted to validate the fuel moisture soft sensor. Two fuels were used to test moisture soft-sensor: spruce bark with an average moisture content of 57%, and dry woodchips (spruce) with the moisture content of 20%. Regardless of disturbances in the process variables, fuel moisture soft-sensor showed a decrease in fuel moisture 20 minutes before drum pressure dropped and thus opening new possibilities to control combustion.

The accuracy of the fuel moisture soft-sensor was investigated in the BioPower 5 CHP plant during the tests by sampling fuel feed and with the FT-IR analyzer. According to the fuel sampling, the wet fuel contained 54.4% moisture per kg fuel on average. To compare, the estimated fuel moisture soft-sensor content resulted in an average fuel moisture content of 54.6% per kg fuel. The dry fuel contained 25.9% moisture per kg fuel on average. The estimated fuel moisture soft-sensor content resulted in an average fuel moisture content of 27% per kg fuel. Therefore, the fuel moisture soft-sensor value and the sampled moisture content matched really well.

The dynamic behaviour of the fuel moisture soft-sensor was studied in the BioPower CHP plant by producing a step function of moisture in the flue gases by feeding dry biomass manually between wet fuel. The results showed that the fuel moisture soft-sensor responds to step changes within 1 minute compared with FT-IR in which the 1 minute filter was used. Thus, the fuel moisture soft-sensor can be used in the control strategy to control both the fuel feed and the combustion air in the furnace.

REFERENCES

Axrup, L., Markides, K., and Nilsson, T. (2000). Using miniature diode array NIR spectrometers for analysing wood chips and bark samples in motion. *Journal of Chemometrics*, 14(5-6), 561–572.

Ayalew, G. and Ward, S.M. (2000). Development of a prototype infrared reflectance moisture meter for milled peat. *Computers and Electronics in Agriculture*, 28(1), 1–14.

Bak, J. and Clausen, S. (2002). FTIR emission spectroscopy methods and procedures for real time quantitative gas analysis in industrial environments. *Measurement Science and Technology*, 13(2), 150–156.

Effenberger, H. (2000). *Dampferzeuger*. Springer, Germany.

Hermansson, S., Lind, F., and Thunman, H. (2011). Online monitoring off fuel moisture-content in biomass-fired furnaces by measuring relative humidity of the flue gases. *Chemical Engineering Research and Design*, 89(11), 2470–2476.

Jaakkola, P.T., Vahlman, T.A., Roos, A.A., Saarinen, P.E., and Kauppinen, J.K. (1998). On-line Analysis of Stack Gas Composition by a Low Resolution FT-IR Gas Analyzer. *Water, Air, & Soil Pollution*, 101(1-4), 79–92.

James, W.L., Yen, Y.H., and King, R.J. (1985). A Microwave Method for Measuring Moisture Content, Density, and Grain Angle of Wood. Technical Report FPL-0250, United States Department of Agriculture, Forest Service.

Kortela, J. and Jämsä-Jounela, S.L. (2010). Fuel quality soft-sensor for control strategy improvement of the Biopower 5 CHP plant. In *Control and Fault-Tolerant Systems (SysTol), 2010*, 221–226. Nice, France, 6-8 October 2010.

Kortela, U. and Lautala, P. (1981). A New Control Concept for a Coal Power Plant. In *Proceedings of the 8th IFAC World Congress*. Kyoto, Japan, 1981.

Kortela, U. and Marttinen, A. (1985). Modelling, Identification and Control of a Grate Boiler. In *American Control Conference, 1985*, 544–549. Boston, MA, USA, 19-21 June 1985.

Lu, S. (1999). Dynamic modelling and simulation of power plant systems. *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, 213(1), 7–22.

Lu, S. and Hogg, B.W. (2000). Dynamic nonlinear modelling of power plant by physical principles and neural

networks. *International Journal of Electrical Power & Energy Systems*, 22(1), 67–78.

Nordell, A. and Vikterlöf, K.J. (2000). Measurements of moisture content in wood fuels with dual energy X-ray. Technical Report ISSN 0282-3772, Värmeforsk.

Nyström, J. and Dahlquist, E. (2004). Methods for determination of moisture content in woodchips for power plants-a review. *Fuel*, 83(7-8), 773–779.

Okamura, S. and Zhang, Y. (2000). New method for moisture content measurement using phase shifts at two microwave frequencies. *Journal of Microwave Power and Electromagnetic Energy*, 35(3), 175–178.

Rosenberg, E., Schatvet, J., and Høydal, K. (2001). In-kiln measurements of moisture content in timber at Moelven Våler AS. In *Third European COST E15 Workshop on wood drying: softwood drying to specific end-uses*, 1–9. Helsinki, Finland, 2001.

Wärtsilä (2005). *Bioenergy solutions*.

Yin, C., Rosendahl, L.A., and Kær, S.K. (2008). Grate-firing of biomass for heat and power production. *Progress in Energy and Combustion Science*, 34(6), 725–754.

Åström, K.J.A. and Bell, R.D. (2000). Drum-boiler dynamics. *Automatica*, 36(3), 363–378.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Temperature Modelling of the Biomass Pretreatment Process

Remus M. Prunescu * **Mogens Blanke** ** **Jakob M. Jensen** ***
Gürkan Sin ****

* *DONG Energy A/S, Innovationscenter, Nesa Allé 1, DK 2820 Gentofte, Denmark (e-mail: rempr@dongenergy.dk)*
** *Department of Electrical Engineering, Automation and Control Group, Technical University of Denmark, Elektrovej, Build. 326, DK 2800 Kgs. Lyngby, Denmark (e-mail: mb@elektro.dtu.dk)*
*** *DONG Energy A/S, Power Concept Optimisation, Nesa Allé 1, DK 2820 Gentofte, Denmark (e-mail: jammu@dongenergy.dk)*
**** *Department of Chemical and Biochemical Engineering, CAPEC, Sølvtofts Plads, Build. 227, Technical University of Denmark (e-mail: gsi@kt.dtu.dk)*

Abstract: In a second generation biorefinery, the biomass pretreatment stage has an important contribution to the efficiency of the downstream processing units involved in biofuel production. Most of the pretreatment process occurs in a large pressurized thermal reactor that presents an irregular temperature distribution. Therefore, an accurate temperature model is critical for observing the biomass pretreatment. More than that, the biomass is also pushed with a constant horizontal speed along the reactor in order to ensure a continuous throughput. The goal of this paper is to derive a temperature model that captures the environmental temperature differences inside the reactor using distributed parameters. A Kalman filter is then added to account for any missing dynamics and the overall model is embedded into a temperature soft sensor. The operator of the plant will be able to observe the temperature in any point of the thermal reactor. Real data sets were extracted from the Inbicon biorefinery situated in Kalundborg, Denmark, and will be utilized to validate and test the temperature model.

Keywords: dynamic modelling, computational fluid dynamics, bioethanol, biomass pretreatment, thermal reactor, biorefinery, Inbicon

1. INTRODUCTION

The worldwide economy is nowadays based on fossil fuels like coal, petroleum and natural gases, which have become increasingly more demanded and difficult to obtain as the current deposits are getting closer to depletion. Fossil fuels are also responsible for most of the climate changes humanity is facing and alternatives to such energy sources receive increasingly more interest. Bioethanol is thought to become the primary renewable liquid fuel (Datta et al., 2011) and solutions to its large scale production from agricultural wastes are intensively investigated.

In this context, DONG Energy built a bioethanol demonstration plant in 2009 at Kalundborg, Denmark, in order to prove that second generation technology of conversion of lignocellulosic biomass waste into ethanol can be profitably applied on a large scale. The conception principle of the plant is the Integrated Biomass Utilization System (IBUS) developed by DONG Energy, which is based on a symbiosis between a biorefinery and a power plant. The IBUS process is commercially exploited by Inbicon A/S, the biomass refinery division of DONG Energy. A detailed description of the refinery process has been documented by Larsen et al. (2008) and is graphically represented in figure 1. The production cycle starts with the pretreatment stage, necessary to break down the biomass into smaller fibres in order to facilitate the subsequent enzymatic digestibility. The pretreatment step is based only on steam from the power plant and recycled water. The next step is the enzymatic liquefaction of the pretreated fibre fraction characterized by a high dry matter content. The resulted slurry is sent to the fermentation tank and is followed by the distillation subprocess. Lignin is recovered as bio-pallets and is utilized as solid fuel in the power plant. Another by-product of the biorefinery is the C5 molasses, a syrup high in nutritional value for livestock.

The successfulness of biomass conversion to ethanol highly depends on the pretreatment stage, which is also responsible for the appearance of inhibitors that affect the enzymatic digestibility (Thomsen et al., 2009). Modelling endeavours have been conducted by Petersen et al. (2009) in order to determine optimal parameters of the pretreatment process. According to (Overend et al., 1987), the pretreatment stage can be characterized by two parameters i.e. the retention time and the steam temperature. These parameters were gathered in a single indicator called the severity factor. Petersen et al. (2009) succeeded in finding a relation between the severity factor and the chemical composition of the slurry that enters the enzymatic treatment stage

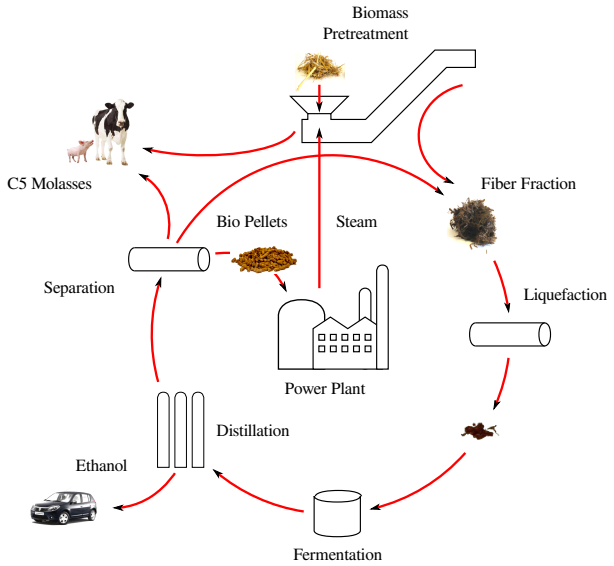


Fig. 1. The IBUS Process (Larsen et al., 2008).

thus offering an estimate of the conversion of biomass to ethanol.

One of the drawbacks of the severity factor developed by Overend et al. (1987) is the fact that it assumes a uniform environment with constant temperature. Therefore, the goal of this article is to find a more accurate mathematical model of the temperature given an irregular moving environment. Knowledge from computational fluid dynamics will be applied in order to account these spatial temperature differences. To simplify simulations, the model is serialized and expressed in a standard state space formulation. In the end, a Kalman filter is added and the overall model will be embedded into a temperature soft sensor that allows the operator to observe how the biomass is treated in any point of the reactor.

2. DESCRIPTION OF THE BIOMASS PRETREATMENT PROCESS

Kristensen et al. (2008) investigated the effects of various pretreatment processes on biomass and his results show that cellulose is not degraded in the hydrothermal pretreatment process but rather becomes more accessible to enzymes due to relocation of lignin and substantial removal of hemicellulose.

The main component of the Inbicon pretreatment process is a pressurized thermal reactor presented in figure 2. Soaked biomass is released from a pressurization unit every 2 min through the left inlet of the tank and, with the help of a motorized snail, the biomass is pushed horizontally with a constant speed till the outlet. The optimal values of the retention time i.e. 15 min and of the reactor temperature i.e. 195 °C($\approx$ 13 bar) were determined based on the experiments of Petersen et al. (2009). The horizontal speed is set to a constant value in order to meet the retention time constraint. The optimal temperature is ensured by a pressure control system that injects saturated steam from the bottom of the reactor through several inlets.

Two temperature measurement belts of 5 sensors each are installed at the beginning and at the end of the tank. The

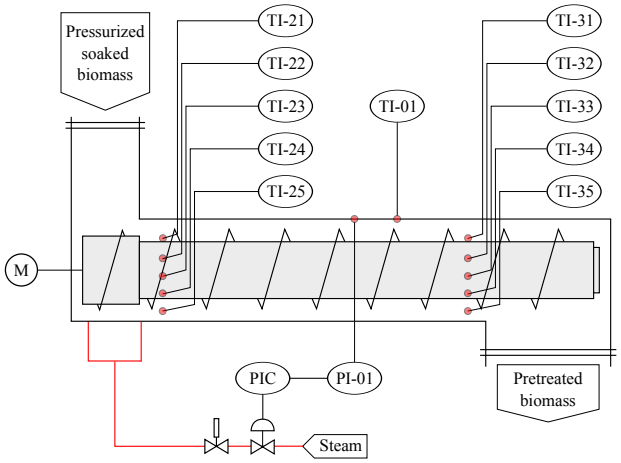


Fig. 2. The thermal reactor schematic diagram with instrumentation.

labels of the temperature sensors, enumerated from top to bottom, are TI-21, TI-22, TI-23, TI-24 and TI-25 for the left group and, respectively, TI-31, TI-32, TI-33, TI-34 and TI-35 for the right series. The reactor pressure is measured by PI-01 and the pressure controller is notated as PIC. A layer of steam is formed in the top part of the reactor as the tank is not fully filled with biomass and its temperature is monitored by TI-01.

3. MATHEMATICAL MODELLING OF THE THERMAL REACTOR

3.1 Preliminary Analysis

The purpose of the thermal reactor modelling is to obtain a temperature gradient that accurately describes the temperature distribution inside the reactor in a two dimensional space. The temperature variations along the width of the reactor are neglected due to its reduced length. Figure 3 contains temperature sensor data that was logged during a nominal operational mode of the plant.

Figure 3 presents temperature variations inside the reactor both on horizontal and vertical axes. A preliminary analysis of the temperature measurements illustrates that the temperature is not uniform but it rather varies considerably mainly on vertical. A difference of about $10 - 15\,^{\circ}\mathrm{C}$ is recorded between the top and the bottom parts of the reactor. The right end of the reactor is open to another subcomponent of the process and it is responsible for a temperature drop on horizontal as the outlet is approached. It is hard to properly model the energy loss due to the open end of the reactor and a Kalman filter will be implemented to account for these effects.

3.2 Temperature Modelling

In order to describe the thermal effects accurately both in space and time, partial differential equations are needed. Computational fluid dynamics collects the tools and methods for describing the heat diffusion in a moving environment such as the heat convection diffusion equation (Egeland and Gravdahl, 2002):

$$\frac{\partial\left(\rho c T\right)}{\partial t} + \nabla^{\mathrm{T}}\left(\rho c \mathbf{u} T\right) = \nabla^{\mathrm{T}}\left(\Gamma c \nabla T\right) + S_T \qquad (1)$$

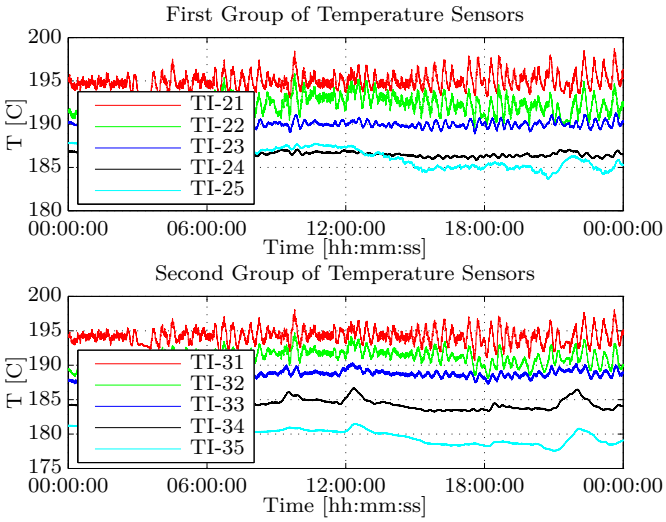


Fig. 3. Preliminary analysis of the temperature inside the reactor. The top subplot contains the temperature recorded by the sensors from the beginning of the reactor while the bottom subplot describes the temperature at the right end of the tank.

where the first term on the left hand side of the equation is the rate of change of temperature T in a fluid element, the second term on the left hand side is the temperature loss due to convection, the first term on the right side is the rate of change due to heat diffusion and the last term is the change caused by the heat source S_T. The density of the material is denoted as ρ, c is the specific heat, $\mathbf{u}$ is the velocity vector and Γ is the diffusion coefficient. The product between the diffusion coefficient Γ and the specific heat c is notated as κ and is called the thermal conductivity of the material:

$$\kappa = \Gamma c \tag{2}$$

In two dimensions, equation (1) can be explicitly written as:

$$\begin{aligned} \rho c\frac{\partial T}{\partial t} + \rho c\frac{\partial\left(\mathbf{u}T\right)}{\partial x} + \rho c\frac{\partial\left(\mathbf{u}T\right)}{\partial y} = \\ = \frac{\partial}{\partial x}\left(\kappa\frac{\partial T}{\partial x}\right) + \frac{\partial}{\partial y}\left(\kappa\frac{\partial T}{\partial y}\right) + S_T \end{aligned} \tag{3}$$

where ρ, c and κ are considered constant. The slurry is pushed horizontally with a constant speed and, therefore, the velocity vector $\mathbf{u}$ has a single constant component u_x on the x axes:

$$\mathbf{u} = \begin{bmatrix} u_x & 0 & 0 \end{bmatrix}^{\mathrm{T}} \tag{4}$$

Any other mixture effects that might occur due to the movement are neglected. Equation (3) is then rewritten considering the movement on a single axes:

$$\rho c\frac{\partial T}{\partial t} + \rho c u_x\frac{\partial T}{\partial x} = \frac{\partial}{\partial x}\left(\kappa\frac{\partial T}{\partial x}\right) + \frac{\partial}{\partial y}\left(\kappa\frac{\partial T}{\partial y}\right) + S_T \tag{5}$$

Equation (5) is parabolic in time and the finite volume method is a way to solve it (Egeland and Gravdahl, 2002). The first step in the finite volume method is to break down the two dimensional space into control volumes or in a grid as in figure 4 where only one central control volume P and its neighbours north N, south S, east E and west W are shown. The central points of the neighbours are denoted with capital letters while the boundaries are notated as north n, south s, east e and west w. The distance from the west w to the east e boundaries is δx_{we} and represents the width of the control volume. The distance from the north n to the south s boundaries is shown as δx_{ns} and is also called the height of the control volume. In order to simplify calculations it is convenient to choose square control volumes i.e. $\delta x_{we} = \delta x_{ns} = \Delta x$.

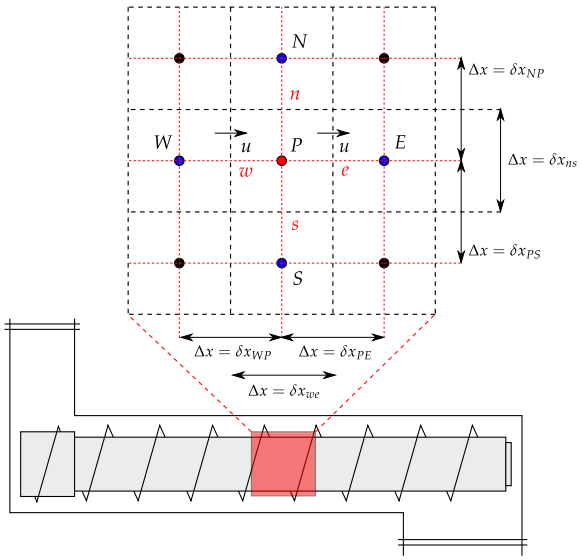


Fig. 4. A control volume with its neighbours from a two dimensional grid.

The Finite Volume Method The solving procedure of equation (5) including its discretization has been moved to appendix A.1 and A.2. The following solution was determined:

$$\begin{aligned} a_P^{n+1}T_P^{n+1} &= a_P^nT_P^n + a_E^nT_E^n + a_W^nT_W^n + a_S^nT_S^n + a_N^nT_N^n + \\ &+ a_E^{n+1}T_E^{n+1} + a_W^{n+1}T_W^{n+1} + a_S^{n+1}T_S^{n+1} + \\ &+ a_N^{n+1}T_N^{n+1} + S_u \end{aligned} \tag{6}$$

with coefficients a_P^n, a_E^n, a_W^n, a_S^n, a_N^n and a_P^{n+1}, a_E^{n+1}, a_W^{n+1}, a_S^{n+1}, a_N^{n+1}:

$$\begin{aligned} a_P^n &= \rho c\frac{\Delta V}{\Delta t} - \theta\left(4D + F\right) \\ a_E^n &= \theta D \\ a_W^n &= \theta\left(D + F\right) \\ a_S^n &= \theta D \\ a_N^n &= \theta D \\ a_P^{n+1} &= \rho c\frac{\Delta V}{\Delta t} + \left(1-\theta\right)\left(4D + F\right) \\ a_E^{n+1} &= \left(1-\theta\right)D \\ a_W^{n+1} &= \left(1-\theta\right)\left(D + F\right) \\ a_S^{n+1} &= \left(1-\theta\right)D \\ a_N^{n+1} &= \left(1-\theta\right)D \end{aligned} \tag{7}$$

where θ is the integration method parameter, D is the diffusion term, F is the convection term, ΔV is the volume of the control object and Δt is the integration step. The integration parameter θ can be 0, 1 or 0.5 corresponding to an implicit Euler, explicit Euler or Crank-Nicholson solver.

A Crank-Nicholson solver is preferred due to its accuracy. Coefficients F and D were explicitly identified as below in appendix A.2:

$$\begin{aligned} F &= \rho c u_x A \\ D &= \frac{\kappa A}{\Delta x} \end{aligned} \tag{8}$$

where A is the side area of the control volume. Because the physical changes of the material along the reactor are neglected i.e. ρ and κ remain constant in any point of the reactor, the diffusion coefficient D becomes also constant regardless of the boundary. If the horizontal speed u_x is held constant then the convection coefficient F becomes constant. If the physical parameters of the biomass change or the horizontal speed is modified then D and F must be updated every simulation step.

The control volumes situated near the borders do not have certain neighbours and are treated differently. The changes that occur in the coefficients are also referred to as the boundary conditions.

Boundary Conditions The boundary conditions are set as suggested by (Bingham et al., 2010). There are two types of boundary conditions depending on whether the temperature is considered known (Dirichlet condition) or the temperature gradient or energy loss is estimated at the border (Neumann type). Figure 5 illustrates the Dirichlet and Neumann setup for a general grid where the boundary condition is notated as T_i or D_i where i is one of the borders i.e. east e, west w, south s or north n.

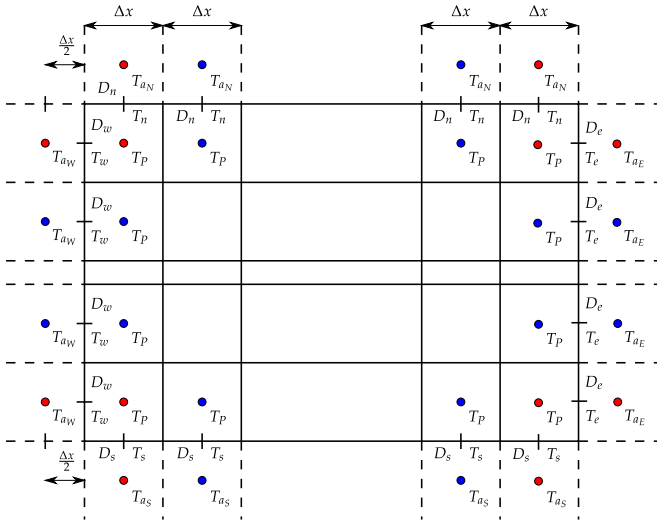


Fig. 5. Dirichlet and Neumann boundary conditions. Dirichlet boundary conditions are denoted as T_i where i is the boundary while Neumann conditions are notated as D_i.

The boundary conditions reflect in several changes of the model coefficients corresponding to the border control volumes. The entire derivation procedure of the coefficients update can be found in appendix A.3. Briefly, a virtual neighbour is created and equation (6) is rewritten considering the extra neighbour.

The reactor is only partially filled with biomass, so the top temperature sensor i.e. TI-01 actually measures the temperature of a steam layer. This measurement constitutes one of the borders and its value is considered the same along the x axes due to the layer of steam. This is a Dirichlet condition and the model coefficients of the northern control volumes change as below:

$$\begin{aligned} \underbrace{(a_P^{n+1} + a_N^{n+1})}_{a_P^{n+1}\leftarrow} T_P^{n+1} &= \underbrace{(a_P^n - a_N^n)}_{a_P^n\leftarrow} T_P^n + a_E^n T_E^n + \\ &+ a_S^n T_S^n + a_W^n T_W^n + a_E^{n+1} T_E^{n+1} + \\ &+ a_S^{n+1} T_S^{n+1} + a_W^{n+1} T_W^{n+1} + \\ &+ \underbrace{S_u + 2T_n(a_N^n + a_N^{n+1})}_{S_u\leftarrow} \end{aligned} \tag{9}$$

After the above updates are performed, coefficients a_N^{n+1} and a_N^n are set to 0 in order to disregard the virtual neighbour. The same Dirichlet conditions apply to the western border of the reactor since the temperature is directly measured.

The energy losses through the bottom part of the reactor are neglected and this fact is translated into a Neumann border condition. The computations can be found in appendix A.3. The following coefficients updates were determined:

$$\begin{aligned} \underbrace{(a_P^{n+1} - a_E^{n+1})}_{a_P^{n+1}\leftarrow} T_P^{n+1} &= \underbrace{(a_P^n + a_E^n)}_{a_P^n\leftarrow} T_P^n + a_W^n T_W^n + a_S^n T_S^n \\ &+ a_N^n T_N^n + a_W^{n+1} T_W^{n+1} + a_S^{n+1} T_S^{n+1} + \\ &+ a_N^{n+1} T_N^{n+1} + \\ &+ \underbrace{S_u - \Delta x D_e(a_E^n + a_E^{n+1})}_{S_u\leftarrow} \end{aligned} \tag{10}$$

The same Neumann conditions are applied to the eastern border of the reactor, which is also considered perfect insulated. The types of boundary conditions have been summarized in figure 6.

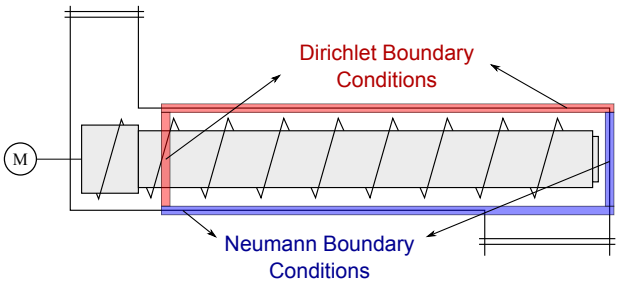


Fig. 6. Dirichlet and Neumann boundary conditions in the thermal reactor case.

An important thing to notice is the fact that, if coefficients (7) do not change in time, then the temperature model described by equation (6) becomes linear and the model can be formulated in a state space manner such that to facilitate fast simulation.

State Space Model In order to derive a state space model, the first step is to serialize and to assign a number to the control volumes from the thermal reactor grid as in figure 7. The first control volume is positioned in the lower left corner of the reactor and the last control volume $n_x n_y$ is situated in the top right corner where n_x and n_y are the dimensions of the grid.

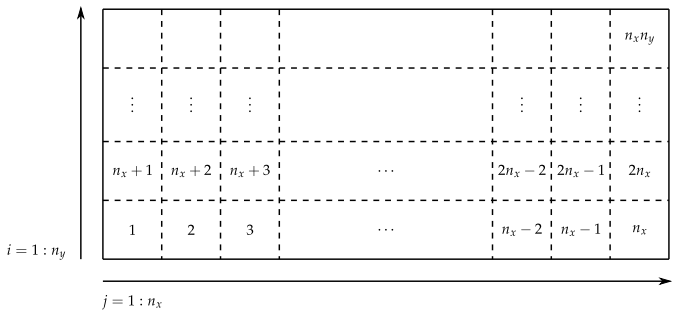


Fig. 7. The control volumes are assigned a number starting from the lowest left corner till the right top corner

Given the coordinates of a volume in the two dimensional space i.e. row i_c and column j_c then the corresponding index i_x in the state vector is found using the relation:

$$i_x = (i_c - 1)n_x + j_c \tag{11}$$

Since the model estimates temperatures, the hat notation will be used from now on to denote the temperature state vector. Let $\hat{x}_T$ comprise the temperatures in the sequentialized control volumes:

$$\begin{aligned} \hat{x}_T^n &= \begin{bmatrix} T_1^n & T_2^n & \cdots & T_{n_x}^n & \cdots & T_{n_x n_y}^n \end{bmatrix}^{\mathrm{T}} \\ \hat{x}_T^{n+1} &= \begin{bmatrix} T_1^{n+1} & T_2^{n+1} & \cdots & T_{n_x}^{n+1} & \cdots & T_{n_x n_y}^{n+1} \end{bmatrix}^{\mathrm{T}} \\ u_T^n &= \begin{bmatrix} S_{u_1}^n & S_{u_2}^n & \cdots & S_{u_{n_x}}^n & \cdots & S_{u_{n_x n_y}}^n \end{bmatrix}^{\mathrm{T}} \end{aligned} \tag{12}$$

where $\hat{x}_T^n$ is the state vector at time step n, $\hat{x}_T^{n+1}$ is the state vector at the next time step and u_T^n gathers all the source terms in the control volumes. Input vector u_T^n will be simplified later because not all of the control volumes have an input source term.

The state space model can be comprised in the following standard equation:

$$\mathbf{E_T}\hat{x}_T^{n+1} = \mathbf{A_T}\hat{x}_T^n + \mathbf{B_T}u_T^n \tag{13}$$

and can be reformulated as:

$$\begin{aligned} \hat{x}_T^{n+1} &= \underbrace{\mathbf{E_T^{-1}A_T}}_{\tilde{\mathbf{A}}_\mathbf{T}}\hat{x}_T^n + \underbrace{\mathbf{E_T^{-1}B_T}}_{\tilde{\mathbf{B}}_\mathbf{T}}u_T^n \\ \hat{x}_T^{n+1} &= \tilde{\mathbf{A}}_\mathbf{T}\hat{x}_T^n + \tilde{\mathbf{B}}_\mathbf{T}u_T^n \end{aligned} \tag{14}$$

where $\tilde{\mathbf{A}}_\mathbf{T}$ is the dynamic matrix and $\tilde{\mathbf{B}}_\mathbf{T}$ is the input matrix of the model.

Matrices $\mathbf{E_T}$, $\mathbf{A_T}$ and $\mathbf{B_T}$ have special structures and are determined by writing equation (6) for each control volume:

$$\mathbf{E_T} = \begin{bmatrix} a_{P_1}^{n+1} & -a_{E_1} & \cdots & \cdots & -a_{N_1} & 0 & \cdots & 0 \\ -a_{W_2} & a_{P_2}^{n+1} & -a_{E_2} & \cdots & 0 & -a_{N_2} & \cdots & 0 \\ \vdots & & & & & & & \\ 0 & & & & & & & \\ -a_{S_{nx+1}} & 0 & & & & & & \\ 0 & -a_{S_{nx+2}} & 0 & & & & & \\ \vdots & & & & & & & \end{bmatrix} \tag{15}$$

$$\mathbf{A_T} = \begin{bmatrix} a_{P_1}^n & a_{E_1} & 0 & \cdots & \cdots & a_{N_1} & 0 & 0 & \cdots & 0 \\ a_{W_2} & a_{P_2}^n & a_{E_2} & 0 & \cdots & 0 & a_{N_2} & 0 & \cdots & 0 \\ 0 & a_{W_3} & a_{P_3}^n & a_{E_3} & \cdots & 0 & 0 & a_{N_3} & \cdots & 0 \\ \vdots & & & & & & & & & \\ 0 & & & & & & & & & \\ a_{S_{nx+1}} & 0 & & & & & & & & \\ 0 & a_{S_{nx+2}} & 0 & & & & & & & \\ \vdots & & & & & & & & & \end{bmatrix} \tag{16}$$

The input matrix $\mathbf{B_T}$ is simplified by keeping only the relevant columns. The temperature of the steam layer is considered the first input into the model and its corresponding column in the $\mathbf{B_T}$ matrix is found from equation (9), where T_n is the input:

$$\mathbf{B_T}(:,1) = \begin{bmatrix} 0 & \ldots & 0 & 2D & 2D & \ldots & 2D \end{bmatrix}^{\mathrm{T}} \tag{17}$$

The temperatures of the western border are directly measured and constitute the remaining inputs:

$$\mathbf{B_T}(:,i) = \begin{bmatrix} 0 & \ldots & 0 & 2(D+F) & 0 & \ldots & 0 \end{bmatrix}^{\mathrm{T}} \tag{18}$$

where i is temperature sensor TI-21, TI-22, TI-23, TI-24 and TI-25. The non-zero coefficient from this column is selected with the help of equation (11) as the coordinates of the control volume where the sensor is positioned are known.

The output $\hat{y}_T^n$ of the system is defined as:

$$\hat{y}_T^n = \tilde{\mathbf{C}}_\mathbf{T}\hat{x}_T^n \tag{19}$$

where $\tilde{\mathbf{C}}_\mathbf{T}$ is the output matrix of the system and is set such to select the temperatures of the control volumes situated on the eastern border.

A system with 6 inputs i.e. TI-01, TI-21, TI-22, TI-23, TI-24 and TI-25 and 5 outputs i.e. TI-31, TI-32, TI-33, TI-34 and TI-35 is, therefore, obtained.

3.3 Kalman Filter

Because the system from equation (14) might be inaccurate when describing the temperature dynamics inside the reactor, a Kalman estimator would be appropriate. Missing dynamics might comprise the rotations of the snail, which contributes to the slurry mixing, gaps in the material where steam can be trapped and contributes to a non-uniform heating or the steam injection from the bottom, which heats the material as the steam travels to the top layer. A static Kalman gain $\tilde{\mathbf{K}}_\mathbf{T}$ is designed such that to ensure best fitting. This is achieved by formulating the state space model as a grey box model with an unknown state noise covariance matrix $\mathbf{R_v}$.

The real system may be written in the following state space formulation:

$$\begin{cases} x_T^{n+1} = \tilde{\mathbf{A}}_\mathbf{T} \cdot x_T^n + \tilde{\mathbf{B}}_\mathbf{T} \cdot u_n + \tilde{\mathbf{G}}_\mathbf{T} v_n \\ y_T^n = \tilde{\mathbf{C}}_\mathbf{T} \cdot x_T^n + e_n \end{cases} \tag{20}$$

where matrices $\tilde{\mathbf{A}}_\mathbf{T}$, $\tilde{\mathbf{B}}_\mathbf{T}$ and $\tilde{\mathbf{C}}_\mathbf{T}$ are defined as in (14) and $\tilde{\mathbf{G}}_\mathbf{T}$ is the state noise propagation matrix. Since all states represent temperatures, it is assumed that the noise is of the same type for all states. All control volumes are affected by its own noise sequence and when augmenting the system together, v_n becomes a column vector with

the same length as x_T^n. The noise vectors v_n and e_n are assumed to be normally distributed white sequences with 0 mean and $\mathbf{R_v}$ and, respectively, $\mathbf{R_e}$ covariance matrices:

$$\begin{cases} v_n \in \mathbf{N}(0, \mathbf{R_v}) \\ e_n \in \mathbf{N}(0, \mathbf{R_e}) \end{cases} \tag{21}$$

The derivation of matrix $\tilde{\mathbf{G}}_\mathbf{T}$ is performed by deviating the temperature variable T from the convection-diffusion equation (5) by a random amount v assumed to be normally distributed with 0 mean and variance R_v. It turns out that $\tilde{\mathbf{G}}_\mathbf{T} = \tilde{\mathbf{A}}_\mathbf{T}$, which is expected since the system was considered linear.

After the static Kalman gain is found, it is desired to bring the model to its innovation form:

$$\begin{cases} \hat{x}_T^{n+1} = \tilde{\mathbf{A}}_\mathbf{T} \cdot \hat{x}_T^n + \tilde{\mathbf{B}}_\mathbf{T} \cdot u_n + \tilde{\mathbf{K}}_\mathbf{T} \cdot e_T^n \\ \hat{y}_T^n = \tilde{\mathbf{C}}_\mathbf{T} \cdot \hat{x}_T^n + e_T^n \end{cases} \tag{22}$$

where e_T^n is the innovation defined as:

$$e_T^n = y_T^n - \tilde{\mathbf{C}}_\mathbf{T} \cdot \hat{x}_T^n \tag{23}$$

3.4 Model Estimation and Validation

Two pairs of input-output data are extracted from the logged sensor information presented in figure 3. One set is used for model estimation and the other one for validation. The grey-box estimation procedure is run and the below state noise variance is found:

$$r_v = 0.0037 \tag{24}$$

which corresponds to a variation of $\pm 0.1566\,^\circ\mathrm{C}$ (2.5758σ). Variance r_v is positioned on the main diagonal of the state noise covariance matrix $\mathbf{R_v}$.

The predicted outputs and the real sensor data are compared. Only the results for one output i.e. TI-35 are presented in figure 8 due to space considerations. The predicted output illustrated with a blue line stays within the yellow confidence interval as shown in the top two subplots. The innovations drawn in the bottom subplots look white and stay within the confidence interval most of the time, which proves that the Kalman filter has a good performance.

Good fitting results were also obtained in the case of the other outputs i.e. TI-34, TI-33, TI-32 and TI-31.

4. SENSOR MOUNTING

The model obtained in equation (22) is embedded into a temperature soft sensor and is connected to the current thermal reactor setup as in figure 9. The soft sensor requires information mainly from the temperature sensor stripes. If the horizontal speed is held constant then the system matrices also remain constant and can be computed offline resulting in a linear model. If the horizontal speed is changed then the system matrices have to be updated as the convection coefficient F needs to be recomputed every time step.

In a block diagram formulation, the temperature soft sensor is connected as in figure 10. The required signals and the outputs of the sensor are explicitly shown in this figure.

The temperature soft sensor provides $\hat{x}_T$, which describes the temperature distribution inside the reactor in a two dimensional space.

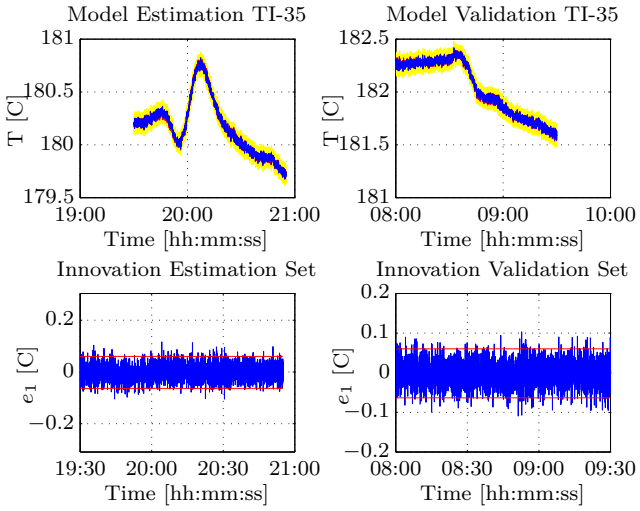


Fig. 8. Model validation TI-35. A fitting of 87% is achieved on both datasets. The top plots present the fitting results while the bottom plots show the innovation or the error estimation.

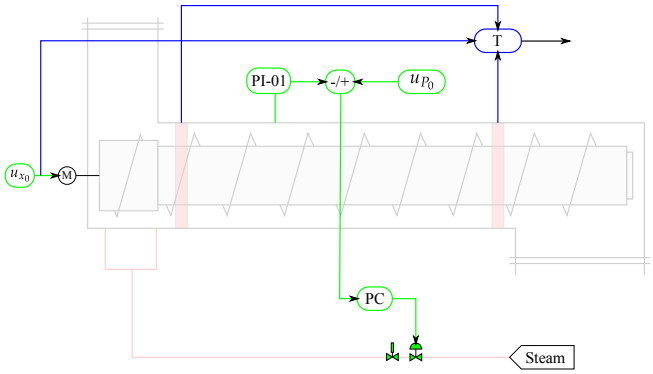


Fig. 9. Sensor mounting. The temperature soft sensor T requires information from the installed temperature measurements and from the nominal horizontal speed. The pressure control system is also illustrated.

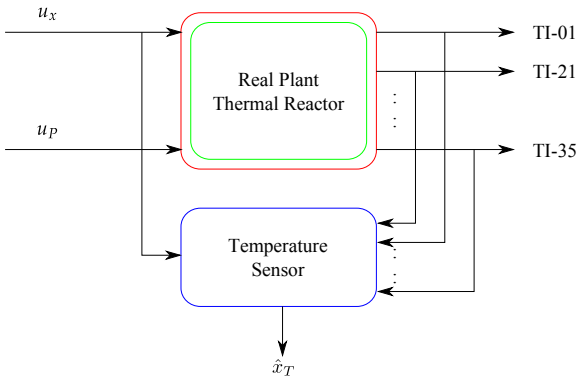


Fig. 10. Temperature sensor mounting block diagram.

5. SIMULATION RESULTS

The temperature soft sensor has been implemented defining a 60x5 grid i.e. 60 divisions on horizontal and 5 on vertical, resulting in an overall model with 300 states. A resolution of 5 divisions on vertical was preferred due to the existing 5 sensors distributed on the height of the reactor.

Real logged data is fed into the sensor and a snapshot is taken during the simulation. Figure 11 displays the temperature distribution constructed from $\hat{x}_T$ as isothermic lines. This figure proves that the temperature model is able to capture the temperature difference between the top and the bottom part of the thermal reactor in real time. As it was seen in figure 3, the temperature is subjected to sinusoidal disturbances. The disturbances are more severe in the upper part of the container and this is also visible in figure 11. The lower right corner of the reactor has an opening towards a downstream component, which acts like a cold source. This is why the temperature decreases as approaching the right ending of the reactor and the effect is well captured by the temperature soft sensor.

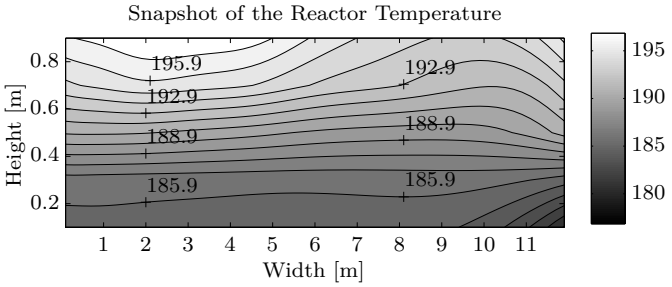


Fig. 11. Snapshot of isothermic lines.

The simulation was run assuming that the thermal reactor is fully filled with biomass, which is, in fact, wrong. There is no direct indication of the biomass level as it is difficult to measure but it could be determined based on the results from the new soft sensor. In figure 11 it is seen that the upper part of the tank changes temperature more rapidly than the lower half portion of the reactor, which rather remains almost constant. This is an indication that the internal environment does not have a constant density. Therefore, the level of biomass can be estimated and, in the current simulation case, a half filled reactor would explain the isothermic lines.

After estimating the level of biomass, the density and thermal conductivity of the control volumes not occupied by biomass can be changed and a more accurate result would be obtained.

6. CONCLUSIONS

A distributed model for the reactor temperature was successfully derived given an irregular environment characterized by temperature differences on both horizontal and vertical axes. The heat convection diffusion equation from computational fluid dynamics proved to give good results in describing the temperature gradient. The Kalman filter also proved to perform well and the obtained temperature gradient captures most of the effects that occur. The state space representation of the model was valuable for efficient simulations and can be further used for monitoring or control purposes.

The operator of the plant is now able to observe how the biomass is treated in any point of the reactor. The operator is also able to analyse the effects of the subsequent components on the reactor temperature, like the temperature drop that occurs near the opening end of the tank.

Valuable information can be obtained from the temperature gradient with respect to the efficiency of the current architecture of the thermal reactor. Ideally, the temperature inside the reactor should be the same in any point of the tank (a uniform environment). The plant operators are now able to observe how efficient the bottom inlets of steam can be and reconfiguration of the reactor can occur in order to achieve a temperature environment closer to the ideal one.

Appendix A. TEMPERATURE MODELLING

A.1 The Finite Volume Method

The key step of the finite volume method is to integrate equation (5) over the control volume:

$$\int_{\Delta V} \rho c \frac{\partial T}{\partial t} dV + \int_{\Delta V} \rho c u \frac{\partial T}{\partial x} dV = \int_{\Delta V} \frac{\partial}{\partial x}\left(\kappa \frac{\partial T}{\partial x}\right) dV + \\ + \int_{\Delta V} \frac{\partial}{\partial y}\left(\kappa \frac{\partial T}{\partial y}\right) dV + \int_{\Delta V} S_T dV \tag{A.1}$$

Using the divergence theorem (Egeland and Gravdahl, 2002, p. 403), the integrals containing partial derivatives with respect to x and y from equation (A.1) can be rewritten like below:

$$\int_{\Delta V} \rho c u \frac{\partial T}{\partial x} dV = \int_{\Delta V} \frac{\partial}{\partial x}\left(\rho c u T\right) dV = \int_{A} \mathbf{n}^{\mathrm{T}}\left(\rho c u T\right) dA \\ \int_{\Delta V} \frac{\partial}{\partial x}\left(\kappa \frac{\partial T}{\partial x}\right) dV = \int_{A} \mathbf{n}^{\mathrm{T}}\left(\kappa \frac{\partial T}{\partial x}\right) dA \\ \int_{\Delta V} \frac{\partial}{\partial y}\left(\kappa \frac{\partial T}{\partial y}\right) dV = \int_{A} \mathbf{n}^{\mathrm{T}}\left(\kappa \frac{\partial T}{\partial y}\right) dA \tag{A.2}$$

where $\mathbf{n}$ is a unit vector normal to the surface. The temperature in the control volume is assumed uniform and this allows the evaluation of integrals (A.2) at the east and west or south and north boundaries respectively:

$$\int_{A} \mathbf{n}^{\mathrm{T}}\left(\rho c u T\right) dA = \left(\rho c u A T\right)_e - \left(\rho c u A T\right)_w \\ \int_{A} \mathbf{n}^{\mathrm{T}}\left(\kappa \frac{\partial T}{\partial x}\right) dA = \left(\kappa A \frac{\partial T}{\partial x}\right)_e - \left(\kappa A \frac{\partial T}{\partial x}\right)_w \\ \int_{A} \mathbf{n}^{\mathrm{T}}\left(\kappa \frac{\partial T}{\partial y}\right) dA = \left(\kappa A \frac{\partial T}{\partial y}\right)_s - \left(\kappa A \frac{\partial T}{\partial x}\right)_n \tag{A.3}$$

The results from (A.3) are substituted into (A.2) and then into (A.1) and the following result yields:

$$\int_{\Delta V} \rho c \frac{\partial T}{\partial t} dV + \left(\rho c u A T\right)_e - \left(\rho c u A T\right)_w = \\ \left(\kappa A \frac{\partial T}{\partial x}\right)_e - \left(\kappa A \frac{\partial T}{\partial x}\right)_w + \\ + \left(\kappa A \frac{\partial T}{\partial y}\right)_s - \left(\kappa A \frac{\partial T}{\partial y}\right)_n + \bar{S}\Delta V \tag{A.4}$$

where the source term in a control volume is averaged as $\bar{S}$ and, therefore, considered constant inside the volume. The

partial derivative with respect to time is isolated on the left hand side of the equation:

$$\begin{aligned}\int_{\Delta V}\rho c\frac{\partial T}{\partial t}dV = \left(\kappa A\frac{\partial T}{\partial x}\right)_e - \left(\kappa A\frac{\partial T}{\partial x}\right)_w + \\ + \left(\kappa A\frac{\partial T}{\partial y}\right)_s - \left(\kappa A\frac{\partial T}{\partial y}\right)_n + \\ (\rho cuAT)_w - (\rho cuAT)_e + \bar{S}\Delta V\end{aligned} \tag{A.5}$$

Equation (A.5) tells that the accumulation in time of thermal energy is a result of three effects i.e. diffusion, convection and generated heat by the control volume itself. Diffusion occurs on horizontal and vertical, from west to east and from north to south respectively. Accumulation of energy is a balance between energy that enters the control volume plus generated energy minus the amount of energy that leaves the volume.

A.2 Discretization

Equation (A.5) is solved by integrating in time and then is discretized using a parameterized method θ:

$$\begin{aligned}\rho c\Delta V T_P^{n+1} = \rho c\Delta V T_P^n + \\ \Delta t\left[\theta f\left(T_P^n, n, x\right) + (1-\theta) f\left(T_P^{n+1}, n+1, x\right)\right]\end{aligned} \tag{A.6}$$

where f is a function that gathers all the terms on the right hand side of equation (A.5). The integration method depends on parameter θ. If $\theta = 1$ then, from equation (A.6), the temperature in control volume P at time step $n+1$ depends only on previous information from time step n. This procedure is also known as the explicit Euler or backward integration method and is equivalent to building a rectangle from time step n. Parameter θ can also be 0 and then the rectangle is built from step $n+1$ to step n or, in other words, the temperature in control volume P at time step $n+1$ depends only on information from time step $n+1$. When θ is 0, the procedure is also called the implicit Euler or forward integration method. A more accurate method is the Crank-Nicolson procedure, which relies on trapezoids instead of rectangles and it corresponds to $\theta = 1/2$.

Equation (A.6) is divided by the integration time Δt:

$$\begin{aligned}\rho c\frac{\Delta V}{\Delta t}T_P^{n+1} = \rho c\frac{\Delta V}{\Delta t}T_P^n + \\ \theta f\left(T_P^n, n, x\right) + (1-\theta) f\left(T_P^{n+1}, n+1, x\right)\end{aligned} \tag{A.7}$$

where $f\left(T_P^n, n, x\right)$ gathers all the variables from time step n. The partial derivatives from (A.5) can be approximated like in the following example:

$$\left.\frac{\partial T}{\partial x}\right|_e \approx \frac{T_E - T_P}{\delta x_{PE}} \tag{A.8}$$

where the partial derivative of temperature T with respect to x is evaluated at the east boundary of the control volume as the ratio between the difference of temperatures in the centers of the control volumes E and P and the distance between their centers δx_{PE}. By conducting these approximations, the following expression for $f\left(T_P^n, n, x\right)$ is found:

$$\begin{aligned}f\left(T_P^n, n, x\right) = \kappa_e A_e\frac{T_E^n - T_P^n}{\delta x_{PE}} - \kappa_w A_w\frac{T_P^n - T_W^n}{\delta x_{WP}} + \\ \kappa_s A_s\frac{T_S^n - T_P^n}{\delta x_{PS}} - \kappa_n A_n\frac{T_P^n - T_N^n}{\delta x_{NP}} + \\ +\rho cuA_w T_w^n - \rho cuA_e T_e^n + \bar{S}\Delta V\end{aligned} \tag{A.9}$$

The constant coefficients from (A.9) are grouped into:

$$\begin{cases} F_e = \rho cuA_e \\ F_w = \rho cuA_w \end{cases} \qquad \begin{cases} D_e = \dfrac{\kappa_e A_e}{\delta x_{PE}} \\ D_w = \dfrac{\kappa_w A_w}{\delta x_{WP}} \\ D_n = \dfrac{\kappa_n A_n}{\delta x_{NP}} \\ D_s = \dfrac{\kappa_s A_s}{\delta x_{PS}} \end{cases} \tag{A.10}$$

where F is also known as the convection term and D is called the diffusion term.

Two more temperatures remain to be evaluated i.e. the temperature at the east boundary T_e and the temperature at the west boundary T_w. The flow direction is known as being from west to east. It was assumed in the introduction section of this chapter based on the real data that there is not a lot of heat diffusion and the convection part has a higher effect. In such cases, it is recommended to use the upwind difference scheme (UDS), which considers the temperatures at the east and west boundaries T_e and T_w as:

$$T_e = T_P \qquad T_w = T_W \tag{A.11}$$

The choice of the difference scheme can be quantized by the use of the Peclet number, which defines the transportiveness and is a measure of the ratio between convection and diffusion:

$$P_e = \frac{F}{D} \tag{A.12}$$

If the Peclet number P_e is greater than 2 then the convection effect is more prominent and an UDS scheme is more appropriate. If $P_e < 2$ then heat diffusion is more important and another approximation scheme is recommended (Bingham et al., 2010) i.e. the central difference scheme (CDS).

The source quantity of volume P from (A.9) can vary in time and it would be preferred to linearize it around T_P^n:

$$\bar{S}\Delta V = S_u + S_P T_P^n \tag{A.13}$$

and then equation (A.9) becomes:

$$\begin{aligned}f\left(T_P^n, n, x\right) = D_e\left(T_E^n - T_P^n\right) - \\ -D_w\left(T_P^n - T_W^n\right) + D_s\left(T_S^n - T_P^n\right) - D_n\left(T_P^n - T_N^n\right) + \\ +F_w T_W^n - F_e T_P^n + S_u + S_P T_P^n\end{aligned} \tag{A.14}$$

$$\begin{aligned}f\left(T_P^n, n, x\right) = \left(S_P - D_e - D_w - D_s - D_n - F_e\right) T_P^n + \\ +D_e T_E^n + (D_w + F_w) T_W^n + \\ +D_s T_S^n + D_n T_N^n + S_u\end{aligned} \tag{A.15}$$

Function f is also evaluated at time step $n+1$:

$$f\left(T_P^{n+1}, n+1, x\right) = D_e\left(T_E^{n+1} - T_P^{n+1}\right) - \\ -D_w\left(T_P^{n+1} - T_W^{n+1}\right) + \\ +D_s\left(T_S^{n+1} - T_P^{n+1}\right) - D_n\left(T_P^{n+1} - T_N^{n+1}\right) + \\ +F_w T_W^{n+1} - F_e T_P^{n+1} + S_u + S_P T_P^{n+1} \\ f\left(T_P^{n+1}, n+1, x\right) = \\ = \left(S_P - D_e - D_w - D_s - D_n - F_e\right) T_P^{n+1} + \\ +D_e T_E^{n+1} + (D_w + F_w) T_W^{n+1} + D_s T_S^{n+1} + D_n T_N^{n+1} + S_u \tag{A.16}$$

(A.17)

The expressions of function f at time steps n and $n+1$ are substituted into equation (A.7) and all constant coefficients are grouped as below:

$$\begin{aligned} a_P^{n+1} T_P^{n+1} &= a_P^n T_P^n + a_E^n T_E^n + a_W^n T_W^n + a_S^n T_S^n + a_N^n T_N^n + \\ &+ a_E^{n+1} T_E^{n+1} + a_W^{n+1} T_W^{n+1} + a_S^{n+1} T_S^{n+1} + \\ &+ a_N^{n+1} T_N^{n+1} + S_u \end{aligned} \tag{A.18}$$

Coefficients a_P^n, a_E^n, a_W^n, a_S^n, a_N^n and a_P^{n+1}, a_E^{n+1}, a_W^{n+1}, a_S^{n+1}, a_N^{n+1} are detailed next:

$$\begin{aligned} a_P^n &= \rho c \frac{\Delta V}{\Delta t} - \theta\left(D_e + D_w + D_s + D_n + F_e - S_P\right) \\ a_E^n &= \theta D_e \\ a_W^n &= \theta\left(D_w + F_w\right) \\ a_S^n &= \theta D_s \\ a_N^n &= \theta D_n \end{aligned} \tag{A.19}$$

$$\begin{aligned} a_P^{n+1} &= \rho c \frac{\Delta V}{\Delta t} + \\ &+ (1-\theta)\left(D_e + D_w + D_s + D_n + F_e - S_P\right) \\ a_E^{n+1} &= (1-\theta) D_e \\ a_W^{n+1} &= (1-\theta)\left(D_w + F_w\right) \\ a_S^{n+1} &= (1-\theta) D_s \\ a_N^{n+1} &= (1-\theta) D_n \end{aligned} \tag{A.20}$$

If coefficients (A.19) and (A.20) remain constant in time then the temperature model described by equation (A.18) is linear.

A.3 Boundary Conditions

The boundary conditions are set as suggested by (Bingham et al., 2010) and will be explained next. There are two types of boundary conditions depending on whether the temperature is considered known at the border or the temperature gradient or energy loss is estimated at the border. When the temperature is set to a value, the boundary condition is also called a Dirichlet condition. Figure 5 illustrates the Dirichlet setup for a general grid where the boundary condition is notated as T_i where i is one of the borders east e, west w, south s or north n.

In the case of the thermal reactor, the temperature on the northern boundary is directly measured with 3 sensors and is considered constant and known on the x axes. This is the indication of a steam layer that is formed on the top part of the reactor. A virtual or auxiliary control volume is created at the boundary with a temperature T_{a_N} that has to be determined such that at the boundary a temperature T_n or T north is reached. A linear interpolation is used between the temperature from the centers of the auxiliary a_N volume and current P volume:

$$T(y) = T_n + \frac{T_P - T_{a_N}}{\Delta y} y \tag{A.21}$$

where $T(y)$ is the temperature along the y axes, T_n is the temperature at the northern border, T_P is the temperature in the control volume P, T_{a_N} is the temperature of the auxiliary control volume and Δy is the distance between P and a_N. Temperature $T(y)$ is evaluated at the center of the auxiliary volume considering that the y axes has its origin at the northern border and pointing downward:

$$T\left(-\frac{\Delta y}{2}\right) = T_{a_N} \Rightarrow \\ T_{a_N} = T_n - \frac{T_P - T_{a_N}}{\Delta y}\frac{\Delta y}{2} = T_n - \frac{T_P - T_{a_N}}{2} \tag{A.22}$$

from where the center temperature of the auxiliary control volume T_{a_N} is found:

$$T_{a_N} = T_n + \frac{T_{a_N}}{2} - \frac{T_P}{2} \Rightarrow \\ \frac{T_{a_N}}{2} = T_n - \frac{T_P}{2} \Rightarrow T_{a_N} = 2T_n - T_P \tag{A.23}$$

Equation (A.18) is evaluated for the control volumes that have a known temperature at the boundaries considering the auxiliary control volumes. In the case of a northern auxiliary neighbor equation (A.18) becomes:

$$a_P^{n+1} T_P^{n+1} = a_P^n T_P^n + a_E^n T_E^n + a_N^n (2T_n - T_P^n) + \\ +a_S^n T_S^n + a_W^n T_W^n + a_E^{n+1} T_E^{n+1} + \\ a_N^{n+1}(2T_n - T_P^{n+1}) + a_S^{n+1} T_S^{n+1} + a_W^{n+1} T_W^{n+1} + S_u \tag{A.24}$$

The source term and the coefficients of T_P^{n+1} and T_P^n of the control volumes subjected to the boundary conditions change as below:

$$\underbrace{\left(a_P^{n+1} + a_N^{n+1}\right)}_{a_P^{n+1}\leftarrow} T_P^{n+1} = \underbrace{\left(a_P^n - a_N^n\right)}_{a_P^n\leftarrow} T_P^n + a_E^n T_E^n + \\ + a_S^n T_S^n + a_W^n T_W^n + a_E^{n+1} T_E^{n+1} + \\ + a_S^{n+1} T_S^{n+1} + a_W^{n+1} T_W^{n+1} + \\ + \underbrace{S_u + 2T_n\left(a_N^n + a_N^{n+1}\right)}_{S_u\leftarrow} \tag{A.25}$$

It is important to set a_N^{n+1} and a_N^n to 0 after the previously indicated updates have been performed in order to disregard the auxiliary control volume.

The same Dirichlet boundary conditions are applied on the western boundary of the reactor. Remember from the process description chapter that there are several sensors i.e. TI-1211621, TI-1211622, TI-1211623, TI-1211624 and TI-1211625 positioned on the y axes at the beginning of the reactor. The temperature inside the reactor will be

modeled starting from this position onward and, therefore, these sensors constitute the western boundary conditions.

If the tank is insulated then the partial derivative of the temperature with respect to the x or y axes is considered null. If there are losses of heat energy due to insulation then this derivative may have a non-zero value. The reactor is considered perfect insulated at all of its borders although one might include the opening towards the hydrocyclone as a loss of energy. It will be later shown that this energy loss can be neglected. Boundary conditions with known temperature derivatives are also called Neumann type conditions and are illustrated in the same figure 5 where the derivatives D_n, D_w, D_e and D_s are positioned at the borders.

An auxiliary control volume is created just as in the Dirchlet case (Bingham et al., 2010) but this time, the derivative at the border is known and approximated as:

$$D_e \approx \frac{T_P - T_a}{\Delta x} \Rightarrow T_{a_E} = T_P - \Delta x D_e \tag{A.26}$$

where D_e is the derivative at the eastern border considered as an example to illustrate the procedure. The temperature of the auxiliary control volume is then easily found as in (A.26). Equation (A.18) is then reevaluated at the eastern boundary:

$$\begin{aligned} a_P^{n+1}T_P^{n+1} = {} & a_P^n T_P^n + a_E^n(T_P^n - \Delta x D_e) + a_W^n T_W^n + \\ & + a_S^n T_S^n + a_N^n T_N^n + \\ & + a_E^{n+1}(T_P^{n+1} - \Delta x D_e) + a_W^{n+1}T_W^{n+1} + \\ & + a_S^{n+1}T_S^{n+1} + a_N^{n+1}T_N^{n+1} + S_u \end{aligned} \tag{A.27}$$

The source term and the coefficients of T_P^{n+1} and T_P^n are updated as below:

$$\begin{aligned} \underbrace{(a_P^{n+1} - a_E^{n+1})}_{a_P^{n+1}\leftarrow} T_P^{n+1} = {} & \underbrace{(a_P^n + a_E^n)}_{a_P^n\leftarrow} T_P^n + a_W^n T_W^n + a_S^n T_S^n \\ & + a_N^n T_N^n + a_W^{n+1}T_W^{n+1} + a_S^{n+1}T_S^{n+1} + \\ & + a_N^{n+1}T_N^{n+1} + \\ & + \underbrace{S_u - \Delta x D_e(a_E^n + a_E^{n+1})}_{S_u\leftarrow} \end{aligned} \tag{A.28}$$

Again, it is important not to forget to set a_E^n and a_E^{n+1} to 0 in order to disregard the auxiliary point. Following the example for the eastern boundary, the southern border can be processed. Notice that, if the derivative is considered null as in this case, the source term does not need to be updated.

REFERENCES

Bingham, H.B., Larsen, P.S., and Barker, A.V. (2010). *Computational Fluid Dynamics*. DTU.

Datta, R., Maher, M.A., Jonesc, C., and Brinkerd, R.W. (2011). Ethanol - the primary renewable liquid fuel. *Journal of Chemical Technology & Biotechnology*, 86, 473–480.

Egeland, O. and Gravdahl, J.T. (2002). *Modeling and Simulation for Automatic Control*.

Kristensen, J.B., Thygesen, L.G., Felby, C., Jørgensen, H., and Elder, T. (2008). Cell-wall structural changes in wheat straw pretreated for bioethanol production. *Biotechnology for Biofuels*.

Larsen, J., Petersen, M.Ø., Thirup, L., Li, H.W., and Iversen, F.K. (2008). The ibus process - lignocellulosic bioethanol close to a commercial reality. *Chemical Engineering & Technology*, 31, 765–772.

Overend, R.P., Chornet, E., and Gascoigne, J.A. (1987). Fractionation of lignocellulosics by steam-aqueous pretreatments. *Phil. Trans. R. Soc. Lond. A*, 321, 523–536.

Petersen, M.Ø., Larsen, J., and Thomsen, M.H. (2009). Optimization of hydrothermal pretreatment of wheat straw for production of bioethanol at low water consumption without addition of chemicals. *Biomass and Energy*, 33, 834–840.

Thomsen, M.H., Thygesen, A., and Thomsen, A.B. (2009). Identification and characterization of fermentation inhibitors formed during hydrothermal treatment and following ssf of wheat straw. *Biotechnological Products and Process Engineering*, 83, 447–455.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Extended abstract for the 17th Nordic Process Control Workshop

Identification of Switching System

Amir H Shirdel* Hannu Toivonen **

Chemical Engineering Department,Abo Akademi University,
Turku,Finland,(e-mail:*amir.shirdel@abo.fi,**hannu.toivonen@abo.fi)

Keywords: System Identification, Support vector regression (SVR), Linear Switching Systems, Hybrid Systems, Reweighting, Sparse Optimization.

Introduction

Switching systems have the property that their dynamical behaviour may switch between a number of different modes. Identification of switching systems consists of identifying both the individual models which describe the system in the various modes, as well as the time instants when the mode changes have occurred. The problem therefore consists of a numerically demanding coupled parameter identification and clustering problem. Support vector machines are a robust and powerful technique for data classification and black-box system identification problems, whereas novel sparse optimization techniques provide an efficient method for finding sparse solutions of identification problems. In this study, support vector regression is applied in combination with sparse optimization to the switching system identification problem.

Method

System identification is the field of modelling dynamic systems to construct a mathematical model from measured data which describes how the input, the output and the disturbances are related. In this paper we are trying to identify the switching system which has the different dynamic form. Switching system described by two sets of states:

discrete state: determines the active mode
continuous state: evolves according to the dynamics of the active mode.

In mathematical form we are trying to find model parameter θ of the system

$$y(k) = \varphi(k)^T \theta_i + e(k)$$

where

$$\varphi(k)^T = [y(k-1), ..., y(k-r), u(k), ..., u(k-r)]$$
$$i \in \{1, 2, ...\}$$

In our method we are going to identify switching system by using Support Vector Regression and Sparse Optimization together.
Support vector regression tries to find parameter such that the maximum number of data points lie within the epsilon-wide insensitivity tube [Vapnik, 1995].

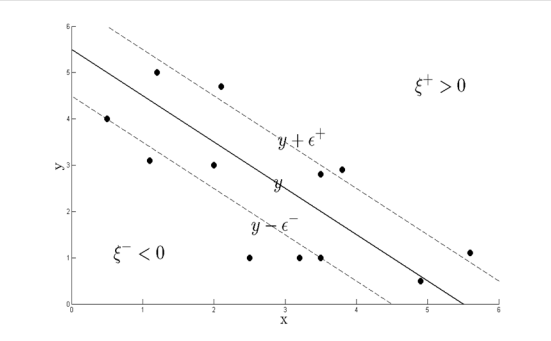


Figure 1: ε-insensitive tube

Identify the parameter θ in $y_k = \varphi_k^T \theta$ by solving :

$$Min \frac{1}{2} \parallel \theta \parallel^2 + C \sum_{K=1}^{N} (\xi_k + \xi_k^*)$$

subject to

$$y_k - \varphi_k^T \leq \varepsilon + \xi_k$$
$$\varphi_k^T - y_k \leq \varepsilon + \xi_k^*$$
$$\xi_k, \xi_k^* \geq 0$$

The key idea is to construct a Lagrange function from the primal objective function and the corresponding constraints, by introducing a dual set of variables. It can be shown that this function has a saddle point with respect to the primal and dual variables at the solution [Smola and Scholkopf ,2003].

In dual problem form, the saddle point condition can be used to eliminate the primal variables$(\theta, \xi, \xi^*)$and Optimal Lagrange multipliers $(\alpha, \alpha^*)$are determined. The optimal parameters are given by:

$$\theta = \sum_{K=1}^{N} (\alpha_k + \alpha_k^*) \varphi_k$$

In sparse optimization we are looking for simple approximate solution of optimization problem. LASSO is a sparse optimization method for linear regression with L1-penalty at the cost of least squares fit. It Shrinks some coefficients and sets others to zero [Tibshirani 1996].

$$Min_u \|y - Au\|_2^2 + \lambda \|u\|_1$$

Dual problem in Support Vector Regression can be written as a LASSO problem for sparse optimization. The crucial question is whether the solution u of the dual problem is sufficiently sparse. Here we can use the fact that our problem is equivalent to LASSO regression. Hence our dual problem approaches the LASSO problem as $C \to \infty$.

$$Min_u \|y - Au\|_2^2 + \varepsilon \|u\|_1$$

subject to

$$-C \leq u \leq C$$

where

$$u_k = \alpha_k + \alpha_k^*$$

With this method, we can select the sparse data points that belong to one mode and find the model parameter of this mode, by applying this method iteratively all of the modes are discovered. As in LASSO, for good results reweighting may have to be performed iteratively.

Example

Consider the simple switching system:

$$y(k) = a(k)y(k-1) + b(k)u(k) + e(k)$$

where

$$a(k) = 0.9 \quad , b(k) = 1 \quad for \quad k = (30n + 1 : 30n + 20)$$
$$a(k) = 0.4 \quad , b(k) = -0.5 \quad for \quad k = (30n + 21 : 30(n + 1))$$

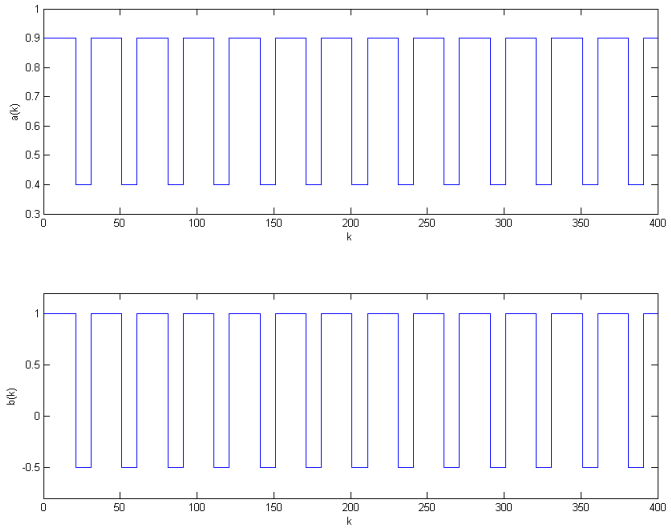


Figure 2: Switching the parameters of system

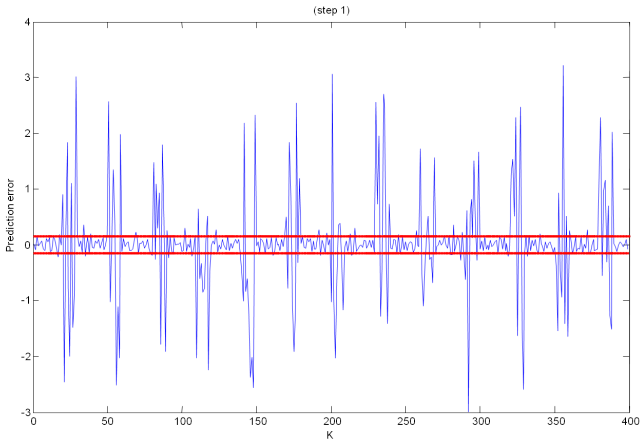


Figure 3: The first iteration to find the first mode and data points

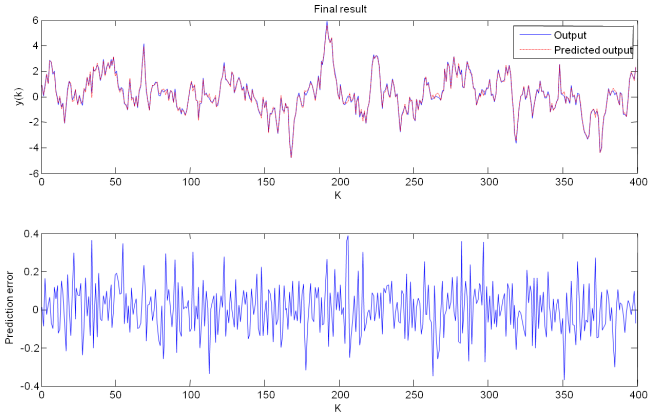


Figure 4: Output,Predicted Output,Predicted Error

U(k) and e(k) are normally distributed with standard deviation 1 and 0.1. SVR design parameters C = 100 and $\varepsilon = 0.15$, the following parameter estimates were obtained:

$$\hat{a}_1(k) = 0.8901 \quad ,\hat{b}_1(k) = 1.0065$$

$$\hat{a}_2(k) = 0.3930 \quad ,\hat{b}_2(k) = -0.4976$$

$$RMSE(Prediction\ error) = 0.1283 \quad RMSE(Noise) = 0.1085$$

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

ModelID, an Interactive Program for Identification of MPC Relevant State-Space Models.

Jørgen K. H. Knudsen *

* *2-control Aps, Frimodtsvej 11, DK-2900 Hellerup, Denmark (e-mail: JoeK@2-control.dk)*

Abstract: This paper describes a practical work flow during identification of linear time invariant state space models for MPC controllers using the new ModelID system identification program. ModelID is designed for process engineering practitioners, who want to develop models without having a detailed knowledge of system identification theory or computer programming. After preprocessing of process data, the user is guided through a set of MISO identifications. After inspection of the impulse responses from the MISO models, the MISO models are combined into a MIMO state space model, using SVD decomposition of Hankel matrices. Finally the state space model is tested in a MPC control loop.
Estimation of time lags in the process are demonstrated using simple (known) processes with different levels of process and measuring noise. The same processes are used to illustrate the problems encountered, when the processes cannot be described as ARX processes. The use of instrumental variable methods and linear filtering provides a solution to these problems.
The identification cycle is supported by many graphical outputs, providing valuable information about the process and the evaluated model.
ModelID is developed for the windows platform, using the C#/.NET based library MPCMath.

Keywords: system identification,instrumental variable methods, state-space models, MPC

INTRODUCTION

The most time consuming task in implementation of MPC control is the development of the model, required for the controller. In many cases linear time invariant models are used for the controller, and typically the models are derived from plant data. Plant data can be historical data from data logging systems, or data generated from carefully planned experiments on the plant. ModelID is an identification program aimed at assisting the development of models. ModelID is designed for people who does not have a deep theoretical background in system identification or computer programming.

ModelD is a GUI based tool, developed for the windows platform, using the C#/.NET based library MPCMath (Knudsen, 2010b).

The purpose of this paper is to illustrate the practical work flow during system identification using ModelID. The paper demonstrates how time delays in the system can be estimated from whitening filtered cross correlations or from impulse responses calculated during the identification cycle. Finally the paper demonstrates the use of instrumental variable algorithms, in cases where the system cannot be adequately described by ARX models

ModelID consist of a set of tools running in a graphical portal. These tools are invoked sequentially during the identification work flow.

The steps and tools in the work flow are:

Tool	Tasks
Data tool	Read data from file and perform initial data treatment.
Model tool	Set model dimensions and process delays. Perform MISO identifications.
Impulse tool	Set impulse response length and calculate impulse responses. Evaluate model quality from impulse responses.
Reduction tool	Perform SVD reduction. Select dimension of state space model and calculate MISO state space models in innovation form.
MPC tool	Initial tuning and test of MPC controller.

DATA TOOL

Process data, read from a comma separated file, is displayed for removal of outliers and removal of fall-outs.

The four tank process is used to illustrate the work-flow. The four tank system, illustrated in Fig. 1, was introduced by Johanson (2000) as a benchmark for control design. Fig. 2 illustrates simulated raw data for this process.

The Controlled variables are the water levels in the four tanks, H1, H2, H3 and H4. The manipulated variables are the two inflows, F1 and F2.

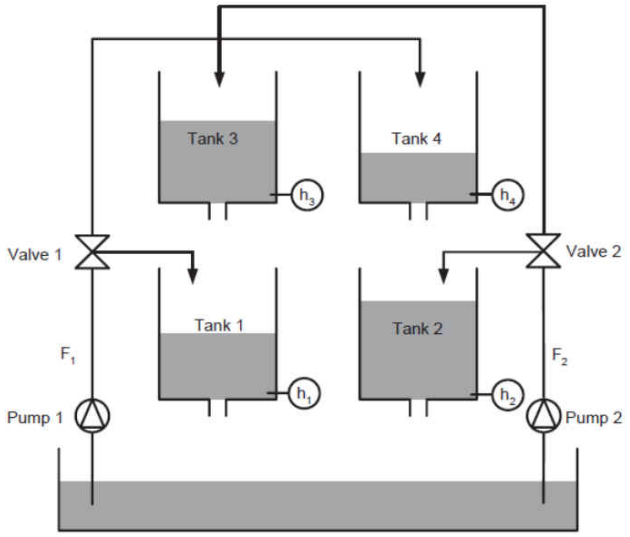


Fig. 1. Four Tank Process used to illustrate process identification work flow

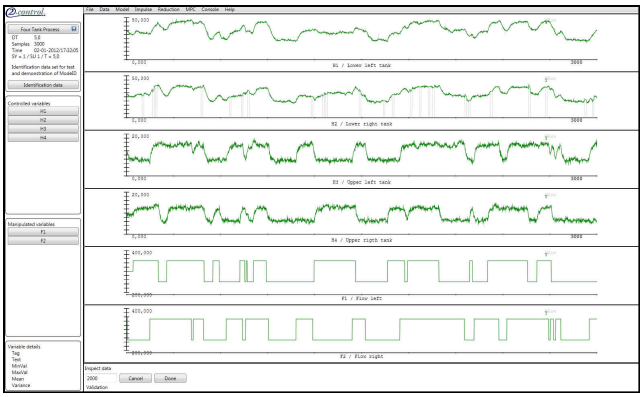


Fig. 2. Simulated raw process data for the four tank process.

The left panel of ModelID's GUI have buttons for the controlled variables and the manipulated variables. These are used to display and edit detailed information for the variables. Fig. 3 illustrates the removal of spikes and fall outs by selection of a threshold parameter for spikes and the value for a fall out situation. Finan et al. (2010) describes the handling of these problems.

After removal of spikes and fall outs the data set is divided into a part used for identification and a part used for validation

MODEL TOOL

The identification procedure

The linear time invariant system is given by

$$Y(t) = G(q)U(t) + H(q)E(t) \tag{1}$$

where

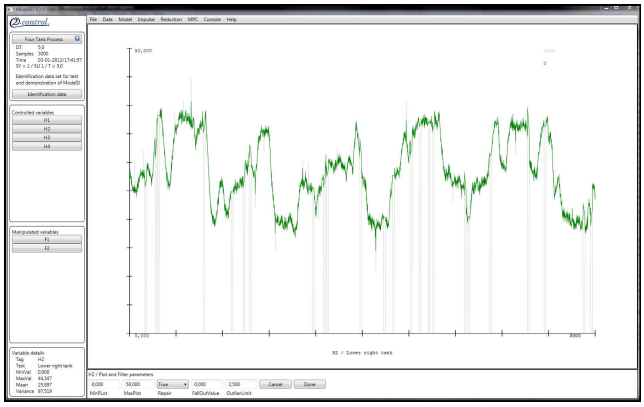


Fig. 3. Spikes and fall outs removal.

$Y(t)$	$\in \mathbb{R}^{n_y}$	Controlled variables.
$U(t)$	$\in \mathbb{R}^{n_u}$	Manipulated variables.
$E(t)$	$\sim N_{iid}(0, I)$	White noise, with I being the identity matrix.
$G(q)$	$\sim \mathbb{R}^{n_y \times n_u}$	Deterministic transfer function between the controlled variables and the manipulated variables.
$H(q)$	$\sim \mathbb{R}^{n_y \times n_y}$	Transfer function of the disturbance model.
q		The time shift operator $qx(t) = x(t+1)$ and $q^{-1}x(t) = x(t-1)$.

The controlled variables, $Y(t)$, are split into a deterministic part $Y_d(t)$ and a stochastic part $W(t)$, with $Y(t) = Y_d(t) + W(t)$. The transfer functions in (1) are assumed to have the structure

$$G(q) = \frac{B(q)}{A(q)} \tag{2}$$

$$H(q) = \frac{\Lambda}{D(q)} \tag{3}$$

where the polynomials are

$$A(q) = I - \sum_{j=1}^{sy} A_j q^{-j} \qquad A_j \in \mathbb{R}^{n_y \times n_y} \tag{4}$$

$$B(q) = \sum_{j=1}^{su} B_j q^{-j} \qquad B_j \in \mathbb{R}^{n_y \times n_u} \tag{5}$$

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_{n_y} \end{pmatrix} \tag{6}$$

$$D(q) = I - \sum_{j=1}^{sd} D_j q^{-j} \qquad D_j \in \mathbb{R}^{n_y \times n_y} \tag{7}$$

The deterministic one step predictor for this system is

$$\hat{Y}(t|t-1) = \sum_{j=1}^{sy} A_j Y(t-j) + \sum_{j=1}^{su} B_j U(t-j) \tag{8}$$

The deterministic predictor for the individual controlled variable is

$$\hat{y}_i(t|t-1) = \sum_{j=1}^{sy} a_{i,j} Y(t-j) + \sum_{j=1}^{su} b_{i,j} U(t-j) \qquad (9)$$

where $a_{i,j}$ and $b_{i,j}$ are the i rows of A_j and B_j , $1 \leq i < n_y$.

The prediction errors for the individual controlled variables are defined by

$$\epsilon_i(t) = y_i(t) - \hat{y}_i(t) \qquad (10)$$

Having n samples of $Y(t)$ and $U(t)$, $0 \leq t < n$, estimated $\hat{A}(q)$ and $\hat{B}(q)$ can be determined minimizing n_y MISO problems

$$V_i = \sum_{t=1}^{n} \ell_i(F_i(q)\epsilon_i(t)) \qquad (11)$$

where ℓ_i are suitable norm functions. In ModelID the $\ell_2, \ell_1, \ell_\infty$ and ℓ_{Huber} norms are provided.

$F_i(q)$ are linear low pass filters reducing the effect of high frequency noise signals, defined by

$$F(q) = \frac{1-f}{1-fq^{-1}} \qquad 0 \leq f < 1 \qquad (12)$$

No filtering is obtained by setting $f = 0$. Selecting $f \to 1$ blocks all information.

Having identified the deterministic part $\hat{G}(q)$ as the combined result of n_y MISO identifications, the stochastic noise signal can be estimated

$$W(t) = Y(t) - \hat{Y}(t) = Y(t) - \hat{G}(q)U(t) \qquad (13)$$

with the one step predictor:

$$\hat{W}(t|t-1) = \sum_{j=1}^{sd} D_j W(t-j) + \Lambda E(t) \qquad (14)$$

Cross correlations and model dimensions

The cross correlations between the manipulated variables and the controlled variables reveal important information about model structure and time delays in the process. Pure cross correlations, as shown in Fig. 4, show that there is no or limited interaction between the pairs F1-H3 and F2-H4.

Filtering the $u(t)$ and $y(t)$ with a filter $F_{wh}(q)$, which tries to make the manipulated variable $u(t)$ as white as possible, increases the information content (as well as the noise level).

$$u_F(t) = F_{wh}(q)u(t) \qquad (15)$$
$$y_F(t) = F_{wh}(q)y(t)$$

The whitening filter $F_{wh}(q)$ is determined by modelling the manipulated variable as an AR-process with dimension 10.

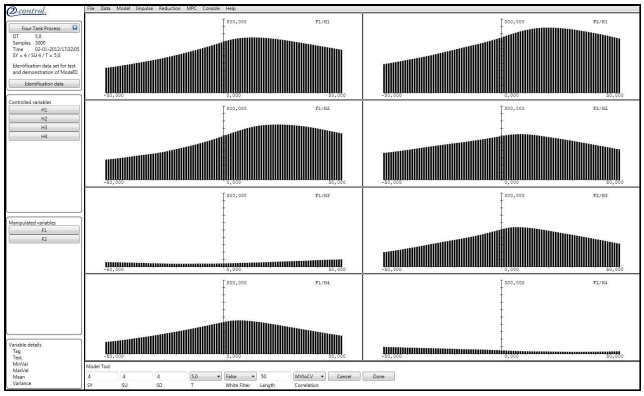


Fig. 4. Model Tool Four Tank Process correlations

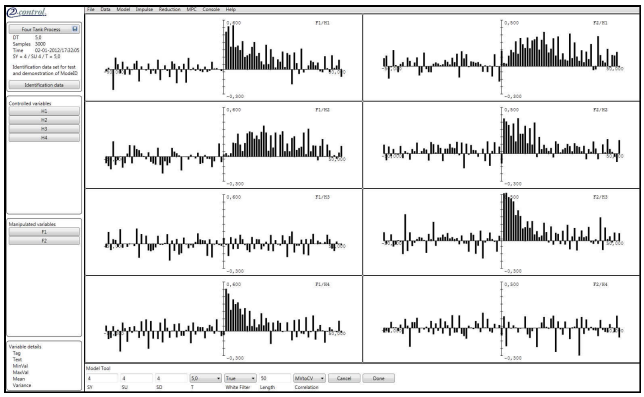


Fig. 5. Model Tool Correlations with white filter option. The correlations show first order dynamic between the pairs F1-H1, F1-H4. F2-H2, F2-H3 and higher order dynamics between F1-H2, F2-H1.

$$F_{wh}(q)u(t) = e(t) \qquad (16)$$

Fig. 5 shows the correlation plot using the white filter option. The correlations indicates that the relations between F1-H1, F1-H4, F2-H2 and F2-H3 could be first order dynamics, and the relations between F1-H2 and F2-H1 second order or higher dynamics. The correlation plot also indicates that there are no pure time delays involved in the plant dynamics.

At this stage the user has to enter estimated delays for the manipulated variables and the dimensions sy, su and sd in equations (8) and (14). This is an iterative process, where the user returns from the subsequent tools to this point, until a satisfactory result is obtained.

Having selected the model dimensions, the results of the individual MISO identifications are displayed by clicking the controlled variable buttons at the left side of the GUI panel. In Fig. 6, the upper left graph shows the plant measurements and the the output of the one step predictor (9). The lower left graph shows the prediction error ϵ_i (10). The upper right graph shows the distribution of ϵ_i, the normal distribution curve corresponding to ϵ_i and the penalty function corresponding to the selected norm.

The lower right quadrant displays the obtained MISO model for the deterministic and stochastic part of the MISO model. At this point the D_j in (14) are assumed

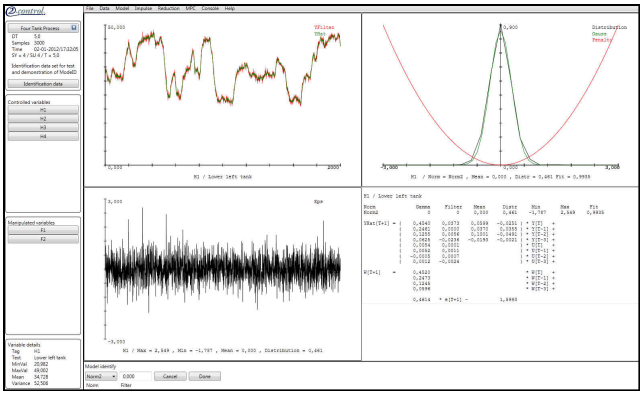


Fig. 6. Model Tool MISO Identification

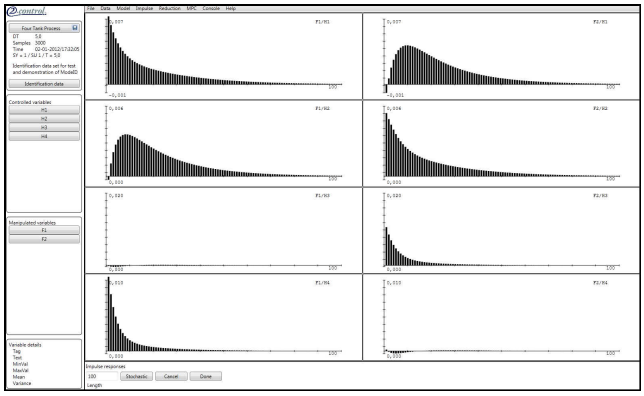


Fig. 7. Impulse responses for the deterministic model.

to be diagonal matrices, giving a "MISO" estimate of the stochastic noise.

At this point the user selects the desired norms and filters for the MISO identifications. Application of the different norms are described in Finan et al. (2010).

The ℓ_2 norm is the standard norm for system identification tasks. The problem can be solved efficiently using either QR or Cholesky factorization algorithms. The disadvantage of the ℓ_2 norm is the sensitivity to outliers. The ℓ_1 norm is less sensitive to outliers and leads to the so called robust identification algorithms.

The Huber Norm, ℓ_{Huber}, defined by (17), is a compromise between ℓ_2 and ℓ_1 norms

$$\ell_{Huber}(\epsilon_i) = \begin{cases} \frac{1}{2}\epsilon_i^2 & |\epsilon_i| \leq \gamma \\ \gamma|\epsilon_i| - \frac{1}{2}\gamma^2 & |\epsilon_i| > \gamma \end{cases} \tag{17}$$

ℓ_1 and ℓ_∞ norms problems are solved as Linear Programming problems. The problems with a Huber-norm are solved as Quadratic Programming problems. Even though parameter estimation using these norms requires more computer resources than least-squares parameter estimation, such problem are solvable in acceptable time on standard laptop computers.

When satisfactory result has been obtained for all the ny identifications, a "MIMO" model for the stochastic noise (3) is calculated before calling the impulse tool. The stochastic model is determined by regression of the predictor equation (14), assuming D_i to be general matrices.

IMPULSE TOOL

The impulse responses for the deterministic part of the identified model are calculated using (8) with

$$\begin{array}{llll} u_i(t) = 1 & t = 0 & 1 \leq i \leq n_u \\ u_i(t) = 0 & t > 0 & \end{array}$$

A sufficiently long length of the calculated impulse responses, M, must be entered, ensuring that they end close to zero.

If the model dimensions sy and su are chosen too high, the impulse responses will start with oscillating components.

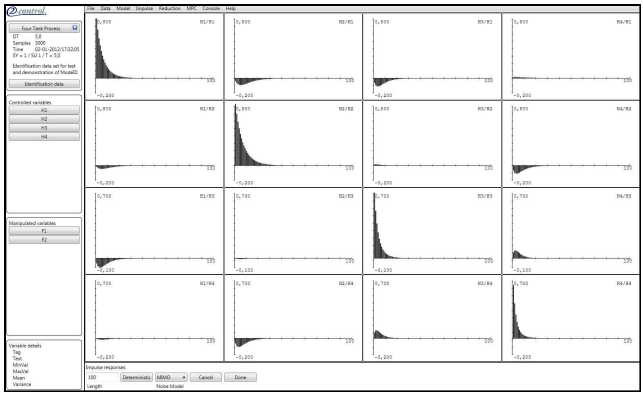


Fig. 8. The impulse responses for the MIMO disturbance model

The impulse responses will also display pure delays between the controlled variables and the manipulated variables. Here the user can return to the Model tool and modify the model dimensions and delays for the manipulated variables.

The impulse responses for the stochastic part are calculated using (14) with

$$\begin{array}{llll} w_i(t) = 1 & t = 0 & 1 \leq i \leq n_y \\ w_i(t) = 0 & t > 0 & \\ E(t) = 0 & t \geq 0 & \end{array}$$

The user can select the desired structure for the stochastic part of the model (D_{MIMO}, D_{MISO} or D_{ARX}). The D_{ARX} option sets $D_i = A_i$ resulting in the noise model of an ARX process. This option is useful if the plant data hos no or negligible stochastic components.

The stochastic model is interesting for two reasons. Primarily it shows to what degree the processes can be described as ARX processes. If the process is not well represented by an ARX structure, the ℓ_2, ℓ_1 and ℓ_{Huber} norm produces biased estimates. The bias increases with the noise level. Secondly, the stochastic model is used to calculate the Kalman gain, K, of the final state space model (18). The Kalman gain is required for tracking between the physical plant and the internal model in the MPC controller. If neither the "MIMO", MISO" or the "ARX" stochastic models are applicable, the Kalman gain must be specified using other methods.

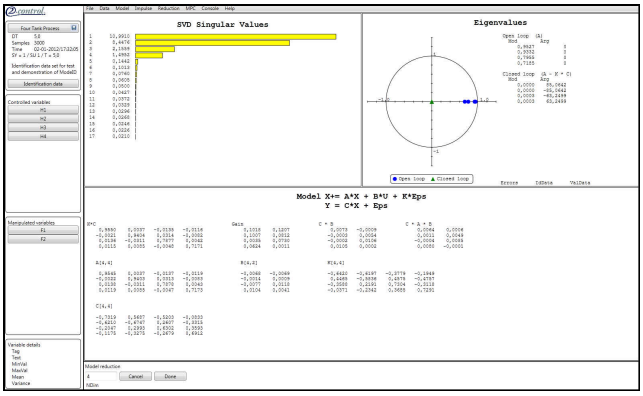


Fig. 9. The reduction tool creates state space models in innovation form.

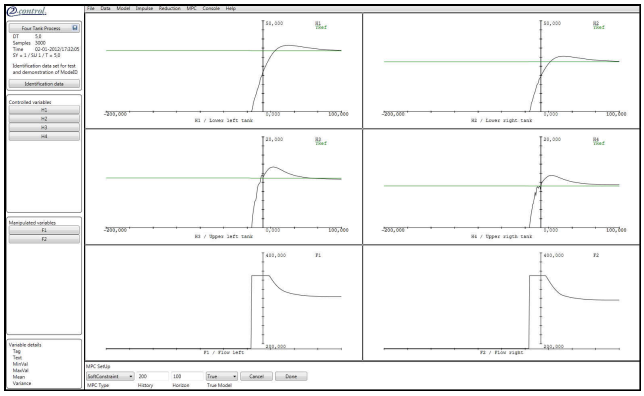


Fig. 10. MPCTool Controlling the Four Tank Process

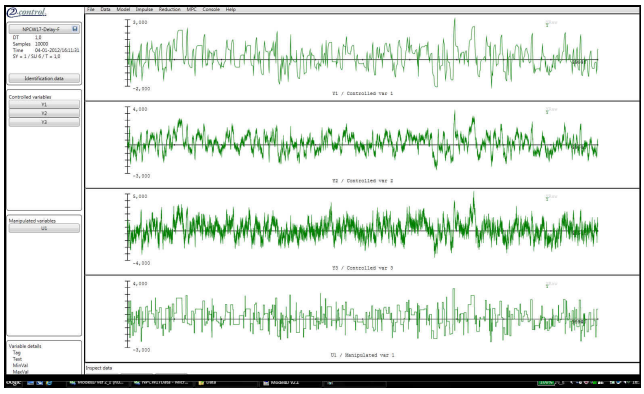


Fig. 11. First order ARX processes signal with increasing noise level

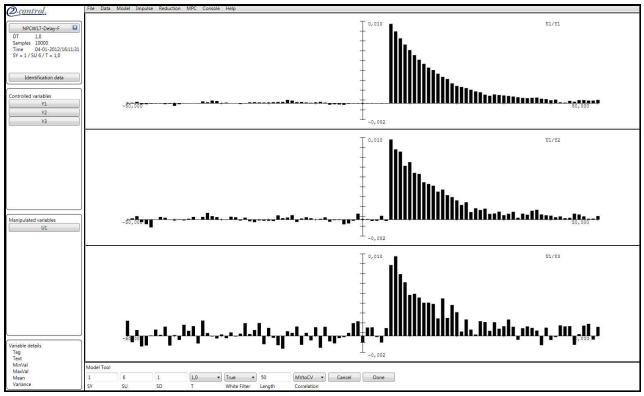


Fig. 12. First order ARX processes cross correlations

REDUCTION TOOL AND STATE-SPACE MODEL

In the reduction tool a block-Hankel matrix is constructed from the impulse responses. The Hankel matrix is factorized using a singular value decomposition (SVD)algorithm. As described in Appendix A, the state space model in innovation form can be realized by model reduction in balanced form using the SVD of the Hankel matrix. The state space model in innovation form is

$$X(t+1) = AX(t) + BU(t) + KE(t) \quad (18a)$$
$$Y(t) = CX(t) + E(t) \quad (18b)$$

The rank of the Hankel matrix is equal to the minimal rank for a state-space system representing the process. For a NDim system, the first NDim singular values are non zero and the subsequent singular values close to zero.

The result of the SVD reduction is shown in Fig. 9. The upper left graph shows the singular values from the SVD. The values show that the system can be represented by a state-space model of dimension NDim = 4. The user enters the desired dimension, NDim, and ModelID calculates the state space model as shown in the lower half of Fig. 9.

The upper right part of the diagram show the open and closed loop eigenvalues ($A - C * K$ in eq. (18))for the state space model.

If the length of the impulse responses is chosen too short, the singular values will decrease gradually, making the selection of the state-space model dimension, NDim, difficult.

MPC TOOL

ModelID includes a MPC module, where tuning of normal MPC and soft constrained MPC (Prasath et al., 2010; Knudsen, 2010a) for the derived State-Space model can be tested simulating the MPC control loop.

ESTIMATING TIME DELAYS

Proper estimation of time delays between the manipulated variables and the controlled variables is important in order to minimize the dimension of the identified state space model. Data from first order and second order processes with pure time delays will be used to illustrate the problem.

The first order system, with a delay of 5 second and a time constant on 10 seconds, is given by

$$y(t) = 0.9048y(t-1) + 0.0952u(t-6) + \sigma * e(t) \quad (19)$$

Fig. 11 shows three responses from (19) with noise variance $\sigma^2 = (0.0, 0.01, 0.1)$. With these noise levels, the time delay of 5 second is shown on all three response. Fig. 13 shows the impulse responses with MISO model dimension $sy = 1$, $su = 6$. All the three responses clearly shows the delay. The user should return to the Model tool at set $su = 1$ and set a delay of 5 seconds for the manipulated variable.

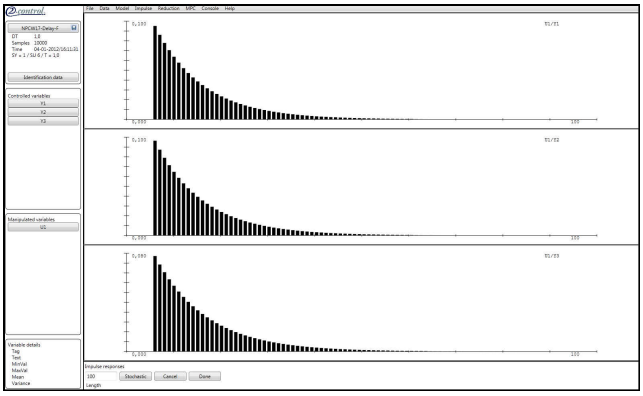


Fig. 13. First order ARX processes Impulse responses

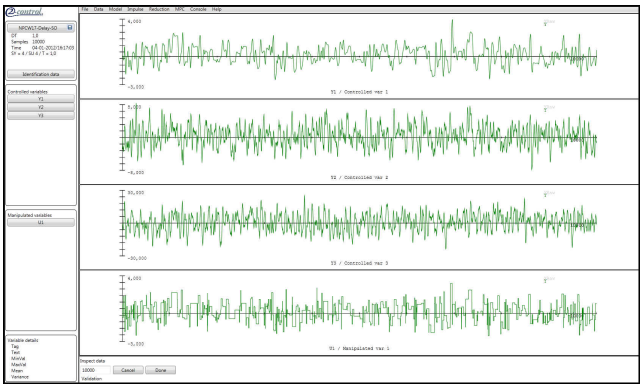


Fig. 14. Second order ARX processes signal with increasing noise level

The second order ARX system, with with a time delay of 5 second, a time constant of 10.0 second and damping of 0.5, is given by

$$y(t) = 1.8953y(t-1) - 0.9048y(t-2) + 0.0047u(t-6) + 0.0047u(t-7) + \sigma e(t) \quad (20)$$

Fig. 14 shows three responses from (20) with noise variance $\sigma^2 = (0.0, 0.01, 0.1)$. With increasing noise level, the proper time delay cannot be determined from the correlation plots on Fig. 15. The impulse responses obtained setting $sy = 2$ and $su = 7$ is shown on Fig. 16. It is evident from Fig. 16 that the proper delay is obtained for all three noise levels. The user should return to the model tool at set $su = 2$ and set a delay of 5 seconds for the manipulated variable.

INSTRUMENTAL VARIABLE METHODS

Regressions using the the $\ell_2, \ell_1, \ell_\infty$ and ℓ_{Huber} norms delivers unbiased estimates if we are dealing with ARX processes. If the process cannot be properly described as an ARX process, the estimate will be biased.

An example is data from a second order Output Error process with a time constant of 10.0 sec and a damping of 1.5. This process is described by

$$y(t) = \frac{b_1 q^{-1} + b_2 q^{-2}}{1 - a_1 q^{-1} - a_2 q^{-2}} u(t) + \sigma e(t) \quad (21)$$

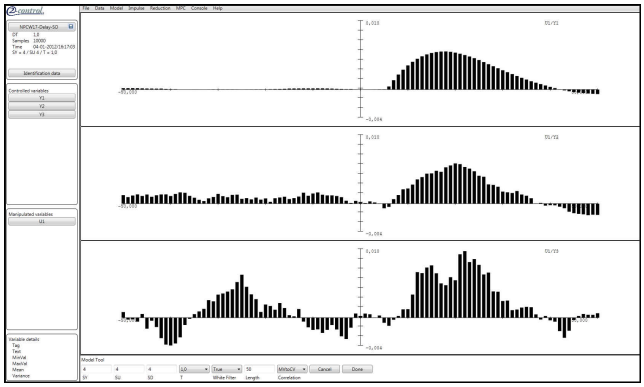


Fig. 15. Second order ARX processes cross correlations.

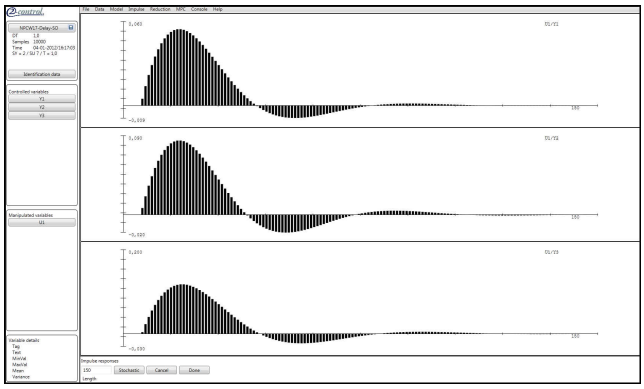


Fig. 16. Second order ARX processes impulse responses.

with $a_1 = 1.7322$, $a_2 = -0.7408$, $b_1 = 0.0045$ and $b_2 = 0.0041$

Regression results with the ℓ_2 norm, using ModelID are shown if Table 1. The parameters for $\sigma^2 = 0.0$ are the correct ones. The results obtained with σ^2= 0.01 and 0.1 are very biased.

Table 1. ℓ_2 norm estimates

σ^2	a_1	a_2	b_1	b_2
0.0	1.7322	-0.7408	0.0045	0,0041
0.01	0.5515	0.3382	0.0028	0,0730
0.1	0.4376	0.4282	0.0262	0,0954

The Instrumental Variable methods (Ljung, 1999; Söderström, 2000) are a possible solutions to his problem. ModelID has implemented the IV4 algorithm (Ljung, 1999) , which can be selected during the MISO identifications with the Model tool. The results obtained with the IV4 algorithm are shown if Table 2.

Table 2. IV4 estimates with Filter = 0.5

σ^2	a_1	a_2	b_1	b_2
0.0	1.7322	-0.7408	0.0045	0,0041
0.01	1.8953	-0.9048	0.0126	0.0031
0.1	1.9390	-0.9447	0.0126	0,0069

The IV4 algorithm gives a much better estimate. In some cases the IV4 algorithm gives an unstable predictor, which luckily is very clearly seen. The predictor can be stabilised using the filter option. If stable, the prediction are rather insensitive to the selected value of the filter.

CONCLUSION

The ModelId is a convenient tool to estimate linear time invariant model for MPC controllers. The aim is to present a tool which is relatively simple to use for the practical user.

In he future it would be natural to include sub-space methods, and support for Identification of closed loop data. Further investigations of Instrumental Variable algorithms is another interesting field.

REFERENCES

Finan, D.A., Jørgensen, J.B., Poulsen, N.K., and Madsen, H. (2010). Robust model identification applied to type 1 diabetes. In *American Control Conference, Marriott Waterfront, Baltimore, MD, USA*.

Johanson, K.H. (2000). The quadruple-tank process: A multivariable laboratory process with an adjustable zero. *IEEE Transactions on control systems technology*, 8(3), 456–465.

Knudsen, J.K.H. (2010a). Implementing model predictive control in the csharp/.net environment. In *Model Based Control Conference. DTU*.

Knudsen, J.K.H. (2010b). Introduction to mpcmath. "`http://www.2-control.dk`".

Ljung, L. (1999). *System Identification, Theory for The User*. Prentice Hall, second edition.

Maciejowski, J. (2002). *Predictive Control with constraint*. Prentice Hall.

Prasath, G., Recke, B., Chidambaram, M., and Jørgensen, J. (2010). Application of soft constrained mpc to a cement mill circuit. In *9th International Symposium on Dynamics and Control of Process Systems,DYCOPS 2010*.

Söderström, T. (2000). *Instrumental variable methods for system identification*. Springer.

Appendix A. GENERATING THE STATE SPACE MODEL

The Markov parameters for the process can be formed from the impulse responses (Maciejowski, 2002).

$$H(t) = \begin{pmatrix} h_{11}(t) & h_{11}(t) & \dots & h_{1n_u}(t) & h^e_{11}(t & \dots & h^e_{1n_y}(t) \\ h_{21}(t) & h_{21}(t) & \dots & h_{2n_u}(t) & h^e_{21}(t) & \dots & h^e_{2n_y}(t) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ h_{n_y1}(t) & h_{n_y1}(t) & \dots & h_{n_yn_u}(t) & h^e_{n_y1}(t) & \dots & h^e_{n_yn_y}(t) \end{pmatrix} \tag{A.1}$$

where $h_{ij}(t)$ is the deterministic impulse response for Controlled variable i to Manipulated variable j at time t and $h^e_{ij}(t)$ is the stochastic impulse response for Controlled variable i to Manipulated variable j at time t

From this the Hankel matrices $\mathcal{H}_{M,M}$ and $\bar{\mathcal{H}}_{M+1,M+1}$ can be formed:

$$\mathcal{H}_{M,M} = \begin{pmatrix} H(1) & H(2) & \dots & H(M) \\ H(2) & H(3) & \dots & H(M+1) \\ \vdots & \vdots & \ddots & \vdots \\ H(M) & H(M+1) & \dots & H(2M-1) \end{pmatrix} \tag{A.2}$$

$$\bar{\mathcal{H}}_{M+1,M+1} = \begin{pmatrix} H(2) & H(3) & \dots & H(M+1) \\ H(3) & H(4) & \dots & H(M+2) \\ \vdots & \vdots & \ddots & \vdots \\ H(M+1) & H(M+2) & \dots & H(2M) \end{pmatrix} \tag{A.3}$$

Singular Value Decomposition, SVD, gives

$$\mathcal{H}_{M,M} = [K_1 K_2]] \begin{bmatrix} \Lambda_1 & \\ & \Lambda_2 \end{bmatrix} [L_1 L_2]' \approx K_1 * \Lambda_1 L_1' \tag{A.4}$$

From this the matrices for the State-Space model in innovation form can be calculated

$$X(t+1) = AX(t) + BU(t) + KE(t) \tag{A.5a}$$
$$Y(t) = CX(t) + E(t) \tag{A.5b}$$

where

$$A = \Lambda_1^{-1/2} * K_1' \bar{\mathcal{H}}_{M+1,M+1} L_1 \Lambda_1^{-1/2} \tag{A.6a}$$
$$\tilde{B} = \Lambda_1^{1/2} [(L_1)_{1:n_u+n_y,:}]' \tag{A.6b}$$
$$B = \tilde{B}_{1:n_u,:} \tag{A.6c}$$
$$K = \tilde{B}_{n_u:n_u+n_y,:} \tag{A.6d}$$

Data Mining for Process Identification

Daniel Peretzki, Alf Isaksson, André Carvalho Bittencourt*

Linköping University,
Department of Electrical Engineering
SE-581 83 Linköping, Sweden

Krister Forsman

Perstorp AB
SE-284 80 Perstorp, Sweden

Abstract

Performing experiments for system identification is often a time-consuming task which may also interfere with the process operation. With memory prices going down, it is more and more common that years of process data are stored (without compression) in a history database. The rationale for this work is that in such stored data there must already be intervals informative enough for system identification. Therefore, the goal of this project was to find an algorithm that searches and marks intervals suitable for process identification (rather than performing completely automatic system identification). For each loop, 4 stored variables are required; setpoint, manipulated variable, process output and mode of the controller.

The proposed method requires a minimum of knowledge of the process and is implemented in a simple and efficient recursive algorithm. The essential features of the method are the search for excitation of the input and output, followed by the estimation of a Laguerre model combined with a chi-square test to check that at least one estimated parameter is statistically significant. The use of Laguerre models is crucial to handle processes with deadtime without explicit delay estimation. The method was tested on three years of data from more than 200 control loops. It was able to find all intervals in which known identification experiments were performed as well as many other useful intervals in closed/open loop operation.

*also ABB Corporate Research, Västerås, Sweden

MPC at Statoils Kalundborg Refinery

Anne-Katrine Ipsen
Statoil Refinery, Process Control Dept.,
Melbyvej 17, 4400 Kalundborg, Denmark
Ph: +45 5957 4500
e-mail: **d01ai@statoil.com**

Abstract

Statoils Kalundborg Refinery has a wide use of MPC (Model Predictive Control).
The Advanced Process Control history at Kalundborg Refinery goes many years back. The first Distributed Control System (DCS) was implemented in 1986, but before that, some Advanced Process Control (APC) was done using a PMX computer. With the DCS system the start came of the current APC use. For the first many years the APC work was done in close cooperation with Exxon, using their experience, know-how and software. Later, in 1999, the use of MPC started, based on a close cooperation with Statoils Research Centre in Trondheim, Norway.

The MPC tool currently used is based on in-house technology called SEPTIC, Statoil Estimation and Prediction Tool for Interactive Control. SEPTIC is used widely within Statoil, both in upstream and downstream oil production.

Since 1999 a total of 17 MPC applications have been implemented at Kalundborg refinery with a total of 119 Manipulated Variables and 247 Controlled Variables. Total incentive in running MPCs is 135 MDkr/year. Still the MPC implementation continues.

The MPC applications include a major system that uses Dynamic Real Time Optimization in order to control the refinery's gasoil production. Another MPC controls the gasoline blending operations which is a batch type process. Most of the MPCs control distillation columns, whereas some optimize whole process units.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Industrial Challenges from the Cement and Mining Industry

B. Recke*, H. Yazdi,*

**FLSmidth A/S,Vigerslev Alle 77, 2500 Valby, Denmark*
DK (Tel: +45 36182700; e-mail:bre@ flsmidth.com).

Abstract: Model predictive control is widely used in industry today and has been used now for several decades. However some challenges in applying the technique still exist and remains as open questions to be addressed by the academic community or in cooperation between academia and industry. This presentation will try to exemplify some of these challenges. This first challenge in applying model based techniques is to get a model that fits well with the plant dynamics. This should preferably be obtained without disturbing the normal operation of the plant. In an industrial environment where the system has large damping effects and all measurements are quite noisy the traditional academic approach of making small perturbations to the actuator signals does not work very well. In this area more work is needed to come up with a good solution that can give 'good enough' models to be used in control. This raises the second issue how well does the model have to fit the plant in order to be good enough for closed loop control? And by which error calculation method do you get the best numbers to judge this? Assuming that a model is somehow obtained the next open question is tuning rules for MPC. These are lacking from the theoretical considerations and that leave the tuning of e.g. weights and limits as more of an art work or gut and experience based assignment. This makes it very difficult to effectively train responsible persons on the plant that should maintain the system how exactly they should be doing this. Furthermore these maintenance people on plants normally do not have a Ph.d.-degree or similar advanced math education which further emphasizes the need for simple and understandable tuning rules. These and other related practical issues will be given in the presentation with the practical application in the industry of Cement and Mining.

*Keywords**: Industrial examples, MPC, tuning rules***.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Application and Development of Advanced Process Control Schemes in Novozymes Fermentation Pilot Plant

Jonas Andersson*, Stuart M. Stocks

**Novozymes A/S, Krogshoejvej 36, DK-2880 Bagsvaerd, Denmark, E-mail: jae@novozymes.com*

Novozymes, arguably the largest enzyme producer in the world, is constantly looking for innovative control strategies and novel methods for optimizing the key unit operation for production of enzymes; the fermentation process. For this purpose the central fermentation pilot plant plays a key role and extensive collaboration with academia has resulted in many fruitful results over the years. This presentation will provide a review of some of the projects that has been conducted within the areas of fed-batch process control, multivariable control, oxygen control in fungal fermentations and probing control. Finally, a general overview of the control set-up used in the Novozymes pilot plant will be given and possibilities for future collaboration with academia presented.

Keywords: Novozymes, fermentation, probing control, fed-batch control, oxygen control, multivariate control

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Using computer models to save energy: An early warning model for tunnel pasteurizer energy consumption

Falko Jens Wagner*

** Sander Hansen Competence Centre, Krones Nordic,*
2840 Holte, Denmark (e-mail: falko.wagner@krones.dk)

Abstract: The Sander Hansen Virtual Pasteurizer (VP) is a system for monitoring pasteurizer performance. It serves as an early warning model for possible component failures by comparing the theoretical and actual utility consumptions. Typically, the utility consumption of a filling line in a brewery is monitored on a long term basis, e.g. 3-6 months. A component failure on any equipment can therefore cause increased utility consumption for a very long period of time, where production typically is running 24-7. Using a mathematical model to calculate the theoretical consumption and comparing this to the measured, actual consumption proves to be an efficient way to detect possible component failures on a per shift basis, thus potentially saving large amounts of utilities.

Keywords: Tunnel pasteurizer, energy consumption, utilities, prediction model, early warning model.

1. INTRODUCTION

Product models to calculate and predict PUs have been used for a long time to control the performance of a tunnel pasteurizer.
The Virtual Pasteurizer takes mathematical modeling one step further by applying an energy model for the entire pasteurizer to monitor energy and water consumptions for the machine.
Until now, energy and water consumptions have only been measured and recorded, but it has been up to the operators to use this information to analyze and judge the performance of the pasteurizer.

Using the Virtual Pasteurizer, the machine can analyze and judge the current consumption by itself and give an alarm message to the operator as soon as a deviation between the actual and theoretical consumption is detected, usually within 1-2 hours. This enables the operator to correct any component failures or malfunctions immediately, thus reducing waste of energy and water for months.

2. PROCESS

Pasteurization means increasing shelf life by killing bacteria through heat treatment. Unfortunately, the heat treatment also impacts and controls the flavor of the product.
The tunnel pasteurization process consists of heating, pasteurizing and cooling filled products, typically cans or bottles, which are moving through the machine on a flat belt conveyor. The heating section increases the temperature of the product from a typical inlet temperature of 5°C to just below the pasteurization temperature, around 50°C. During the pasteurization phase the products are heated to the final pasteurization temperature slightly above 60°C, where the products remain for a certain period of time. The cooling section reduces the temperature to a typical discharge temperature of 30-35°C for subsequent handling. Each section is divided into several zones, to keep the temperature jumps, and thus thermal tensions, to a minimum.
To minimize the total energy consumption, the heating and cooling sections are regeneratively connected and in energy balance, i.e. the total amount of heating energy equals the total amount of cooling energy during steady state operation. So for normal operation heat is only added in the pasteurization section of the machine. Examples of typical temperatures are shown in Fig. 1.

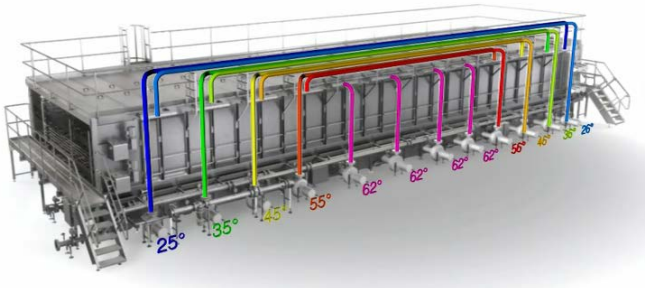


Fig. 1: Picture of Krones Shield Pasteuriser with typical temperatures

A time-temperature representation is shown in Fig. 2. The blue curve represents the spray temperature of the individual steps (zones), the red curve represents the product temperature and the yellow curve represents the degree of pasteurization, or pasteurization units (PU). The green curve shows the prediction model, which is used by the PU Control to control the temperature set points of the pasteurization zones.

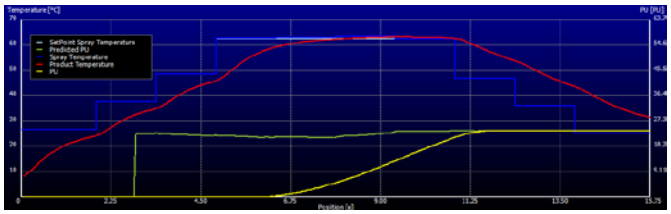


Fig. 2: Pasteurization process

2.1 Pasteurization Units

As indicated in Fig. 2, the goal of the process is to give the products a certain level of pasteurization, measured in Pasteurization Units (PU). Pasteurization Units are empirically defined by the equation:

$$\frac{dPU}{dt} = 10^{\left(\frac{T-X}{Z}\right)}$$

where X = 60°C, Z=6.94 and PUs are measured per minute. The function is normalized around 60°C giving 1 PU per minute and increasing exponentially with temperature. Therefore the typical pasteurization temperature lies just above 60°C. For beer, the typical lower limit of Pasteurization Units is 10 PUs.

2.2 Energy flows

With the given process as described in the beginning of this chapter, the energy flows in the system can be described as depicted in Fig. 3:

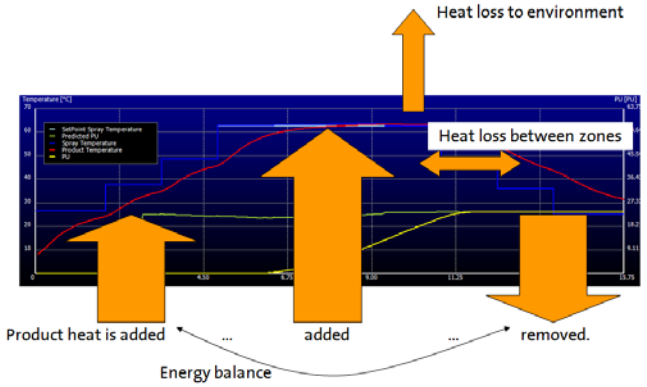


Fig. 3: Energy flows in system

During the heating phase at the beginning of the process, heat is added to the entering cold products through the spraying water. In the central pasteurization phase, further heat is added through heating valves supplying hot water (approx. 85°C) for mixing with the spray water, while keeping the temperatures at an exact set point that reflects the desired Pasteurization Unit uptake. Through the final cooling phase, heat is removed from the products through the spraying water. This spraying water is in turn exchanged with the zones in the heating phase, while the temperatures of these zone pairs are matched carefully so the energy balance is maintained, i.e. the amount of energy added to the products in the heating phase closely matches the amount of energy removed during the cooling phase. The energy is transported by the spraying water from the cooling phase to the heating phase for each connected zone pair, see also Fig. 1.

Throughout the system energy is lost to the environment, as well as heat exchange takes place between adjacent zones with different temperatures. The residual energy is removed by the warm products leaving the machine.

The entire process is controlled by the PU Control system, which monitors and calculated the PU uptake in real-time, while providing the optimal set points for the pasteurization zones to ensure the optimal product treatment. For this purpose a prediction model is used, see also Fig. 2.

2.3 Non-steady-state operation

Now, while the typical steady-state operation of the machine generates a constant energy consumption and thus is easy to calculate, the goal of the early warning model described in this paper is to portrait the energy consumption throughout non-steady-state operation accurately as well.

While the temperatures in the pasteurization zones are kept at a constant temperature during uninterrupted operation, in order to obtain the minimum PU requirement, it is necessary to take action during interruptions that causes the pasteurizer transport belt to stand still. Without taking action, i.e. keeping the temperatures at their normal set points, when the machine is standing still, the PU uptake will increase above the minimum requirement. However, there is also a maximum requirement, which should not be exceeded. To accomplish this, temperatures in the pasteurization zones need to be lowered to a certain level, where the PU uptake is no longer significant. This lower level is typically around 50°C.

When the machine is restarting the conveyors, thus moving the products through the machine, the temperatures in the pasteurization zones are again increased to their typical level for pasteurization during steady-state operation.

3. MATHEMATICAL MODEL OF ENERGY CONSUMPTION

The following section describes the measurement of energy in detail, followed by the calculation of the theoretical consumptions.

3.1 Measurement of actual energy consumption

The actual energy consumption is measured by the flow through and temperature difference across the CHESS (Central Heat Exchanger Supply System). The total energy consumption is calculated by

$$Q_H = m\, C_p\, dT\, \eta$$

where m denotes the mass flow through the CHESS, Cp is the heat capacity for water, dT is the temperature difference across the CHESS and η is the efficiency of the heat exchanger in the CHESS.

Given the total actual energy, this must be distributed to the individual zones. The total flow is the sum of the flows to the individual zones, given by

$F_{Total} = \Sigma(f_i v_i \varphi_i) = m$

where F is the total flow, fi is the maximum flow through the heating valve of the particular zone i, vi is the opening of the heating valve of zone i (between 0 and 1) and φi is the design factor (see also section 6) for the particular valve in zone i depending on the situation in the water system.

The flow through the individual zone is then found by

$f_i = F_{Total} (v_i \varphi_i) / \Sigma(v_i \varphi_i)$

The total energy is then distributed to the individual zones by

$q_i = Q_H f_i / F_{Total}$

3.2 Calculation of theoretical energy consumption

The theoretical energy consumption consists of three parts: energy uptake in products, energy loss to environment and energy loss between zones.

3.3 Energy uptake in products

For each step, the energy uptake in one product is calculated by the energy difference between the temperature before and after ($T_{content}$ and $T_{content\_OLD}$). Since the cycle time for each step is 1s, the energy flow into one product is easily calculated by

$$Q_P = m_P C_P (T_{content} - T_{content\_OLD})$$

which in turn has to be multiplied by number of products standing side by side in the machine

$$n = \frac{D_P b n_{decks} \theta}{\frac{D_P^2}{4}\pi}$$

which reduces to

$$n = \frac{4 b n_{decks} \theta}{D_P \pi}$$

(D_P: diameter of products, b: treatment width of pasteurizer, n_{decks}: number of decks, θ: fill factor)

This has to be accumulated for all products in each zone

$$Q_{P,Zone} = \sum_{Rows\ in\ zone} Q_P n$$

One important information required for the calculation to be accurate is the fill factor. Therefore a measurement of the number of products entering the pasteurizer must be supplied. This can either be accommodated by a product counter at the infeed of the pasteurizer or by using the speed of the filler before the pasteurizer as a delayed signal to the pasteurizer itself.

3.4 Energy loss to environment

For each zone the energy loss to the environment is calculated by

$$Q_L = kA_Z (T_Z - T_{env})$$

(k: heat transfer coefficient to environment, T_Z: temperature in zone, T_{env}: temperature of environment, A_Z: surface area of zone)

The most important information here is the existence of an environment temperature measurement, which must be accommodated.

3.5 Energy loss to neighboring zones

The energy loss to neighboring zones is calculated by the heat transfer coefficient between zones and the temperature difference of the particular zone to the preceding zone and to the following zone:

$$Q_N = k_N A_N (T_{Z-1} - T_Z) + k_N A_N (T_Z - T_{Z+1})$$

(k_N: heat transfer coefficient between zones, T_Z: zone temperature, A_N: area of heat transfer between zones (cross section area))

3.6 Total energy consumption of zone

Finally, the total energy consumption of each zone is the sum of all energy consumptions calculated above:

$$Q_Z = Q_P + Q_L + Q_N$$

4. APPLICATION

To analyze the running state of the pasteurizer, the actual energy consumption of each zone is compared to the theoretical energy consumption of that zone. If the difference is greater than a fixed percentage for a specific period of time, a warning message is displayed. Performing this analysis on zone level gives the operator an indication of the location of the problem.

4.1 Examples of trouble shooting

If one particular zone is using too much energy, compared to the theoretical value, this could indicate one of the following problems:

- The cooling valve in a zone could be leaking. The additional infusion of cold water, which is not detected by the system, is counteracted by an increased heating energy consumption of that zone.

- Water mixing between the zone and its neighbouring (colder) zone could exist. The additional infusion of cold

water, which is not detected by the system, is counteracted by an increased heating energy consumption of that zone.

On the machine scale, the total actual energy consumption can also be used to find a problem when comparing it to the theoretical total energy consumption. If the actual energy consumption of the entire pasteurizer is higher than the total theoretical energy consumption, this could indicate one of the following problems:

- The heat exchanger is clogged and cannot give the desired flow through the system.

- A structural problem in the buffer system of the pasteurizer has occurred and water mixing in the buffer system is taking place.

5. RESULTS

The following picture (Fig. 4) shows the recorded actual energy consumption compared to the theoretical consumption.

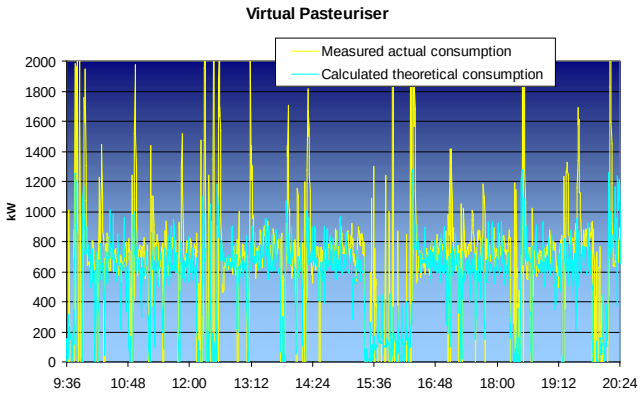


Fig. 4: Comparison of actual (measured) and theoretical (calculated) energy consumption

During normal operation, the result shows a good match. (VP Heat Energy C++ denotes the theoretical energy consumption calculated by the Virtual Pasteuriser, VP Heat Energy PLC denotes the measured actual energy consumption) (Fig. 5):

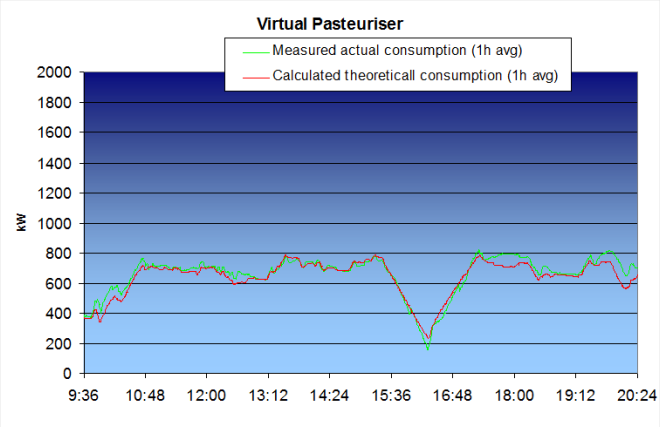


Fig. 5: Actual and theoretical energy consumption based on 1 hour average

6. CASE

The following pictures (Fig. 6) show how a mismatch in the theoretical and actual consumption can be used to display an alarm, telling the operator where to look for possible component failures in the system. First, the measured energy consumption (PLC) in zone 7 is higher than the theoretical (C++). This is displayed on a bar graph and indicated with a red error:

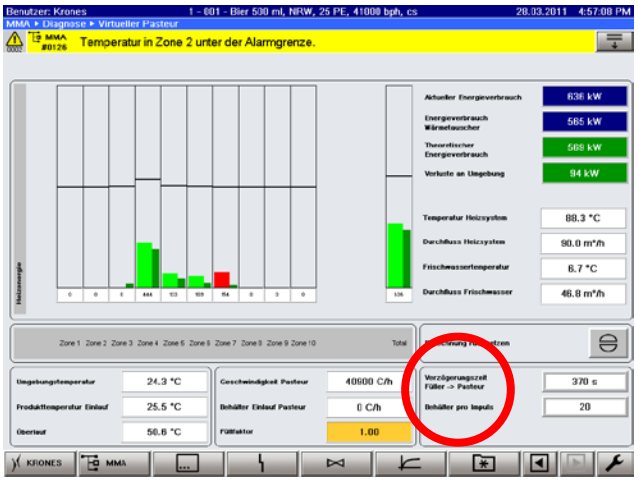


Fig. 6: Visualization of Virtual Pasteurizer

The alarm indicates a possible component failure in the vicinity of zone 7. A closer investigation of the machine in this area revealed the following problem (Fig. 7):

Fig. 7: Water overflow from one zone to another

Water is constantly overflowing from the neighboring, colder zone (8). This causes a decrease in zone temperature of zone 7, which is counteracted by the control system by opening a hot water valve in zone 7 in order to keep the zone at the desired set point.

The normal control system (PU control) does not realize this as being a failure of the machine. Its goal is only to keep a constant temperature by means of adding hot or cold water.

The Virtual Pasteurizer, however, interprets this as a possible error, since the actual energy flow into zone 7 (water from hot water valve) is larger than the theoretical process dictates.

In this case the malfunction could be retraced to a defect pump in zone 8. This disturbed the water balance between the connected zones, causing water to overflow on one end.

As another example a leaking valve should be mentioned, as it can easily have the same effects as described in this case.

6. CHALLENGES

The first tests and installations have shown a working model, but present also some challenges for future development.

6.1 Water distribution

The water distribution model is too simple and assumes certain linearity in the behavior of the flow. For the first implementation the design factor φi was chosen to be the diameter or kws value of the individual valve. However the flow through any given valve is also depending on the other valves and their respective openings in the system.

The challenge here is to derive a model that captures the behavior more accurately.

6.2 Average calculation

As shown above, the average calculation is necessary to get a sufficiently clear picture of the actual status. Especially the digital valves present a challenge, since the flow through these are either very high (40-60 m^3/h) or nothing. And because the valves are only open for approx. 20 seconds at a time, the fluctuations in the assumed energy flow are extreme.

Here the period and/or method of the average calculation plays a significant role and must be chosen carefully. On one hand the smoothing should cancel out the noise, while on the other hand allow to portrait the actual fluctuations accurately enough to avoid accumulation of errors.

The period should be matched to the intrinsic transportation delays in the system as well as production plans and operating schedules.

7. CONCLUSION

By calculating and comparing the theoretical energy and water consumptions of the individual zones and the total pasteurizer with the actual, measured values, an early warning model for possible component failures as well as other malfunctions is established.

This can drastically reduce the energy and water consumption of a tunnel pasteurizer because it gives the operator an immediate warning. Instead of waiting several months before the energy and water household is analyzed and compared to earlier periods, the problem can now be realized (directly on the operator control panel) and fixed the same day as it occurs, thus minimizing the waste of energy and water.

Cases have proven the application of the model, as the model was able to detect a possible failure.

Improvements of the model are needed to make it more accurate. Particularly the water distribution model needs to be extended from its assumed linearity to the actual behavior.

An extension of the model to also include electrical and compressed air consumption is planned.

REFERENCES

EBC Technology & Engineering Forum (1995). *Beer Pasteurisation,* Getränke-Fachverlag.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Optimal Controlled Variables for Parallel Process Units

Johannes Jäschke, Sigurd Skogestad

November 16, 2011

During the lifetime of a process plant, the capacity of the plant is often increased due to a rising demand for products. Starting with some initial capacity, a common solution to cater for increased demand is to add one or more parallel process lines. The resulting parallel plants will be similar, but not entirely identical.

In many cases, the parallel processes share one or several common feed streams, which are distributed to the different lines. An important question which has to be answered for these kinds of parallel processes is how to distribute the feed stream optimally between the individual lines. In practice, it cannot be expected that the optimal value of the split remains constant. It rather is a function of changing disturbances and operating conditions in each line.

This work deals with finding a simple control structure, which guarantees the optimal split between the parallel lines. In particular, this implies finding controlled variables, which can be controlled at a constant setpoint, and which will yield optimal operation even if a disturbance changes the performance in one of the lines.

We assume that the overall operating cost is composed of the sum of the n individual flow dependent costs for each line, $J = \sum_i^n J_i$, and that the total flow q is fixed, i.e. the total flow is the sum of the flow through all lines $q = \sum_i^n q_i$. This results in $n-1$ degrees of freedom to optimize the total cost J, namely the split of the flow rates between the n lines.

A basic result from [1] states that optimal operation with respect to the split can be achieved by ensuring that the individual gradients for each pair of lines are equal, that is $\delta J_i/\delta F_i = \delta J_k/\delta F_k$ is controlled to zero for each pair (i,k). The interpretation of this is that at the optimum, an increase in production of one line must lead to an equally large decrease in performance on another line. In this work, this insight is exploited further, and we show using the ideas from [2] how the individual gradients can be expressed as functions of measurements only. The measurement expressions can then be controlled using simple feedback controllers. We present some case studies, where the optimality condition is used to find simple combinations of controlled variables, which result in the optimal split under varying operating conditions. Examples include parallel operation of CSTRs, furnaces, and optimized oil production from several oil wells.

References

[1] James J. Downs and Sigurd Skogestad. An industrial and academic perspective on plantwide control. *Annual Reviews in Control*, 35(1):99 – 110, 2011.

[2] Johannes Jäschke and Sigurd Skogestad. Optimal controlled variables for polynomial systems. *Journal of Process Control*, Article in press, 2011.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

On Performance for Tracking MPC

J. P. Maree [*] **J. B. Rawlings** [**] **L. S. Imsland** [*]

[*] *Department of Engineering Cybernetics, Norwegian University of Science and Technology, Norway*
[**] *Department of Chemical and Biological Engineering, University of Wisconsin-Madison, USA*

EXTENDED ABSTRACT

Model predictive control (MPC) has been employed extensively in industrial process control systems to deal with constrained multi-variable control problems. In highly automated plants, the goal of optimal economic operation is often addressed by an integrated two-layer control structure, encapsulated within a multi-layer hierarchical control architecture. The first layer, usually referred to as a real-time optimization (RTO) layer, takes economic objectives into account, and subsequently performs a static optimization. In the second layer, MPC is utilized as an advanced regulatory control strategy, whose objective is to dynamically regulate the process to some steady-state operating conditions, as determined in the RTO layer, and furthermore reject any process disturbances.

Control performance monitoring and assessment is an important asset-management technology in maintaining highly efficient operation of any multi-layer hierarchical structure as implemented in process industries. Monitoring of process plants implies watching specific process statistics that reflects the control performance over a period of time. In an regulatory control layer, such statistics can typically include the variance of regulation to desired steady-state set-points, and steady-state tracking errors. These statistical measures give an indication of how well a control strategy is performing its task of regulating the process at desirable set-points. However, what these statistics do not give, is insight into how different control strategies will perform with respect to each other, as they asymptotically steer the process to these set-points in closed-loop operation, given some economic objectives (the latter gives rise to the term *closed-loop asymptotic performance* used in this text).

Since the ultimate objective of the regulatory control layer, incorporating MPC, is to promote economic objectives, any attempt to evaluate MPC control performance should be quantified in terms of closed-loop asymptotic performance. Furthermore, it is desirable to have stabilizing MPC formulations that can serve as a safeguard against sporadic process behavior, which can cause plant-wide instabilities and performance deterioration. It is therefore desirable to find a global control Lyapunov function that establish sufficient conditions for global stability. However, for systems with state and control constraints, it is difficult to obtain a global control Lyapunov function. Hence, we restrict ourselves to the goal of obtaining a local control Lyapunov function in conjunction with some control invariant region. The latter gives rise to terminal penalty and constraint MPC formulations, whose value functions admit a Lyapunov candidate value function.

In this work we investigate, first-most, case studies where the operational set-points (as propagated from the RTO layer), or their respective control invariant regions, are unreachable for a N-step receding horizon policy due to a short receding horizon. Proposed solutions in literature, which adopted the name *Safe-park* strategies, address such infeasible scenarios. Characteristics of such MPC strategies is typically a shifted stage cost (replacing unreachable set-points received from RTO layer with reachable set-points) such that the process is admissibly steered to the closest feasible steady-state. The MPC strategy then penalizes the intermediate feasible steady-state operational set-point with respect to how far the process is operating from the desirable set-point (termed as some offset cost). Such strategies, however, is counter-intuitive when one considers MPC whose primary objective is to promote economic objectives, which is based on the optimal operational set-points propagated from a RTO layer. The proposed MPC strategy in this work is formulated with a stage cost that points to the desirable steady-state set-point, as received from the RTO, without the addition of an additional offset cost. We, furthermore, extend the terminal constraint region to be the convex hull of control invariant regions of the manifold of steady-states, therefore ensuring a feasible MPC solution. We, lastly show that by finding a terminal control law for the extended terminal constraint region, such that we satisfy a basic stability assumption, we can formulate a terminal cost that admits cost drop, using standard MPC stability theory.

We conclude this work by initiating future investigations in how to express closed-loop asymptotic performance when we have different control laws, which both admit acceptable control performance (convergence to desirable set-point and minimal tracking error), however, with different economic cost objectives. Being able to quantify, and understand closed-loop asymptotic performance better, enables us to understand how different control laws and strategies contribute to increased economic performance in closed-loop operation. Initial attempts only establish upper bounds on closed-loop asymptotic performance measures. However, through simulation case studies it can been shown that the structure of a terminal cost function may provide some additional insight.

[1] Corresponding author (J.P. Maree) *phillip.maree@itk.ntnu.no*

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Single Shooting and ESDIRK Methods for adjoint-based optimization of an oil reservoir ⋆

Andrea Capolei * **Carsten Völcker** * **Jan Frydendall** *
John Bagterp Jørgensen *

* *Department of Informatics and Mathematical Modeling & Center for Energy Resources Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark.*
(e-mail: {acap,cv,jf,jbj}@imm.dtu.dk).

Abstract: Conventional recovery techniques enable recovery of $10-50\%$ of the oil in an oil field. Advances in smart well technology and enhanced oil recovery techniques enable significant larger recovery. To realize this potential, feedback model-based optimal control technologies are needed to manipulate the injections and oil production such that flow is uniform in a given geological structure. Even in the case of conventional water flooding, feedback based optimal control technologies may enable higher oil recovery than with conventional operational strategies. The optimal control problems that must be solved are large-scale problems and require specialized numerical algorithms. In this paper, we combine a single shooting optimization algorithm based on sequential quadratic programming (SQP) with explicit singly diagonally implicit Runge-Kutta (ESDIRK) integration methods and a continuous adjoint method for sensitivity computation. We demonstrate the procedure on a water flooding example with conventional injectors and producers.

Keywords: Optimal Control, Optimization, Numerical Methods, Oil Reservoir

1. INTRODUCTION

The growing demand for oil and the decreasing number of newly discovered significant oil fields require more efficient management of the existing oil fields. Oil fields are developed in two or three phases. In the primary phase, the reservoir pressure is large enough to make the oil flow to the production wells. In the secondary phase, water must be injected to maintain pressure and move the oil towards the producers. In some cases, a tertiary phase known as enhanced oil recovery is considered. Enhanced oil recovery includes technologies such as in situ combustion, surfactant flooding, polymer flooding, and steam flooding (Thomas, 2008). After the secondary phase, typically the oil recovery is somewhere between 10% and 50% (Chen, 2007; Jansen, 2011).

Optimal control technology and Nonlinear Model Predictive Control have been suggested for improving the oil recovery of the secondary phase (Jansen et al., 2008). In such applications, the controller adjusts the water injection rates and the bottom hole well pressures to maximize oil recovery or a financial measure such as net present value. In the oil industry, this control concept is also known as closed-loop reservoir management (Jansen et al., 2009). The controller in closed-loop reservoir management consists of a state estimator for history matching and an optimizer that solves a constrained optimal control problem for the production optimization. The main difference of the closed-loop reservoir management system from a traditional Nonlinear Model Predictive Controller (Binder et al., 2001) is the large state dimension (10^6 is not unusual) of an oil reservoir model. The size of the problem dictates that the ensemble Kalman filter is used for state estimation (history matching) and that single shooting optimization algorithms compute gradient based on adjoints (Jansen, 2011; Jørgensen, 2007; Sarma et al., 2005; Suwartadi et al., 2011; Völcker et al., 2011).

In this paper, we propose a high order temporal integration method (Explicit Singly Diagonally Implicit Runge-Kutta, ESDIRK) for forward computation of the initial value problem and for backward solution of the associated continuous-time adjoint. Conventional practice by commercial reservoir simulators is limited to the use of first order temporal implicit or semi-implicit integrators for the initial value problem and the adjoints. Völcker et al. (2010a,b, 2009) introduce high order ESDIRK methods in two phase reservoir simulation. The high order scheme allows larger steps and therefore faster solution of the reservoir model equations. To compute the gradient of the objective function in a single shooting optimization method, Völcker et al. (2011) propose a method based on adjoints for the discretized equations. Cao et al. (2002) and Jansen (2011) provide an overview of gradient computation using the adjoint. Brouwer and Jansen (2004) and Sarma et al. (2005) explain and demonstrate gradient computation by the adjoint equations based on the implicit Euler discretization. Kourounis et al. (2010) suggest the

⋆ This research project is financially supported by the Danish Research Council for Technology and Production Sciences. FTP Grant no. 274-06-0284

continuous-time high order adjoint equations for gradient computation in production optimization. Nadarajah and Jameson (2007) compare gradients computed by discrete and continuous adjoints for problems arising in aerodynamics. They conclude that the gradients computed from continuous adjoints is accurate enough to be used in optimization algorithms. Since computation of gradients based on continuous time adjoints is faster than gradients based on discrete adjoints, this conclusion implies that the gradient computations can be accelerated by using the continuous time adjoint equations.

The novel contribution in this paper is an extension of the adjoint based optimization method suggested by Völcker et al. (2011) to include gradient computation based on the continuous-time adjoint equation. Using a conventional oil field as case study, we demonstrate the new single-shooting optimization algorithm based on ESDIRK integration of the initial value problem and ESDIRK integration of the continuous-time adjoint equation. The case study illustrates the potential of optimal control for production optimization of water flooded oil reservoirs by maximizing the net present value. We do a parameter study to illustrate the sensitivity of the optimal solution to the discount factor.

The paper is organized as follows. Section 2 states the general constrained optimal control problem using a novel representation of the system dynamics. The ESDIRK algorithm for solution of the differential equation systems is described in Section 3, while Section 4 presents the continuous adjoint method. Section 5 describes the numerical case study and discusses the sensitivity of the optimal solution to the discount factor in the net present value. Conclusions are presented in Section 6.

2. OPTIMAL CONTROL PROBLEM

In this section, we present the continuous-time constrained optimal control problem and its transcription by the single shooting method to a finite dimensional constrained optimization problem. First we present the continuous-time optimal control problem. Then we parameterize the control function using piecewise constant basis functions, and finally we convert the problem into a constrained optimization problem using the single shooting method.

Consider the continuous-time constrained optimal control problem in the Bolza form

$$\min_{x(t),u(t)} J = \hat{\Phi}(x(t_b)) + \int_{t_a}^{t_b} \Phi(x(t),u(t))dt \tag{1a}$$

subject to

$$x(t_a) = x_0 \tag{1b}$$

$$\frac{d}{dt}g(x(t)) = f(x(t),u(t)),\ t \in [t_a,t_b], \tag{1c}$$

$$u(t) \in \mathcal{U}(t) \tag{1d}$$

$x(t) \in \mathbb{R}^{n_x}$ is the state vector and $u(t) \in \mathbb{R}^{n_u}$ is the control vector. The time interval $I = [t_a, t_b]$ as well as the initial state, x_0, are assumed to be fixed. (1c) represents the dynamic model and includes systems described by index-1 differential algebraic equations (DAE). (1d) represents constraints on the input values, e.g. $u_{\min} \leq u(t) \leq u_{\max}$, $c(u(t)) \geq 0$, and some constraints related to rate of movement that are dependent on the input parametrization.

Path constraints

$$\eta(x(t),u(t)) \geq 0 \tag{2}$$

may render the optimization problem infeasible. For this reason and due to computational efficiency considerations when computing the sensitivities by the adjoint method (Capolei and Jørgensen, 2012; Jørgensen, 2007), we include these constraints as soft constraints using the following smooth approximation

$$\chi_i(x(t),u(t)) = \frac{1}{2}\left(\sqrt{\eta_i(x(t),u(t)^2 + {\beta_i}^2)} - \eta_i(x(t),u(t)\right) \tag{3}$$

to the exact penalty function $\max(0, -\eta_i(x(t)))$ for $i \in \{1,\ldots,n_\eta\}$.With this approximation of the path constraints, the resulting stage cost, $\Phi(x(t),u(t))$, used in (1a) consist of the inherent stage cost, $\tilde{\Phi}(x(t),u(t))$, and terms penalizing violation of the path constraints (2)

$$\Phi(x,u) = \tilde{\Phi}(x,u) + \|\chi(x,u)\|_{1,Q_1} + \frac{1}{2}\|\chi(x,u)\|^2_{2,Q_2} \tag{4}$$

2.1 Discretization

Control Parametrization Let T_s denote the sample time such that an equidistant mesh can be defined as

$$t_a = t_0 < \ldots < t_S < \ldots < t_N = t_b \tag{5}$$

with $t_j = t_a + jT_s$ for $j = 0,1,\ldots,N$. We use a piecewise constant representation of the control function on this equidistant mesh, i.e. we approximate the control vector on every subinterval $[t_j, t_{j+1}]$ by the zero-order-hold parametrization

$$u(t) = u_j,\ u_j \in \mathbb{R}^{n_u},\ t_j \leqslant t < t_{j+1},\ j \in 0,\ldots,N-1 \tag{6}$$

Input Constraints The input constraints (1d) include bound constraints $u_{\min} \leq u_k \leq u_{\max}$. In the discrete problem using the zero-order-hold parametrization, we also include rate of movement constraints in the form $\Delta u_{\min} \leq \Delta u_k \leq \Delta u_{\max}$ with $\Delta u_k = u_k - u_{k-1}$.

2.2 Single Shooting Optimization

For the single shooting approach (control vector parametrization), we introduce the function

$$\begin{aligned}
&\psi(\{u_k\}_{k=0}^{N-1}, x_0) = \\
&\quad\Big\{ J = \int_{t_a}^{t_b} \Phi(x(t),u(t))dt + \hat{\Phi}(x(t_b)) : \\
&\qquad x(t_0) = x_0, \\
&\qquad \frac{d}{dt}g(x(t)) = f(x(t),u(t)),\ t_a \leq t \leq t_b, \\
&\qquad u(t) = u_k,\ t_k \leq t < t_{k+1},\ k = 0,1,\ldots,N-1 \Big\}
\end{aligned} \tag{7}$$

such that (1) can be approximated with the finite dimensional constrained optimization problem

$$\min_{\{u_k\}_{k=0}^{N-1}} \quad \psi = \psi(\{u_k\}_{k=0}^{N-1}, x_0) \tag{8a}$$

$$s.t. \quad u_{\min} \leq u_k \leq u_{\max} \qquad k \in \mathcal{N} \tag{8b}$$

$$\Delta u_{\min} \leq \Delta u_k \leq \Delta u_{\max} \qquad k \in \mathcal{N} \tag{8c}$$

$$c_k(u_k) \geq 0 \qquad k \in \mathcal{N} \tag{8d}$$

with $\mathcal{N} = \{0,1,\ldots,N-1\}$.

3. ESDIRK METHODS

In this section, we describe our implementation of the ESDIRK method for the computation of $\psi(\{u_k\}_{k=0}^{N-1}, x_0)$ in (7). Computation of $\psi(\{u_k\}_{k=0}^{N-1}, x_0)$ consists of two major operations: 1) For each integration step we first compute the model states $x(t)$ solving the initial value problem (1c), 2) and then we compute, using the same quadrature points, the value of the Lagrange term

$$\bar{\psi}(t) := \int_{t_a}^{t} \Phi(x(t), u(t))dt \qquad t_a \leq t \leq t_b. \tag{9}$$

in the cost function (1a). Let $\tilde{t}_n$ denote the integration times chosen by the step size controller in the integrator. Each integration step size, h_n, is chosen such that it is smaller than or equal to the sample time, T_s. Therefore, one sample interval contains many integration steps. The numerical solution of the IVP (1c) by an s-stage, stiffly accurate, Runge-Kutta ESDIRK method with an embedded error estimator, may in each integration step $[\tilde{t}_n, \tilde{t}_{n+1}]$ be denoted (Capolei and Jørgensen, 2012; Völcker et al., 2010a)

$$T_1 = \tilde{t}_n, \quad T_i = \tilde{t}_n + c_i h_n \tag{10a}$$
$$X_1 = x_n \tag{10b}$$
$$\phi_i(\{X_j\}_{j=1}^{i-1}, u) = g(X_1) + h_n \sum_{j=1}^{i-1} a_{ij} f(X_j, u) \tag{10c}$$
$$g(X_i) = \phi_i(\{X_j\}_{j=1}^{i-1}, u) + h_n \gamma f(X_i, u) \tag{10d}$$
$$x_{n+1} = X_s \tag{10e}$$
$$e_{n+1} = h_n \sum_{j=1}^{s} d_i f(X_j, u) \tag{10f}$$

with $i = 2, \ldots, s$. X_i denotes the numerical solution at time T_i for $i \in \{1, \ldots, s\}$. x_{n+1} is the numerical solution at time $\tilde{t}_{n+1} = \tilde{t}_n + h_n$. e_{n+1} is the estimated error of the numerical solution, i.e. $\|e_{n+1}\| \approx \|g(x_{n+1}) - g(x(\tilde{t}_{n+1}))\|$.

Subsequent to solution of (10), we compute the numerical solution of the cost function (9)

$$\bar{\psi}(\tilde{t}_{n+1}) = \bar{\psi}(\tilde{t}_n) + h_n \sum_{i=1}^{s} b_i \Phi(X_i, u) \tag{11}$$

When $\tilde{t}_{n+1} = t_b$, we add the Mayer term of (1a) such that

$$\psi(\{u_k\}_{k=0}^{N-1}, x_0) = \psi(t_b) = \bar{\psi}(t_b) + \hat{\Phi}(x(t_b)) \tag{12}$$

The main computational effort in the ESDIRK method is solution of the implicit equations (10d) using a Newton based method. (10d) is solved by sequential solution of

$$R_i(X_i) := [g(X_i) - h_n \gamma f(X_i, u)] - \phi_i(\{X_j\}_{j=1}^{i-1}, u) = 0 \tag{13}$$

for $i = 2, \ldots, s$. (13) is solved using an inexact Newton method. Each iteration in the inexact Newton method for solution of (13) may be denoted

$$M \Delta X_i^{[l]} = -R_i(X_i^{[l]}) \tag{14a}$$
$$X_i^{[l+1]} = X_i^{[l]} + \Delta X_i^{[l]} \tag{14b}$$

The iteration matrix, M, is an approximation

$$M \approx J(X_i^{[l]}) \tag{15}$$

to the Jacobian of the residual function

$$J_i(X_i) = \frac{\partial R_i}{\partial X_i}(X_i) = \frac{\partial g}{\partial x}(X_i) - h_n \gamma \frac{\partial f}{\partial x}(X_i, u) \tag{16}$$

The iteration matrix, M, and its LU factorization is updated adaptively by monitoring the convergence rate of the inexact Newton iterations. Convergence of the inexact Newton iteration is measured by

$$\|R_i(X_i^{[l]})\| = \max_{j \in 1, \ldots, n_x} \frac{|(R_i(X_i^{[l]})_j|}{\max\{\text{atol}_j, \text{rtol}_j g_j(X_i^{[l]})\}} < \tau \tag{17}$$

where atol is the absolute tolerance and rtol is the relative tolerance. Steps are accepted if this measure of the residual is smaller than $\tau \approx 0.1$ In case of divergence or slow convergence, the iterations are terminated, the step size, h_n, is decreased and the Jacobian of the iteration matrix is re-evaluated and factorized. As explained in e.g. Völcker et al. (2010b) and Capolei and Jørgensen (2012), the step size controller adjust the temporal step sizes such that the error estimate satisfies a norm similar to the norm used in (17).

4. CONTINUOUS ADJOINT METHOD

Gradient based methods such as sequential quadratic programming (SQP) methods for solution of (8) require the gradient of the objective function (7) with respect to the control vector parameters, i.e. $\partial\psi/\partial u_k$ for $k = 0, 1, \ldots, N-1$. In this section, we describe a continuous-time adjoint based method for computation of these gradients.

Proposition 1. (Gradients based on Continuous Adjoints). Consider the function $\psi = \psi(\{u_k\}_{k=0}^{N-1}; x_0)$ defined by (7).

The gradients, $\partial\psi/\partial u_k$, may be computed as

$$\frac{\partial \psi}{\partial u_k} = \int_{t_k}^{t_{k+1}} \left(\frac{\partial \Phi}{\partial u} - \lambda^T \frac{\partial f}{\partial u} \right) dt \qquad k = 0, 1, \ldots, N-1 \tag{18}$$

in which $x(t)$ is computed by solution of (1b)-(1c) and $\lambda(t)$ is computed by solution of the adjoint equations

$$\frac{d\lambda^T}{dt}\frac{\partial g}{\partial x} + \lambda^T \frac{\partial f}{\partial x} - \frac{\partial \Phi}{\partial x} = 0 \tag{19a}$$
$$\frac{\partial \hat{\Phi}}{\partial x}(x(t_b)) + \lambda^T(t_b)\frac{\partial g}{\partial x}(x(t_b)) = 0 \tag{19b}$$

Proof. See Appendix A.

Remark 2. (Computation using ESDIRK). $x(t)$ is computed using the ESDIRK method applied to (1b)-(1c) and integration forwards. This solution is stored. The same ESDIRK method is applied for computation of $\lambda(t)$ by solving (19) integrating backwards in time.

Remark 3. (Gradients Computed by Continuous Adjoints). The gradients computed using the continuous adjoints are not the exact gradients, $\partial\psi/\partial u_k$, when the involved differential equations and integrals are computed by discretization using the ESDIRK method. However, they can be made sufficiently precise for the optimizer such that they do not affect the convergence (Nadarajah and Jameson, 2007). The advantage of the continuous adjoint equations (19) is that they can be solved faster than the adjoint equations for the discretized system (10)-(12).

5. PRODUCTION OPTIMIZATION FOR A CONVENTIONAL OIL FIELD

In this section, we apply our algorithm for constrained optimal control problems to production optimization in a

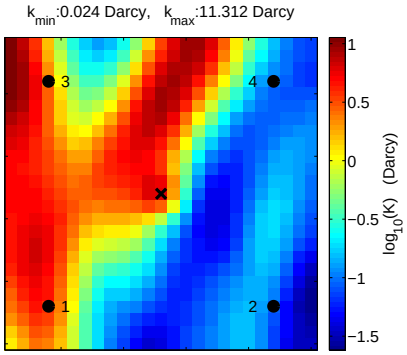


Fig. 1. The permeability field and the location of wells. A circle indicates the location of an injector and a cross indicates the location of a producer.

Table 1. Parameters for the two phase model and the discounted state cost function (20).

Symbol	Description	Value	Unit
ϕ	Porosity	0.2	-
c_r	Rock compressibility	0	Pa^{-1}
ρ_o	Oil density (400 atm)	800	kg/m^3
ρ_w	Water density (400 atm)	1000	kg/m^3
c_o	Oil compressibility	10^{-5}	1/atm
c_w	Water compressibility	10^{-5}	1/atm
μ_o	Dynamic oil viscosity	$2 \cdot 10^{-3}$	$\text{Pa} \cdot \text{s}$
μ_w	Dynamic water viscosity	$1 \cdot 10^{-3}$	$\text{Pa} \cdot \text{s}$
S_{or}	Residual oil saturation	0.1	-
S_{ow}	Connate water saturation	0.1	-
n_o	Corey exponent for oil	1.5	-
n_w	Corey exponent for water	1.4	-
P_{init}	Initial reservoir pressure	400	atm
S_{init}	Initial water saturation	0.1	-
r_o	Oil price	100	USD/m^3
r_w	Water production cost	20	USD/m^3

conventional horizontal oil field that can be modeled as two phase flow in a porous medium (Chen, 2007; Völcker et al., 2009). The reservoir size is 450 m × 450 m × 10 m. By spatial discretization this reservoir is divided into $25 \times 25 \times 1$ grid blocks. The configuration of injection wells and producers as well as the permeability field is illustrated in Fig. 1. As indicated in Fig. 1, the four injectors are located in the corners of the field, while the single producer is located in the center of the field. The specification of the two phase oil model consists of the injector ($i \in \mathcal{I}$) and the producer ($i \in \mathcal{P}$) location, the permeability parameters indicated in Fig. 1, and the parameters listed in Table 1. The initial reservoir pressure is 400 atm everywhere in the reservoir. The initial water saturation is 0.1 everywhere in the reservoir. This implies that initially the reservoir has a uniform oil saturation of 0.9.

The inherent discounted stage cost function (see (4))

$$\begin{aligned}\tilde{\Phi}(t) &= \tilde{\Phi}(x(t), u(t)) \\ &= -\frac{1}{(1+b)^{t/365}} \sum_{j \in \mathcal{P}} \left(r_o(1-f_w) - f_w r_w\right) q_j(t)\end{aligned} \tag{20}$$

accounts for the value of the oil produced minus the processing cost of the produced water. In this cost function, we have neglected the processing cost of injected water as well as the effect of pressure on injecting water. b is the discount factor. The fractional flow of water, $f_w = \lambda_w/(\lambda_w + \lambda_o)$, indicates the relative flow of water. $\lambda_w = \rho_w k k_{rw}/\mu_w$ and $\lambda_o = \rho_o k k_{ro}/\mu_o$ are the water and oil mobilities, respectively. In the problems considered, we do not have any cost-to-go terms, i.e. $\hat{\Phi}(t_b) = 0$. Neither do we have any path constraints (2). Therefore, maximizing the net present value of the oil field corresponds to minimization of

$$J(t_b) = -\text{NPV}(t_b) = \int_{t_a}^{t_b} \Phi(x(t), u(t)) dt \tag{21}$$

with $\Phi(x(t), u(t)) = \tilde{\Phi}(x(t), u(t))$. The optimizer maximizes the net present value by manipulating the injection of water at the injectors and by manipulation of the total fluid production (oil and water) at the producers. Hence, the manipulated variable at time period $k \in \mathcal{N}$ is $u_k = \{\{q_{w,i,k}\}_{i \in \mathcal{I}}, \{q_{i,k}\}_{i \in \mathcal{P}}\}$ with $\mathcal{I}$ being the set of injectors and $\mathcal{P}$ being the set of producers. For $i \in \mathcal{I}$, $q_{w,i,k}$ is the injection rate (m^3/day) of water in time period $k \in \mathcal{N}$ at injector i. For $i \in \mathcal{P}$, $q_{i,k}$ is the total flow rate (m^3/day) at producer i in time period $k \in \mathcal{N}$. Therefore, at producer $i \in \mathcal{P}$, the water flow rate is $q_{w,i,k} = f_w q_{i,k}$ and the oil flow rate is $q_{o,i,k} = (1 - f_w) q_{i,k}$.

The bound constraints (8b) appear in the production optimization problem because the water injected at injectors and the production at the producers must both be positive and because each production facility has a maximum flow capacity. In the considered problem we have

$$0 \le q_{w,i,k} \le q_{\max} \qquad i \in \mathcal{I},\ k \in \mathcal{N} \tag{22a}$$
$$0 \le q_{i,k} \le q_{\max} \qquad i \in \mathcal{P},\ k \in \mathcal{N} \tag{22b}$$

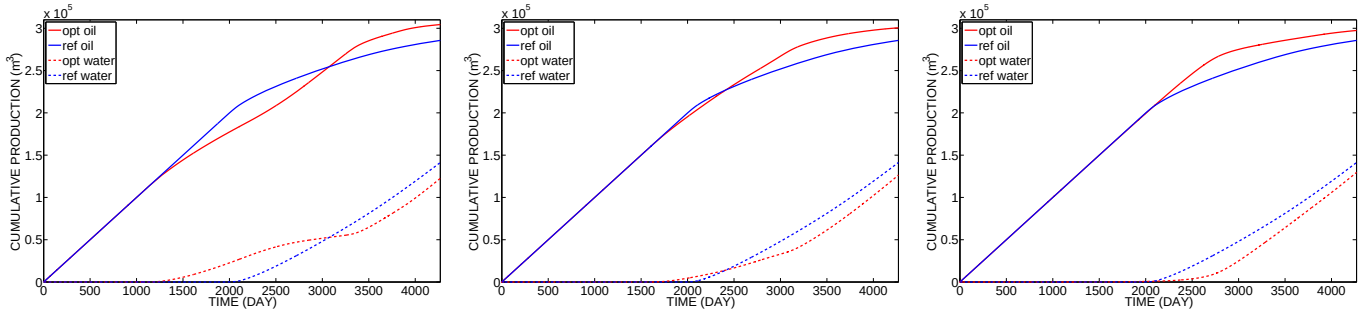
The maximum flow capacity, $q_{\max}$, is the same for all injectors and producers in this case study. The rate of change for all injectors and producers are $|q_{i,k} - q_{i,k-1}| \le 5$ for $i \in \mathcal{I} \cup \mathcal{P}$ and $k \in \mathcal{N}$. Since the injection of oil is zero, $q_{o,i,k} = 0$ for $i \in \mathcal{I}$, we get $|q_{w,i,k} - q_{w,i,k-1}| \le 5$ for $i \in \mathcal{I}$ and $k \in \mathcal{N}$. This leads to the rate of movement constraints (8c). In addition we use a voidage replacement constraint (Brouwer and Jansen, 2004; Jansen, 2011)

$$\sum_{i \in \mathcal{I}} q_{i,k} = \sum_{i \in \mathcal{I}} q_{w,i,k} = \sum_{i \in \mathcal{P}} q_{i,k} \qquad k \in \mathcal{N} \tag{23}$$

and enforce a constant total injection, $\sum_{i \in \mathcal{I}} q_{w,i,k} = q_{\max}$ for $k \in \mathcal{N}$. This translates into constraints of the type (8d). By the total injection constraint, the optimization problem reduces to a problem of redistributing the flows among the injectors.

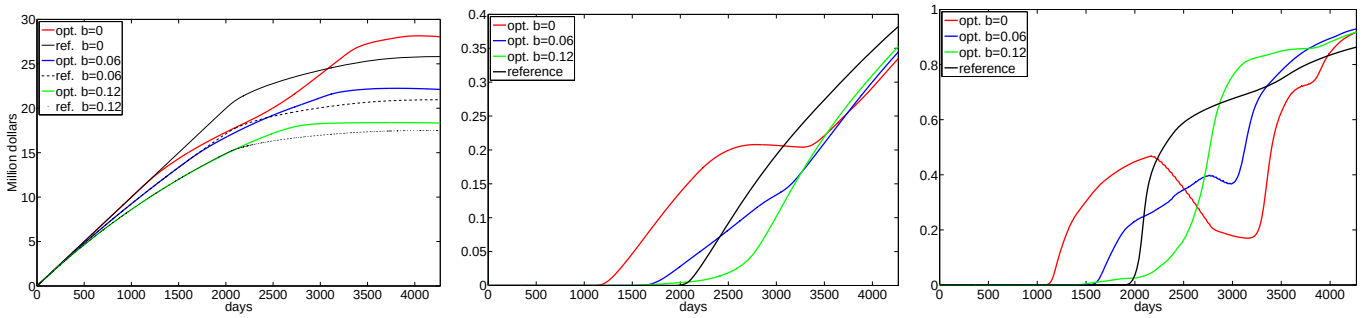
The prediction and control horizon is $t_b = 4270$ days and the sampling period is $T_s = 35$. Hence the prediction and control horizon corresponds to $N = 122$ periods. With a total injection at each time period of $q_{\max} = 100$ m^3/day, these specifications corresponds to injection of 1.05 pore volume during operation of the reservoir. The prediction horizon is optimal in the reference case for a total injection of 100 m^3/day.

The optimal water injection rates computed by solution of the constrained optimal control problem (1) for different discount factors, b, are illustrated in Fig. 2. In addition, a base case with constant and equal water injection rates is illustrated. It is evident that the optimal injection rates are very sensitive to the discount factor, b. The corresponding cumulative oil and water production are plotted in Fig. 3. Independent of the discount factor value, the optimized strategies produce more oil than the base case. For the high discount factor case, $b = 0.12$, less oil is recovered than in



(a) Discount factor $b = 0$. (b) Discount factor $b = 0.06$. (c) Discount factor $b = 0.12$.

Fig. 3. Cumulative oil and water productions for different discount factors, b.



(a) NPV. (b) Water cut. (c) Water fraction (kg water/kg fluid).

Fig. 4. The net present value (NPV), water cut (accumulated water production per produced fluid), and the water fraction as function of time for the scenarios considered.

Table 2. Key indicators for the optimized cases. Improvements are compared to the base case.

b	NPV 10^6 USD	ΔNPV %	Cum. Oil 10^5 m^3	ΔOil %	Cum. water 10^5 m^3	ΔWater %	Oil Rec. factor %	ΔOil Rec. factor %-point
0	28.0	+8.7	3.05	+6.5	0.122	−13.2	83.7	+5.2
0.06	22.1	+5.6	3.01	+5.2	0.126	−10.5	82.6	+4.1
0.12	18.3	+4.8	2.98	+4.1	0.129	−8.2	81.7	+3.2

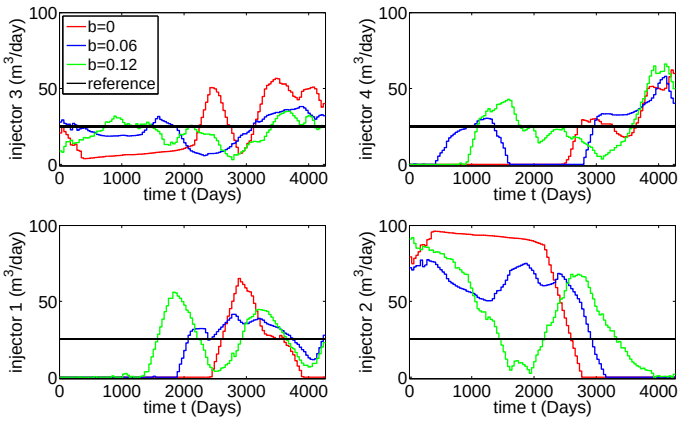


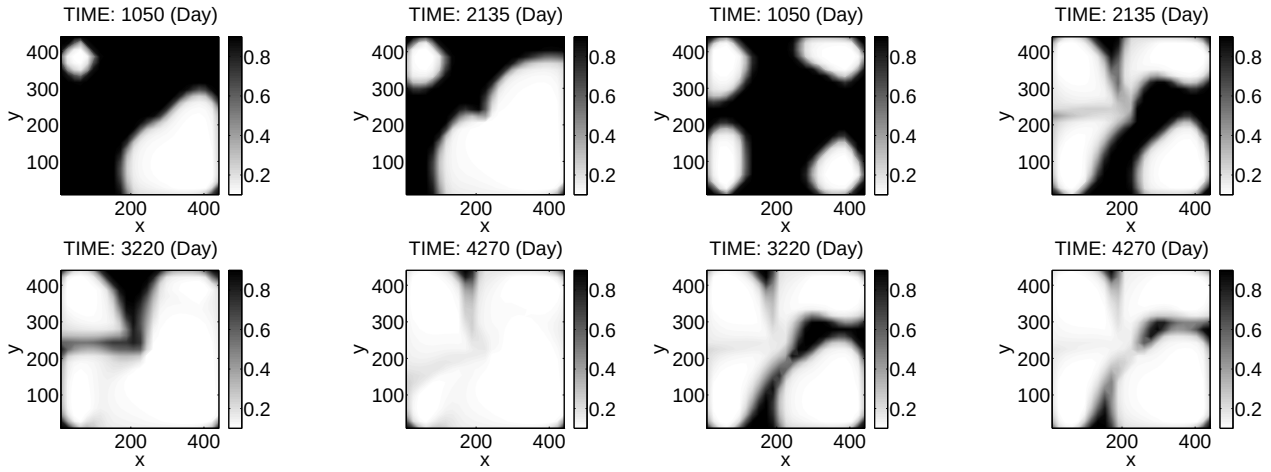
Fig. 2. Optimal water injection rates for different discount factors, b.

the low discount factor case, $b = 0$. However, the produced oil is always above the reference case when $b = 0.12$. This is not the case for $b = 0$ and $b = 0.06$. Fig. 4 illustrates the net present value, the water cut and the water fraction for the base case scenarios as well as the optimized scenarios. The plot of NPV demonstrates that when $b = 0$, the NPV is lower than the base case NPV at some time during the production. At the end of the production the optimized NPV is largest. In order to recover the maximum amount of oil less oil must be produced at some times. This is also confirmed by the water fraction curves. The results are summarized in Table 2. Table 2 shows that most oil is recovered in the case without discount ($b = 0$), while least oil is recovered when the discount factor is high ($b = 0.12$).

Fig. 5 illustrates the evolution of the oil saturation for the optimized case ($b = 0$) and the base case. The figures show that initially, less oil is produced from the upper left corner in the optimal case compared to the base case. This gives a better sweep of the oil field and results ultimately in higher oil recovery.

6. CONCLUSIONS

In this paper, we solve constrained optimal control problems using a single shooting method based on a quasi-Newton implementation of Powell's sequential quadratic programming (SQP) algorithm. The system of differential equations are formulated in a novel way to ensure mass conservation and the resulting initial value problem (1c) is solved with tailored ESDIRK integration methods. We also introduce a high order continuous adjoint system for efficient computation of the gradients. The algorithm is implemented in Matlab.



(a) Optimal solution ($b = 0$). (b) Reference solution.

Fig. 5. Oil saturations at different times for the optimal solution and the reference solution.

The resulting algorithm is tested on a production optimization problem for an oil reservoir with two phase flow. For all cases considered, the dynamic optimization increase the net present value of the oil field and give increased oil production. However, the optimal injection rates are very sensitive to the discount factor.

REFERENCES

Binder, T., Blank, L., Bock, H.G., Burlisch, R., Dahmen, W., Diehl, M., Kronseder, T., Marquardt, W., Schlöder, J.P., and von Stryk, O. (2001). Introduction to model based optimization of chemical processes on moving horizons. In M. Grötschel, S. Krumke, and J. Rambau (eds.), *Online Optimization of Large Scale Systems*. Springer.

Brouwer, D.R. and Jansen, J.D. (2004). Dynamic optimization of waterflooding with smart wells using optimal control theory. *SPE Journal*, 9(4), 391–402.

Cao, Y., Li, S., Petzold, L., and Serban, R. (2002). Adjoint sensitivity analysis for differential-algebraic equations: The adjoint DAE system and its numerical solution. *SIAM Journal on Scientific Computing*, 24(3), 1076–1089.

Capolei, A. and Jørgensen, J.B. (2012). Solution of constrained optimal control problems using multiple shooting and ESDIRK methods. In *2012 American Control Conference*. Submitted.

Chen, Z. (2007). *Reservoir Simulation. Mathematical Techniques in Oil Recovery*. SIAM, Philadelphia, USA.

Jansen, J.D., Douma, S.D., Brouwer, D.R., Van den Hof, P.M.J., Bosgra, O.H., and Heemink, A.W. (2009). Closed-loop reservoir management. In *2009 SPE Reservoir Simulation Symposium*, SPE 119098. The Woodlands, Texas, USA.

Jansen, J.D., Bosgra, O.H., and Van den Hof, P.M.J. (2008). Model-based control of multiphase flow in subsurface oil reservoirs. *Journal of Process Control*, 18, 846–855.

Jansen, J. (2011). Adjoint-based optimization of multiphase flow through porous media - A review. *Computers & Fluids*, 46, 40–51.

Jørgensen, J.B. (2007). Adjoint sensitivity results for predictive control, state- and parameter-estimation with nonlinear models. In *Proceedings of the European Control Conference 2007*, 3649–3656. Kos, Greece.

Kourounis, D., Voskov, D., and Aziz, K. (2010). Adjoint methods for multicomponent flow simulation. In *12th European Conference on the Mathematics of Oil Recovery*. Oxford, UK.

Nadarajah, S.K. and Jameson, A. (2007). Optimum shape design for unsteady flows with time-accurate continuous and discrete adjoint methods. *AIAA Journal*, 45(7), 1478–1491.

Sarma, P., Aziz, K., and Durlofsky, L.J. (2005). Implementation of adjoint solution for optimal control of smart wells. In *SPE Reservoir Simulation Symposium, 31 January-2 Feburary 2005, The Woodlands, Texas*.

Suwartadi, E., Krogstad, S., and Foss, B. (2011). Nonlinear output constraints handling for production optimization of oil reservoirs. *Computational Geosciences*. doi: 10.1007/s10596-011-9253-3. Accepted.

Thomas, S. (2008). Enhanced oil recovery - an overview. *Oil & Gas Science and Technology*, 63, 9–19.

Völcker, C., Jørgensen, J.B., Thomsen, P.G., and Stenby, E.H. (2010a). Explicit singly diagonally implicit Runge-Kutta methods and adaptive stepsize control for reservoir simulation. In *ECMOR XII - 12th European Conference on the Mathematics of Oil Recovery*. Oxford, UK.

Völcker, C., Jørgensen, J.B., Thomsen, P.G., and Stenby, E.H. (2010b). Adaptive stepsize control in implicit Runge-Kutta methods for reservoir simulation. In M. Kothare, M. Tade, A.V. Wouwer, and I. Smets (eds.), *Proceedings of the 9th International Symposium on Dynamics and Control of Process Systems (DYCOPS 2010)*, 509–514. Leuven, Belgium.

Völcker, C., Jørgensen, J.B., and Stenby, E.H. (2011). Oil reservoir production optimization using optimal control. In *50th IEEE Conference on Decision and Control and European Control Conference*. Orlando, Florida. Accepted.

Völcker, C., Jørgensen, J.B., Thomsen, P.G., and Stenby, E.H. (2009). Simulation of subsurface two-phase flow in an oil reservoir. In *Proceedings of the European Control Conference 2009*, 1221–1226. Budapest, Hungary.

Appendix A. PROOF OF PROPOSITION 1.

The idea in the proof stems from Cao et al. (2002). Define

$$\begin{aligned} G(x,\dot{x},u) &= \frac{d}{dt}g(x(t)) - f(x(t),u(t)) \\ &= \frac{\partial g}{\partial x}\dot{x}(t) - f(x(t),u(t)) = 0 \end{aligned} \tag{A.1}$$

and introducing the Lagrange multiplier, $\lambda(t)$, to define the augmented objective function as

$$J_A = J + \int_{t_a}^{t_b} \lambda^T(t)\, G(x,\dot{x},u)dt \tag{A.2a}$$

(A.1) implies that the derivative of the augmented objective function, J_A, can be expressed as

$$\begin{aligned} \frac{dJ_A}{du_k} = \frac{dJ}{du_k} &= \frac{\partial\hat{\Phi}}{\partial x}\frac{\partial x}{\partial u_k}\bigg|_{t_b} \\ &+ \int_{t_a}^{t_b}\left(\frac{\partial\Phi}{\partial u}\frac{\partial u(t)}{\partial u_k} + \frac{\partial\Phi}{\partial x}\frac{\partial x(t)}{\partial u_k}\right)dt \\ &+ \int_{t_a}^{t_b}\lambda^T(t)\left(\frac{\partial G}{\partial u}\frac{\partial u(t)}{\partial u_k} + \frac{\partial G}{\partial x}\frac{\partial x(t)}{\partial u_k} + \frac{\partial G}{\partial\dot{x}}\frac{\partial\dot{x}}{\partial u_k}\right)dt \end{aligned} \tag{A.3}$$

where

$$\frac{\partial u(t)}{\partial u_k} = \begin{cases} I & t_k \le t < t_{k+1} \\ 0 & otherwise \end{cases} \tag{A.4}$$

Integrating by part

$$\begin{aligned} \int_{t_a}^{t_b}\lambda^T(t)\frac{\partial G}{\partial\dot{x}}\frac{\partial\dot{x}}{\partial u_k}dt &= \left[\lambda^T\frac{\partial G}{\partial\dot{x}}\frac{\partial x(t)}{\partial u_k}\right]\bigg|_{t_a}^{t_b} \\ &- \int_{t_a}^{t_b}\frac{d}{dt}\left(\lambda^T\frac{\partial G}{\partial\dot{x}}\right)\frac{\partial x}{\partial u_k}dt \end{aligned} \tag{A.5}$$

and using $\frac{\partial x}{\partial u_k}(t_a) = 0$ in our case, we can rearrange equation (A.3) as

$$\begin{aligned} \frac{dJ}{du_k} &= \int_{t_k}^{t_{k+1}}\left(\frac{\partial\Phi}{\partial u} + \lambda^T\frac{\partial G}{\partial u}\right)dt \\ &+ \int_{t_k}^{t_{k+1}}\left(\frac{\partial\Phi}{\partial x} + \lambda^T\frac{\partial G}{\partial x} - \frac{d}{dt}\left(\lambda^T\frac{\partial G}{\partial\dot{x}}\right)\right)\frac{\partial x}{\partial u_k}dt \\ &+ \left[\left(\frac{\partial\hat{\Phi}}{\partial x} + \lambda^T\frac{\partial G}{\partial\dot{x}}\right)\frac{\partial x}{\partial u_k}\right]\bigg|_{t_b} \end{aligned} \tag{A.6}$$

This expression gives the derivative dJ/du_k for any value (not just the optimal one) of $\lambda(t)$. We choose $\lambda(t)$ such that it satisfies

$$\frac{\partial\Phi}{\partial x} + \lambda^T\frac{\partial G}{\partial x} - \frac{d}{dt}\left(\lambda^T\frac{\partial G}{\partial\dot{x}}\right) = 0 \tag{A.7a}$$

$$\left[\frac{\partial\hat{\Phi}}{\partial x} + \lambda^T\frac{\partial G}{\partial\dot{x}}\right]\bigg|_{t_b} = 0 \tag{A.7b}$$

and gives a simple expression for evaluation of dJ/du_k

$$\frac{\partial\psi}{\partial u_k} = \frac{dJ}{du_k} = \int_{t_k}^{t_{k+1}}\left(\frac{\partial\Phi}{\partial u} - \lambda^T\frac{\partial f}{\partial u}\right)dt \tag{A.8}$$

(A.1) implies

$$\frac{\partial G}{\partial x} = \frac{d}{dt}\left(\frac{\partial g}{\partial x}\right) - \frac{\partial f}{\partial x} \tag{A.9a}$$

$$\frac{\partial G}{\partial\dot{x}} = \frac{\partial g}{\partial x} \tag{A.9b}$$

such that (A.7) can be rearranged to

$$\frac{d\lambda^T}{dt}\frac{\partial g}{\partial x} + \lambda^T\frac{\partial f}{\partial x} - \frac{\partial\Phi}{\partial x} = 0 \tag{A.10a}$$

$$\frac{\partial\hat{\Phi}}{\partial x}(x(t_b)) + \lambda^T(t_b)\frac{\partial g}{\partial x}(x(t_b)) = 0 \tag{A.10b}$$

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Optimization of lift gas allocation in a gas lifted oil field as non-linear optimization problem

Roshan Sharma* Kjetil Fjalestad**
Bjørn Glemmestad*

**Department of Electrical Engineering, Information Technology and Cybernetics, Telemark University College Porsgrunn, Norway (e-mail: roshan.sharma,bjorn.glemmestad@hit.no).*
***Statoil Research Center, Porsgrunn, Norway*
(e-mail: kfja@statoil.com)

Abstract: Proper allocation and distribution lift gas is necessary for maximizing total oil production from a field with gas lifted oil wells. When the supply of the lift gas is limited, the total available gas should be optimally distributed among the oil wells of the field such that the production of oil from the field is maximized. This paper describes a non-linear optimization problem with constraints associated with the optimal distribution of the lift gas. A non-linear objective function is developed using a simple dynamic model of the oil field where the decision variables represents the lift gas flow rate set points of each oil wells of the field. The lift gas optimization problem is solved using the 'fmincon' solver found in MATLAB. For verification, a modified hill climbing method has also been utilized for solving the optimization problem. Comparison and discussion of the simulation results from both the methods of solving the optimization problem is also included in the paper. Using both of these methods, it has been shown that after optimization the total oil production is increased by about 5%. The paper also discusses on the possibility that one single optimization is sufficient to operate the oil field in an optimal state when the total supply of lift gas is taken as disturbance. Simulation results show that repeated optimization performed after the first time optimization has no effect in the total oil production.

Keywords: Optimization, non-linear programming, objective function, gas lifted oil well, fmincon, hill climbing, constraints.

1. INTRODUCTION

For a gas lifted oil field where multiple oil wells share the lift gas supplied by the common source (see Figure 3), proper distribution of the available gas is an important issue for maximizing the total oil produced from the oil wells. The set points for the mass flow rate of the lift gas for each of the gas lift choke valves have to be allocated in a way that the distribution yields maximum oil production using the available gas. Stable or steady operation does not guarantee optimal operation i.e. even though the field is operating in a stable manner, the lift gas might have been distributed among the oil wells in a non-optimal way and the wells might have been producing less than what they can actually produce.

For the Norne oil field with five gas lifted oil wells, the objective is to distribute the available gas ensuring optimal production of oil. The amount of lift gas available is assumed to be limited. *Thus, optimization for the oil field for this case is the task of finding out the optimal set points of the five flow controllers*. Each flow controller controls the opening and closing of the gas lift choke valve of each well. In other words, the decision variables for the optimization problem are the nominal set points of the flow controllers. Using these optimal set points for the mass flow rate of the lift gas, the total oil produced from the five oil wells should be maximized using the limited amount of gas available.

A non-linear gas lift optimization problem has been formulated using a simple model of the oil field. The main objective of the paper is to solve this non-linear optimization problem using the *'fmincon'* solver of the MATLAB Optimization toolbox as described in Section 4. In addition, the optimization problem has also been solved using modified hill climbing method as described in Section 5. The focus of this paper is to show through simulation results that both the optimization methods cause an increase in the total oil production from the field. A brief explanation of the model of the oil field is given in Section 2. Section 3 contains details about the development of the non-linear objective function with constraints. Comparison of the two optimization methods is described in Section 6.

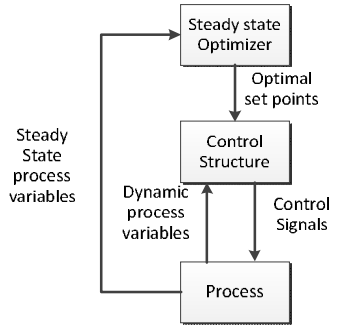


Figure 1: Optimization hierarchy

The optimization procedure works along with the control structure used in the oil field for the control of lift gas distribution and oil production. The optimizer loop rests on top of the control loop and provides the optimal set points to the control loop as shown in Figure 1.

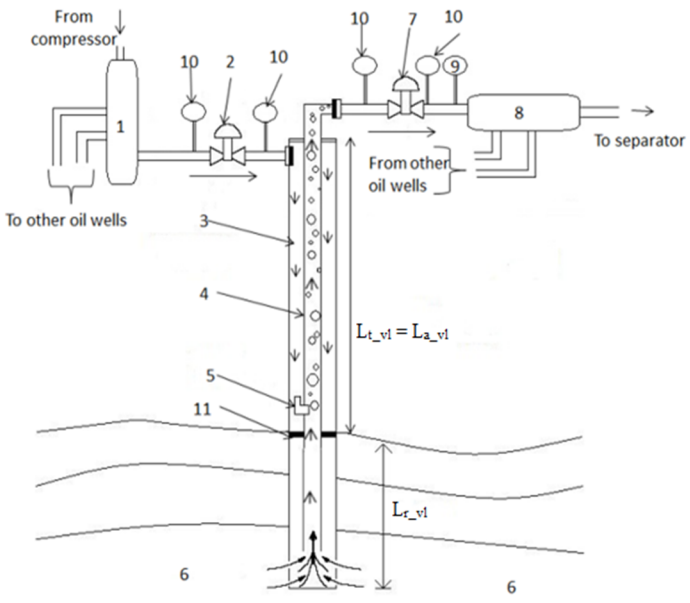
Out of the four control structures applied to the oil field as proposed by Sharma et al. (2011), optimization with the cascade control structure is discussed in this paper.

Optimization of gas lifted oil field has been a research of interest to many authors. Penalty function or Sequential Unconstrained Minimization Technique (SUMT) which can accommodate both the equality constraints and inequality constraints needed to solve the non-linear optimization model of the gas allocation to a gas lifted oil field was proposed by Zhong et al., (2004). Daily well scheduling in gas lifted petroleum fields has been formulated and solved by using mixed integer nonlinear (MINLP) model (Kosmidis et al., 2005) where the discrete decisions include the operational status of wells, the allocation of wells to manifolds or separators and the allocation of flow lines to separators, and the continuous decision include the well oil rates and the allocation of gas-to-gas lift wells. Dynamic programming has been be used for solving a gas-lift optimization problem (Camponogara and Nakashima, 2006) where the gas-lift optimization problem can be casted as a mixed integer nonlinear programming problem whose integer variables decide which oil well should produce, while the continuous variables allocate the gas-compressing capacity to the active ones. Computational scheme using genetic algorithm has been used to find optimum gas injection rate (Saepudin et al., 2007; Ray and Sarker, 2007) for gas lifted oil filed and also for dual gas lift system (Sukarno et al., 2009). For gas lift optimization, a high dimensional problem has been reduced into one single variable problem by using Newton reduction method based on upper convex profile (Rashid, 2010). Gas-lift optimization has been formulated and solved by using an objective function considering the annualized capital costs on compressor, turbine and gas pipelines, the operating costs related to fuel and the revenue from produced oil (Souza et al., 2010).

2. MODEL OF THE OIL FIELD

A simple model of a gas lifted oil well where all the necessary and important components are taken into account is shown in Figure 2. The oil field consists of five gas lifted oil wells which share the lift gas from the common distribution manifold. Figure 3 shows a schematic of the oil field with five oil wells. In this paper, the details about the development of the model of the oil field have not been discussed and only the description of the oil field is included. The model proposed by Sharma et al. (2011) has been used here in this paper.

The compressor outputs a highly pressurized lift gas into the gas distribution pipeline. The lift gas enters into the annulus of each well from this common gas distribution manifold. The amount of lift gas to be injected into each well is controlled by the gas lift choke valve present in the well head of each oil well.



1.Lift gas distribution manifold 2. Gas lift choke valves 3.Annulus 4. Tubing 5. Gas injection valve 6. Reservoir 7. Production choke valve 8. Gathering manifold 9. Multiphase meters 10. Pressure and temperature transducer 11. Packer

$L_{t\_vl} = L_{a\_vl}$ = vertical length of tubing/annulus above the gas injection point

L_{r_vl} = vertical length of tubing below the gas injection point

Figure 2: Different components of a gas lifted oil well

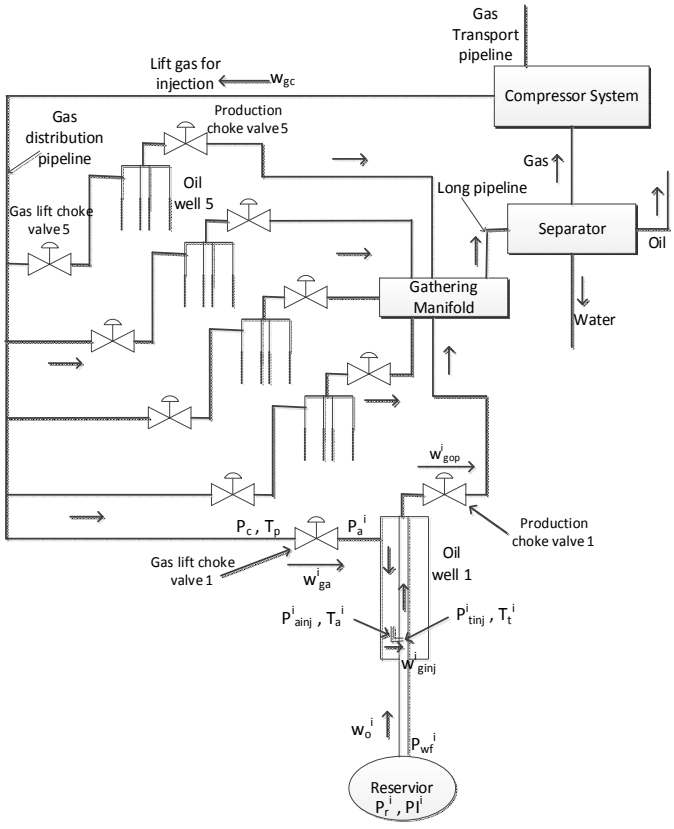


Figure 3: Schematic of an oil field with five oil wells

From the annulus, the high pressure lift gas is injected into the tubing at a proper depth through the gas injection valve (see Figure 2). The gas injection valve is designed in a way that the back flow of fluid into the annulus from the tubing does not occur through it.

The injected gas mixes with the multiphase fluid (crude oil, water and gas produced from the reservoir) in the tubing at the point of injection thereby reducing its density and the weight of the liquid column in the tubing. This causes the differential pressure between the reservoir (P_r^i, superscript 'i' represents the ith oil well for all symbols used in this paper) and bottom hole pressure (P_{wf}^i) to be increased causing the liquid column to flow upwards to the surface. The production choke valve controls the flow rate of the fluid (w_{gop}^i) produced from the reservoir. In this paper, it is left at 100% fully open and has not been implemented for control purpose. The mixture of the gas, water and oil flowing out of each of the wells through the production choke valves is collected together in the common gathering manifold and finally transported to the separator where they are separated into their respective constituents. The gas is then sent back to the compressor system and recycled to be used for lifting purpose.

Pressure and temperature transducers measure the pressure and temperature both downstream and upstream the production choke vale and gas lift choke valve. A multiphase flow meter is installed downstream the production choke valve and is used to measure the flow rate of oil, gas and water individually. The packer is used to seal the bottom of the casing annulus, which funnels all of the production into the tubing string, so all of the available gas energy is utilized to lift the fluid.

For the purpose of gas injection distribution, control and optimization, friction losses have not been taken into account. All phases of the multiphase fluid are assumed to be evenly distributed with no slugging. The temperature of lift gas and the multiphase fluid in all sections of pipeline is assumed to be constant at 280 K and the reservoir pressure is kept constant at 150 bar. It is also assumed that flashing does not occur.

3. DEVELOPMENT OF OBJECTIVE FUNCTION

A simple notation of the optimization problem is (Edgar et al., 2001):

Minimize:	$f(\boldsymbol{x})$	Objective function
Subject to:	$h(\boldsymbol{x}) = 0$	Equality constraints
	$g(\boldsymbol{x}) \geq 0$	Inequality constraints

Here $\boldsymbol{x}$ is the vector of decision variables. For the case of the oil field, the decision variables are the flow rates of lift gas through each of the gas lift choke valves.

The amount of oil produced from the reservoir (w_o^i) is a function of the amount of lift gas (w_{ga}^i) injected into the well.

$$w_o^i = f(w_{ga}^i) \qquad 1.$$

Due to limited supply of lift gas, the sum of the total gas injected into the five oil wells should be equal to that supplied by the compressor. This gives rise to a linear equality constraint as,

$$\sum_{i=1}^{5} w_{ga}^i = w_{gc} \qquad 2.$$

w_{ga}^i is the mass flow rate of lift gas through the gas lift choke valve of ith well and w_{gc} is the mass flow rate of lift gas supplied by compressor.

In order to find the lower and the upper limit of the lift gas injection rate, it is assumed that the gas lift choke valve of each well will be at least 10% open and to the maximum 80% open. This range could have been chosen to be from 0% (fully closed) to 100% (fully open), however, seldom are the choke valves operated at their full throttle. Also shutting down one of the oil wells completely would give rise to other factors (for e.g. re-starting the well by following the well unloading procedures). Moreover, the open loop as well as the closed loop simulation results (Sharma et al., 2011) of the simple model of the oil field used in this paper does not show the opening of the valves below 10% and above 80% under normal operating conditions. For calculating the gas mass flow rate at 10% and 80% opening of the valves, steady state operating condition of the field has been considered. In normal operating condition, the pressure upstream the gas lift choke valve (P_c) is assumed to be 200 bar and the pressure downstream the gas lift choke valve (P_a^i) is assumed to be 170 bar (this assumption is made after analyzing a year's data from real oil field).

Mass flow rate through the gas lift choke valve (w_{ga}^i) is obtained by using the standard flow equation developed by Instrument Society of America (ANSI/ISA S75.01, 1989),

$$w_{ga}^i = \frac{N_6 C_v(u_1^i) Y_1^i \sqrt{\rho_{gp} \max(P_c - P_a^i, 0)}}{3600} \quad \frac{kg}{sec} \qquad 3.$$

N_6=27.3 is the valve constant, u_1^i is valve opening of the ith gas lift choke valve expressed in percentage, P_c and P_a^i are the pressures upstream and downstream of the ith gas lift choke valve in bars, ρ_{gp} is the density of gas in the distribution pipeline in kg/m^3 which is a function of the upstream pressure P_c. Y_1^i is the gas expansion factor and $C_v(u_1^i)$ is the valve characteristic as a function of its opening.

We assume the gas expansion factor (Y_1^i) to be:

$$Y_1^i = 1 - \alpha_Y \left(\frac{P_c - P_a^i}{\max(P_c, P_c^{min})} \right) \qquad 4.$$
$$\alpha_y = constant = 0.66$$

P_c^{min} is the minimum pressure in the gas distribution pipeline. For P_c= 200 bar, P_a^i = 170 bar and P_c^{min} = 10 bar, the value of the gas expansion factor Y_1^i = 0.901.

Valve characteristic as a function of its opening ($C_v(u_1^i)$) is modelled by three linear equations as shown in Equation (5).

The function in Equation (5) is fitted to the data supplied by the choke supplier.

$$C_v(u_1^i) = \begin{bmatrix} 0 & u_1^i < 5 \\ 0.111u_1^i - 0.556 & 5 < u_1^i < 50 \\ 0.5u_1^i - 20 & u_1^i > 50 \end{bmatrix} \quad 5.$$

For u_1^i= 80%, $C_v(80\%) = 20$

For u_1^i= 10%, $C_v(10\%) = 0.554$

The density of the lift gas in the distribution pipeline (ρ_{gp}) at P_c = 200 bar can be expressed using the gas law as,

$$\rho_{gp} = 10^5 \frac{MP_C}{RT_p z|_{P_c=200}} \quad \frac{kg}{m^3} \quad 6.$$

Here, M = molecular weight of the lift gas = $20 \times 10^{-3} kg$, P_c = pressure of gas distribution manifold = 200 bar,R = Universal gas constant = 8.314 kg/mole, T_p = gas temperature in the distribution manifold = 280 K and $z|_{P_c=200}$ = gas compressibility factor at a pressure of 200 bar.

The gas compressibility factor given by Equation (7) is expressed as a polynomial function of gas pressure P in bar (assuming constant temperature of 280 K at the bottom of the sea). It is a curve fitted (LSQ-method) to calculations from PVTsim (PVTsim, 2008) using the lift gas composition and assuming constant temperature.

$$z = -2.572 \times 10^5 P^{-8} + 2.322 \times 10^{-5} P^2 - 0.005077P + 1 \quad 7.$$

For a pressure P =*200 bar*, the gas compressibility factor z = 0.7076. The density of the lift gas in the distribution pipeline from Equation (6) is then, ρ_{gp}= 242.83 kg/m^3.

Then using equation (3), the flow rate of the lift gas through the gas lift choke valve is,

For u_1^i = 80% (gas lift choke valve opening of 80%),

$$w_{ga}^i = 11.66 \quad \frac{kg}{sec}$$

For u_1^i = 10% (gas lift choke valve opening of 10%),

$$w_{ga}^i = 0.323 \frac{kg}{sec}$$

Thus the lower and upper bounds for the gas injection rate is given by,

$$0.323 \le w_{ga}^i \le 11.66 \ \frac{kg}{sec} \quad 8.$$

Since the production of the oil from the reservoir is a function of the gas injection rate, let us express w_o^i as a function of w_{ga}^i. Flow rate of the lift gas from the annulus into the tubing through the gas injection valve (w_{ginj}^i) is given by (ANSI/ISA S75.01, 1989),

$$w_{ginj}^i = \frac{K^i Y_2^i \sqrt{\rho_{ga}^i \max(P_{ainj}^i - P_{tinj}^i, 0)}}{3600} \frac{kg}{sec} \quad 9.$$

K^i is the gas injection valve constant, P_{ainj}^i is the pressure upstream the gas injection valve in the annulus and P_{tinj}^i is the pressure downstream the gas injection valve in the tubing, ρ_{ga}^i is the average density of gas in the annulus. Y_2^i is the gas expandability factor given by,

$$Y_2^i = 1 - \alpha_Y \left(\frac{P_{ainj}^i - P_{tinj}^i}{max(P_{ainj}^i, P_{ainj}^{min})} \right)$$

$\alpha_y = constant = 0.66$

P_{ainj}^{min} is the minimum pressure of gas in the annulus at the point of injection. Arranging and solving equation (9) we get,

$$P_{tinj}^i = P_{ainj}^i - \frac{\left(\frac{w_{ginj}^i \times 3600}{K^i Y_2^i} \right)^2}{\rho_{ga}^i} \quad \text{bar} \quad 10.$$

P_{ainj}^i can also be expressed by adding the hydrostatic pressure drop due to lift gas inside the annulus to the pressure P_a^i as,

$$P_{ainj}^i = P_a^i + \frac{\rho_{ga}^i g L_{a\_vl}^i}{10^5} \quad \text{bar} \quad 11.$$

$L_{a\_vl}^i$ is the vertical depth of the annulus from the well head to the point of injection in meters. Combining equation (10) and (11) we get,

$$P_{tinj}^i = P_a^i + \frac{\rho_{ga}^i g L_{a\_vl}^i}{10^5} - \frac{\left(\frac{w_{ginj}^i \times 3600}{K^i Y_2^i} \right)^2}{\rho_{ga}^i} \quad 12.$$

Similarly arranging and solving equation (3) we get,

$$P_a^i = P_c - \frac{\left(\frac{w_{ga}^i \times 3600}{N_6 C_v(u_1^i) Y_1^i} \right)^2}{\rho_{gp}} \quad \text{bar} \quad 13.$$

The bottom hole pressure or well flow pressure $\boldsymbol{P_{wf}^i}$ is,

$$P_{wf}^i = P_{tinj}^i + \frac{\rho_o g L_{r\_vl}^i}{10^5} \quad \text{bar} \quad 14.$$

$L_{r\_vl}^i$ is the vertical length of the tubing below the gas injection point up to reservoir opening in meters and $\rho_o$is the density of crude oil in kg/m^3. From equation (10) and (14) we get,

$$P_{wf}^i = P_a^i + \frac{\rho_{ga}^i g L_{a\_vl}^i}{10^5} - \frac{\left(\frac{w_{ginj}^i \times 3600}{K^i Y_2^i} \right)^2}{\rho_{ga}^i} + \frac{\rho_o g L_{r_{vl}}^i}{10^5} \quad \text{bar} \quad 15.$$

From equation (13) and (15) we get,

$$P_{wf}^i = P_c - \frac{\left(\frac{w_{ga}^i \times 3600}{N_6 C_v(u_1^i) Y_1^i} \right)^2}{\rho_{gp}} + \frac{\rho_{ga}^i g L_{a_{vl}}^i}{10^5} - \frac{\left(\frac{w_{ginj}^i \times 3600}{K^i Y_2^i} \right)^2}{\rho_{ga}^i} + \frac{\rho_o g L_{r_{vl}}^i}{10^5} \quad \text{bar} \quad 16.$$

The mass flow rate of crude oil flowing from the reservoir into the tubing (w_o^i) is calculated using the PI (Productivity Index) model of the well (Brown and Beggs, 1977, American Petroleum Institute, 1994).

$$w_o^i = \frac{PI^i max\left(P_r^i - P_{wf}^i, 0\right)}{3600} \quad \frac{kg}{sec} \qquad 17.$$

P_r^i is the reservoir pressure which is assumed to be constant at 150 bar. Then from equation (16) and (17) we get,

$$w_o^i = \frac{PI^i max\left(P_r^i - P_c + \frac{\left(\frac{w_{ga}^i \times 3600}{N_6 C_v\left(u_1^i\right) Y_1^i}\right)^2}{\rho_{gp}} - \frac{\rho_{ga}^i g L_{a_{vl}}^i}{10^5} + \frac{\left(\frac{w_{ginj}^i \times 3600}{K^i Y_2^i}\right)^2}{\rho_{ga}^i} - \frac{\rho_o g L_{T_{vl}}^i}{10^5}, 0\right)}{3600} \frac{kg}{sec} \qquad 18.$$

Finally the objective function for the optimization problem considering all the five oil wells of the field can be expressed as,

Maximize

$$f\left(w_{ga}^i\right) = \frac{PI^i max\left(P_r^i - P_c + \frac{\left(\frac{w_{ga}^i \times 3600}{N_6 C_v\left(u_1^i\right) Y_1^i}\right)^2}{\rho_{gp}} - \frac{\rho_{ga}^i g L_{a_{vl}}^i}{10^5} + \frac{\left(\frac{w_{ginj}^i \times 3600}{K^i Y_2^i}\right)^2}{\rho_{ga}^i} - \frac{\rho_o g L_{T_{vl}}^i}{10^5}, 0\right)}{\sum_{i=1}^{5} 3600} \frac{kg}{sec} \qquad 19.$$

Subject to the linear equality constraint,

$$\sum_{i=1}^{5} w_{ga}^i = w_{gc} \qquad 20.$$

and the bounds,

$$0.323 \leq w_{ga}^i \leq 11.66 \quad \frac{kg}{sec} \qquad 21.$$

Optimization problem formulated in equations (19), (20) and (21) is a non- linear programming with constraints. In this paper, two approaches to solve the optimization problem have been discussed. The first approach is with the use of *'fmincon'* solver from MATLAB optimization toolbox and the second approach is with the use of modified *'hill climbing'* method.

4. OPTIMIZTION USING SOLVER FROM MATLAB OPTIMIZATION TOOLBOX

'fmincon' is a built-in function in MATLAB for finding the minimum of a constrained non-linear objective function of several variables starting at user specified initial estimate. However, for the case of maximizing oil production, the objective function has to be maximized instead of minimizing. This can be achieved by using *'fmincon'* to the objective function reflected along the x-axis i.e. by taking the negative of the objective function.

In our case, since we have the bound constraints (Equation 21) as well as the linear constraints (Equation 20), *'fmincon'* uses sequential quadratic programming (SQP) method with the *active set* optimization algorithm (MathWorks Inc., 2011). Moreover, to get the optimal points as close as to the global maximum, global search algorithm has been used along with *'fmincon'*. The global search algorithm starts the *'fmincon'* solver from multiple start points (MathWorks Inc., 2011). For details about how global search algorithm and the SQP method are implemented in *'fmnicon'*, refer to the documentation of MATLAB.

To use *fmincon'* matrices containing coefficients of the linear equalities ($A_{eq} and\ B_{eq}$) and the vectors for the lower and upper bounds ($lb\ and\ ub$) have to be passed as argument to it. Equation (20) can be arranged in matrix form $A_{eq}x = B_{eq}$ as,

$$\underbrace{[1 \quad 1 \quad 1 \quad 1 \quad 1]}_{A_{eq}} \underbrace{\begin{bmatrix} w_{ga}^1 \\ w_{ga}^2 \\ w_{ga}^3 \\ w_{ga}^4 \\ w_{ga}^5 \end{bmatrix}}_{x} = \underbrace{w_{gc}}_{B_{eq}} \qquad 22.$$

Similarly, from Equation (21), the lower and upper bounds are,

$$\underbrace{0.323}_{lb} \leq w_{ga}^i \leq \underbrace{11.66}_{ub} \qquad 23.$$

4.1 Optimization with the cascade control structure

Optimization of the lift gas distribution in a gas lifted oil well is performed along with the cascade control structure. A schematic diagram of the cascade control structure is shown in Figure 4. In cascade control strategy, a pressure transducer measures the pressure of the common gas distribution manifold which is then taken as feedback to a pressure controller. The set point to this pressure controller is 200 bar. The output of the pressure controller gets added to the nominal set points of five flow controllers, the result of which is then given as the current set point to the five flow controllers.

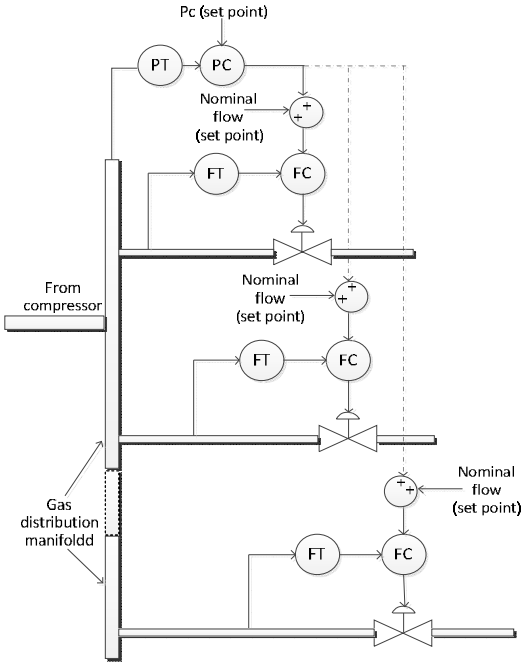


Figure 4: Schematic for cascade control structure.

The five flow controllers control the rate of flow of lift gas through each of the gas lift choke valves. The pressure controller is responsible for maintaining a fairly constant pressure of the lift gas in the gas distribution manifold by manipulating proper lift gas flow rates through the valves. The optimal flow rate set points calculated by the optimization algorithm is actually utilized by the cascade control structure as the nominal flow set points for five flow controllers.

4.2 Discussion on simulation results

The non-linear objective function with linear equality constraints and inequality bounds was solved using MATLAB optimization toolbox. It was assumed that the total available lift gas was 40,000 Sm^3/hr. At first the nominal set-points were distributed randomly (non-optimally) among well 1 to well 5 consuming 15%, 17%,25%,23% and 20% of the total available lift gas respectively. The non-optimal flow set points are listed in Table 1.

Table 1: Non- optimal distribution of lift gas flow rate set points

Well no.	Well 1	Well 2	Well 3	Well 4	Well 5	Unit
Nominal flow set points	1.38	1.57	2.31	2.12	1.84	—

The process was first allowed to reach the steady state with the controllers running alone. Optimizer loop was activated once the process reached the steady state. The values of the different process variables (like pressure of gas distribution manifold, valve openings, Productivity Index (*PI*) values etc.) used in the objective function (Equation 19) were taken as the steady state values.

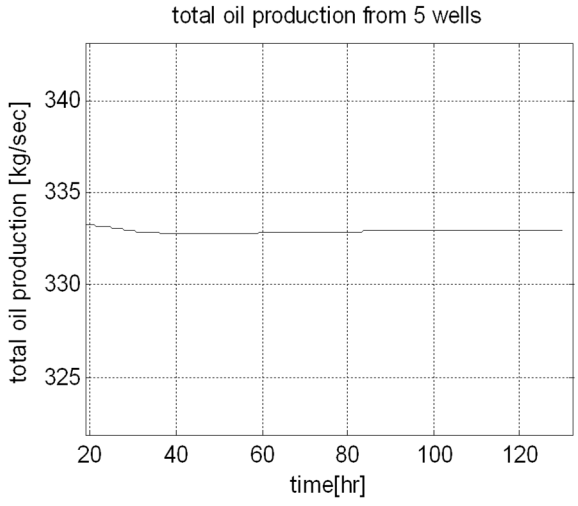


Figure 5(a): Total oil production without optimization

In Figure 5(a), the total oil produced from the five oil wells without any optimization was about 332 kg/sec which resembled to the total oil production of the real oil filed at Norne. However, after the process reached the steady state, at t = 50 hours, the optimizer loop was activated once to find the optimal set points for the flow controller. The optimal set points returned by the optimizer were used and the simulation was continued with the controllers alone. The total oil produced was increased by 14 kg/sec from about 332.5 kg/sec to about 346.5 kg/sec as shown in Figure 5(b) using the new optimal set points. The optimal flow set points returned by the optimizer are listed in Table 2.

Table 2: Optimal flow set points calculated by optimizer

Well no.	Well 1	Well 2	Well 3	Well 4	Well 5	Unit
Nominal flow set points	2.12	1.37	1.37	4.05	0.32	—

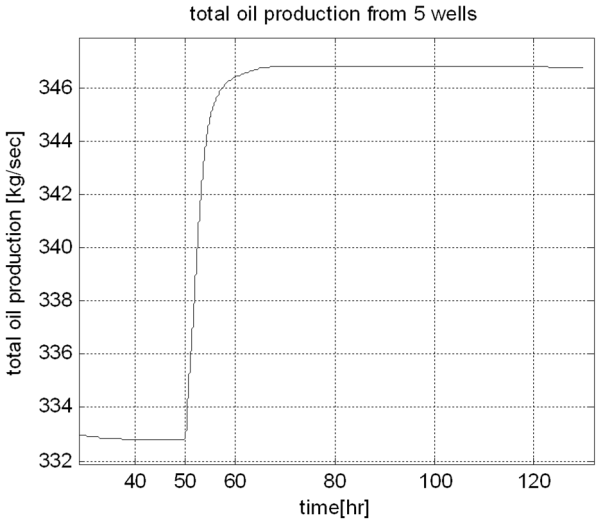


Figure 5(b): Total oil production with optimization turned 'ON' at t = 50 hours.

The distribution of the lift gas through the five gas lift choke valve after the application of the new optimal flow set points and with the controllers still active is shown in Figure 6(a). Similarly the distribution of the lift gas through the choke valves without any optimization is shown in Figure 6(b).

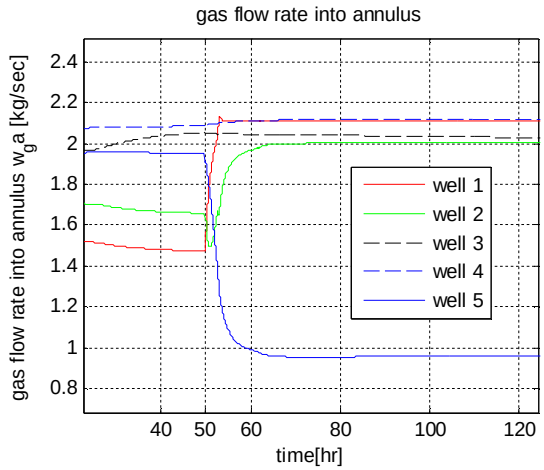


Figure 6(a): Lift gas distribution among wells with optimization.

As can be seen from Figure 6(b), when the optimizer was not activated at t = 50 hours after the process reached the steady

state, there was no re-distribution of the lift gas and the process continued to operate in its steady state.

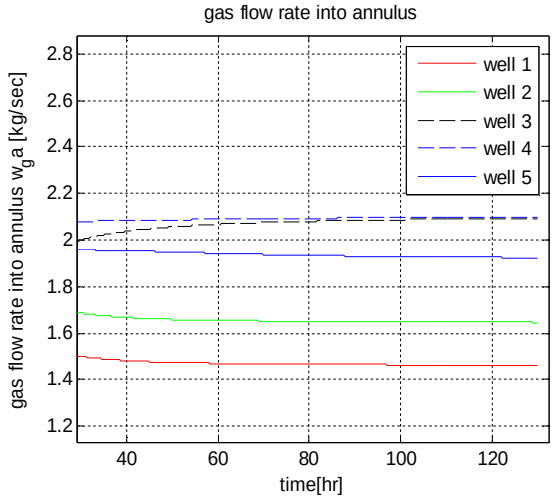


Figure 6(b): Lift gas distribution among wells without optimization.

However, when the optimizer was turned 'ON' at t = 50 hours as in Figure 6(a), due to the new optimal set points of the flow controllers generated by the optimizer, the lift gas distribution was changed resulting in an increased oil production.

Since re- distribution of the lift gas took place at t = 50 hours as a result of optimization, the oil produced from each oil well was also re-distributed in accordance to the distribution of the lift gas as shown in Figure 7.

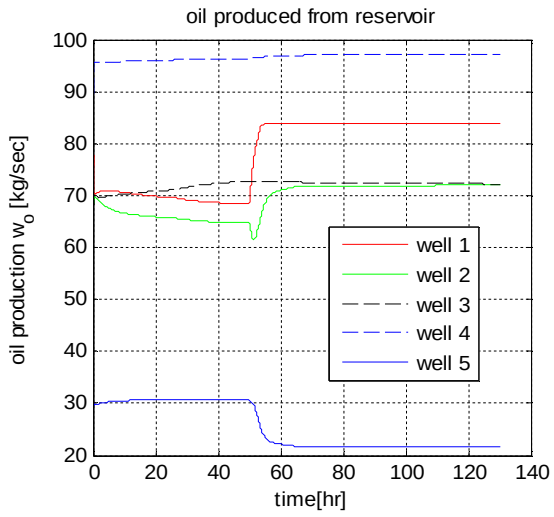


Figure 7: Oil produced from each oil well after optimization.

Production of oil from oil well 1 and oil well 2 increased because the lift gas supplied to these wells was increased after optimization (see Figure 6(a)). Similarly, the oil flow rate from well 5 decreased due to less injection of lift gas after optimization (see Figure 7).

Finally, it can be concluded from the simulation results that the outcome of optimization using the built-in solver in MATLAB is an increased total oil production. Expressing it in percentage, the increased production due to process optimization was approximately 4.21%.

4.3 How often to perform optimization?

The effect on the total oil production when optimization is performed multiple times is an interesting topic to discuss. To check this, optimization was performed at t = 50 hours under the availability of 40000 Sm^3/hr of gas supply (see Figure 8). The process with increased oil production due to this first place optimization was then allowed to reach the steady state. After the total oil production flow rate reached steady state, at t = 90 hours, the supply of lift gas was reduced to 36000 Sm^3/hr. Due to the reduction in supply gas, the total oil production started to decrease. The process was again allowed to reach the steady state at the lower supply of lift gas. Then at t = 120 hours, when the total oil production flow rate reached steady state, optimization loop was activated for the second time.

Figure 8 shows that when the optimization was activated for the second time after the application of input disturbance, it had no effect on the total oil production and the total oil production rate remained the same. The initial values given to the optimizer loop turned ON at t = 120 hours were the steady state flow rate values through each gas lift choke valves. The optimizer loop after performing calculations returned back the same initial values as the optimal values.

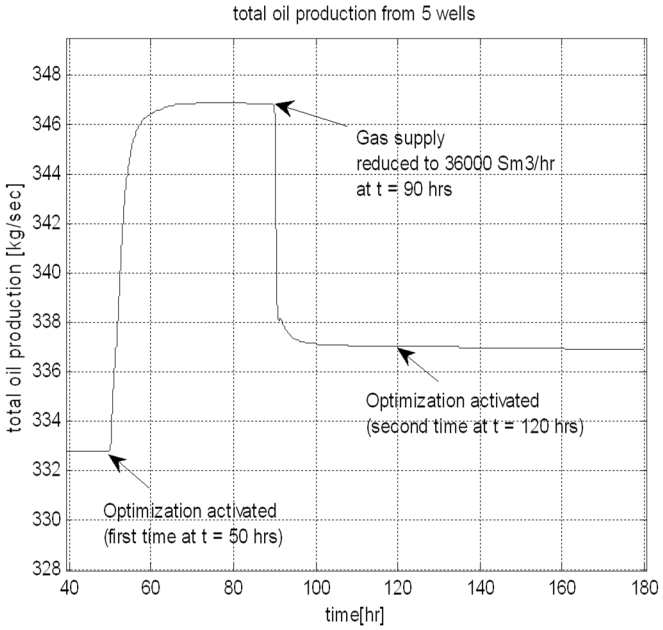


Figure 8: Total oil production when optimization was performed for the second time after input disturbance.

This could be due to the effect of the optimizer run for the first time at t = 50 hours. The process may have already obtained the optimal flow rate set points for all the five flow controllers when the optimizer was activated for the first time. When input disturbance was given at t = 90 hours, the controllers may have decreased the gas flow rates through the valves in an optimal way because of the already available optimal flow rate set points. So the reduced total production

due to reduced gas supply may have already acquired the optimal steady state at t = 120 hours.

4.4 Optimization after the reduction of the supply of lift gas.

An obvious question that can probably occur is the doubt of whether the optimizer loop failed to function properly for a reduced supply of lift gas. To justify this question about behaviour of the optimizer activated for the second time, an interesting point would be to see how the total oil production will be affected when the optimizer is activated for the first time only after the application of the input disturbance i.e. without any previous activation of the optimizer. In this case, at t = 60 hours when the process reaches steady state for 40000 Sm^3/hr supply of gas, instead of activating the optimizer, the gas supply was reduced to 36000 Sm^3/hr. The reduced oil production due to reduced supply of gas was again allowed to reach the steady state. Then at t = 120 hours, the optimizer was activated for the first time.

As can be seen from Figure 9, when the optimizer was activated for the first time after applying the input disturbance without any prior optimizations, the total production of the oil increased by around 15kg/sec from about 321 kg/sec to about 336 kg/sec. This clearly implies that the optimizer loops functions properly even for reduced supply of gas.

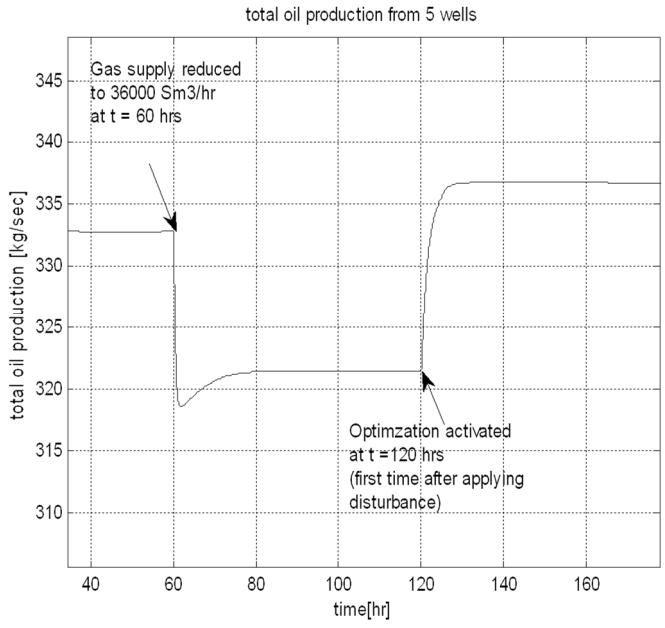


Figure 9: Total oil production when optimization was performed for the first time, only after input disturbance was applied.

It can thus be briefly concluded that one single optimization is sufficient enough to bring the process to optimal operating condition and further optimization is not necessary even for changing total gas supply as input disturbances.

5. OPTIMIZATION USING HILL CLIMBING METHOD

Hill Climbing is an iterative method of finding the maximum/minimum of a function with the decision variable(s) where only one of the decision variables is changed at a time keeping all the others unchanged. Iteration is started with the initial values of the decision variables provided by the user. The algorithm then tries to find better value of by increasing/decreasing only a single decision variable at one time. If the change provides a better solution than before, the decision variable is slightly incremented and the new solution is calculated. Hill climbing aims to ascend to a peak by repeatedly moving to an adjacent state with a higher fitness (Juels and Wattenberg, 1994). This step is repeated until the increment of the decision variable provides no further better solution. The whole process is repeated for each remaining decision variable. Finally, the set of decision variables is the optimal set of variables.

However, one difficulty in using hill climbing method in the case of the oil field where the decision variables are the gas flow rate through the five gas lift choke valves is that, when an increment is made in the gas flow rate of only one of the gas lift choke valve, gas flow rate through another (or all of the remaining 4 valves) has to be decreased exactly by the same amount by which the increment was made to maintain the gas flow rate constraint (the supply gas flow rate should be equal to the sum of the gas flow rates through the five gas lift choke valves). Since at least two decision variables have to be manipulated at the same time, this approach can be considered to be a modified version of the general hill climbing method. Furthermore, the hill climbing method for solving the optimization problem has been done to check the performance of the optimization problem solved by using *'fmincon'* solver of MATLAB optimization toolbox as described in Section 4.

5.1 Strategy for hill climbing method

The gas lift choke valves are assumed to be never fully closed and never fully open. They are assumed to be open from 10% to 80%. If the supply gas flow rate is denoted by *'total flow'* then 10% gas lift valve opening of one of the well means that amount of gas flowing through that well is 3.5% of *'total flow'* and that for 9.22 kg/sec gas flow rate (total gas flow rate at normal condition) through the well, it means that the amount of gas flowing through the well is 100% of *'total flow'*.

Initially all the five oil wells have an equal lift gas distribution of 20% of *'total flow'*. The strategy is to decrease the flow rate in one of the wells (also called here as *'starting well'*) from initial equal distribution of 20% of *'total flow'* to 3% of *'total flow'* in smaller steps of 1% of *'total flow'* at each iteration. At the same time the flow rate of another well (also called here as *'helping well'*) is increased from 20% of *'total flow'* to 37% of *'total flow'* with the same time step of 1% of *'total flow'* at each iteration. 18 iterations are performed and the total oil production at each iteration are calculated and stored. After this, keeping the same oil well as *'starting well'*, the remaining three oil wells perform the role of the *'helping well'* turn by turn. A sub-total of 72 iterations will be performed and for each iteration, the total oil production is calculated and stored.

When all the remaining three oil wells have completed working as *'helping well'*, the role of the *'starting well'* is undertaken by next oil well. The remaining oil wells will again function as *'helping well'* for this new *'starting well'*. The whole process is repeated until each of the five oil wells work as *'starting well'*. At the end, a total of 360 iterations will be performed. The set of gas flow rates which gives the highest oil production among these 360 iterations is considered to be the optimal gas flow rates.

It should be noted that the step change of 1% of *'total flow'* is equivalent to 5% change in the flow rate of lift gas in each well. This step of 5% change in the lift gas flow rate in each well is assumed to provide observable change in the total oil production. Moreover, for each oil well, the iteration swings from 3% of total production to 37% of total production, which is sufficient enough to provide a set of good local optimal set points.

5.2 Discussion on simulation results

Hill climbing method was used to solve the non-linear objective function with linear equality constraints and inequality bounds. It was assumed that the total available lift gas was 40,000 Sm^3/hr. At first the nominal set-points were distributed randomly (non-optimally) among well 1 to well 5 consuming 15%, 17%,25%,23% and 20% of the total available lift gas respectively as listed in Table 1.

The cascade control structure of Figure (4) was used along with the optimizer for optimal control and distribution of lift gas and for optimal production of oil. All the assumptions and conditions used for optimization using built-in *'fmincon'* solver from MATLAB optimization toolbox have also been used for hill climbing method.

The process was first allowed to reach the steady state with the controllers running alone. At t = 60 hours, when the process reached the steady state, the optimizer loop was activated once to find the optimal set points for the five flow controllers. The optimal set points returned by the optimizer were used as the new nominal set points and the simulation was continued with the controllers alone. The total oil produced from the oil field was increased to about 347 kg/sec as shown in Figure 10.

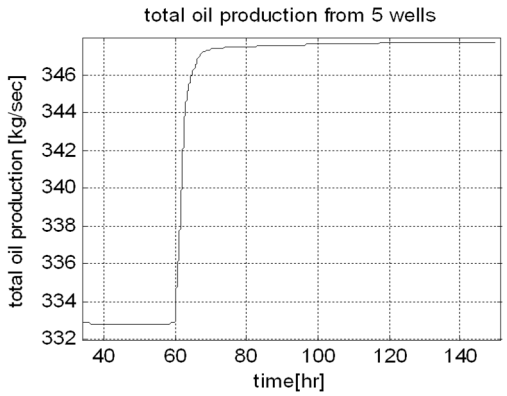


Figure 10: Total oil production with hill climbing optimization turned 'ON' at t = 60 hours.

Without any optimization, the total oil production was 332 kg/sec as shown in Figure 5(a). So optimization caused an increase in the total oil production by about 15 kg/sec. When expressed in percentage, the oil production was increased by about 4.5%, which is very much similar to what was obtained with optimization using solver from MATLAB optimization toolbox.

The optimal flow set points returned by hill climbing optimizer are listed in Table 3.

Table 3: Optimal flow set points calculated by hill climbing optimizer.

Well no.	Well 1	Well 2	Well 3	Well 4	Well 5	Unit
Nominal flow set points	3.32	1.84	1.84	1.84	0.37	—

The initial values given to the optimizer loop was the equal gas distribution of 1.8444 kg/sec (with 40000 Sm^3/hr of supply gas). Since with hill climbing method used for this case, only two decision variables gets modified at a time, the optimizer returned a set of gas flow rates where only two decision variables were altered with the remaining having values equal to 1.8444 kg/sec as shown in Table 3.

The distribution of lift gas among the five gas lift choke valves after the application of the new optimal flow set points and with the controller still active is shown in Figure 11.

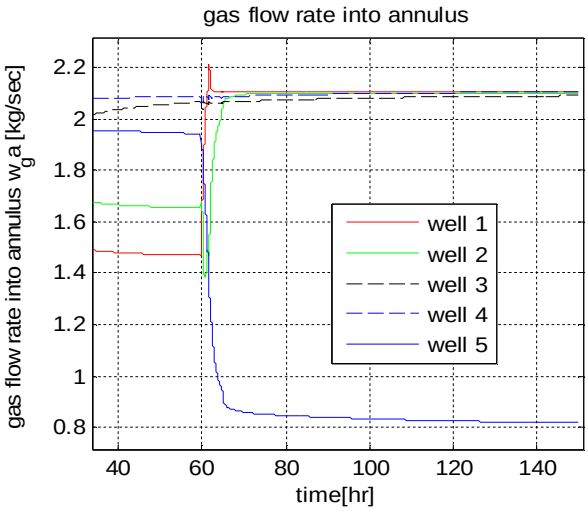


Figure 11: Gas distribution among wells using hill climbing optimizer turned 'ON' at t = 60 hours.

The gas flow rate though well 5 which is the least producing well (Sharma et. al., 2011) has decreased and that through well 1 and well 2 have increased after the hill climbing optimizer was turned 'ON' at t = 60 hours. Without any optimization, there was no re-distribution of the lift gas among the wells and the process continued to operate in its steady state as shown in Figure 6(b).

Since re-distribution of the lift gas took place at t = 60 hours as a result of optimization, the oil produced from each oil well were also re-distributed in accordance to the distribution of the lift gas and is shown in Figure 12.

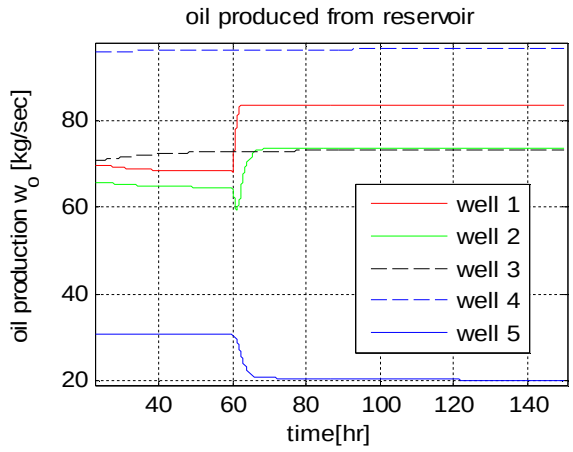


Figure 12: Oil produced from each oil well after optimization with hill climbing method.

Production of oil from oil well 1 and oil well 2 increased because the lift gas supplied to these wells was increased after optimization (see Figure 11). Similarly, the oil flow rate from well 5 decreased due to less injection of lift gas after optimization (see Figure 12) which is very similar to what was obtained using the built-in solver of MATLAB optimization toolbox.

Finally, it can be concluded from the simulation results that the outcome of optimization using the hill climbing method is an increase in the total oil production. The simulation results obtained using the hill climbing method can also be used as the verification of the results obtained from built-in optimization solver.

6. COMPARISON OF THE TWO OPTIMIZATON METHODS

Both the methods of optimization were successfully tested through simulations and both of them could optimize the process resulting in increased oil production. The amount of total oil production increased after the optimization was almost the same for both the methods. So it is difficult to compare these two methods based on how much oil production was increased by each optimization method.

The built-in solver of MATLAB optimization toolbox uses the mathematical model of the oil field. Mathematical model of a complex process is just an approximation of the real process. Models of complex process will always have some assumptions and it cannot represent the real process completely. Moreover, as time passes and as the process becomes older, the dynamics of the process may change over time due to several factors. The mathematical model developed and tested when the plant was young may not be a good representation of the older model. Under this situation, optimization methods based on the mathematical model of the process might not at all provide optimal results in the real oil fields. Thus hill climbing optimization procedure which is independent on mathematical modeling of the process is advantageous in this regard over the use of MATLAB solver for solving an optimization problem which is an important strength for real applications.

'fmincon' which is a built-in solver of Optimzation toolbox in MATLAB uses *active set* optimization algorithm utilizing sequential quadratic programming (MathWorks Inc, 2011). This method uses the information of the gradient of each decision variables such that all the decision variables can be changed in each iteration. In case of hill climbing method, only single variable can be altered at a time (the other variable(s) were changed only to meet the constraints). Also, the built-in solver uses *'global solution'* class which has a tendency to provide better optimal solution (more towards global solution) than the hill climbing method which only provides local optimal solution.

A number of field experiments should be performed if hill climbing method is to be applied to real oil field. The normal operation of the existing oil field has to be disrupted for performing these tests during which time there might be some loss of total production of oil. However, this loss of oil during the test period (if they are not very long) can be compromised with an increase in the total oil production for a longer period of time due to optimization.

To conclude the use of *'fmincon'* as built-in solver from MATLAB optimization toolbox is recommended when simulator (mathematical model of the process) is used for optimization. But, for real oil fields, hill climbing method might be more beneficial and realistic to use.

7. CONCLUSION

For optimal distribution of the available lift gas among the five oil wells in order to maximize the total oil production, a non-linear optimization problem with linear constraints and inequality bounds was formulated using the model of the process at steady state. The optimization problem was then solved using two methods; one was using the MATLAB optimization toolbox and another was by using the hill climbing method of optimization. Both the optimization methods could increase the total oil production by about 4.2%. One advantage of doing optimization with hill climbing method in a real oil field is that it does not require a mathematical model of the oil field so it is free of modeling errors and assumptions. However, it does require a number of experiments to be performed in the real field for which the normal operation of the oil field might have to be obstructed.

Optimization with the *'fmincon'* solver was performed including the global search algorithm, so this method has the tendency to provide better optimal solution (more towards global solution) than the hill climbing method which only provides local optimal solution. It can also be briefly concluded that one time optimization is sufficient enough to bring the oil field to an optimal state and multiple optimization is not necessary to be done when total gas supply is varied as input disturbance.

REFERENCES

American Petroleum Institute (1994). API gas lift manual, Book 6 of the Vocational Training series, Third Edition. *American Petroleum Institute, Exploration and Production Department, Washington, DC 20005-4070, USA.*

ANSI/ISA S75.01, (1989). Flow Equations for Sizing Control Valves, Standards and Recommended Practices for Instrumentation and Control, *10th edition, Vol. 2.*

Brown, K. E. & Beggs, H. D. (1977). The technology of artificial lift methods, Volume 1, Inflow Performance, Multiphase flow in pipes, The flowing well, *PennWell Publishing Company, Tulsa, Oklahoma, ISBN: 0-87814-031-X.*

Camponogara, E. & Nakashima, P. H. R. (2006). Solving a gas-lift optimization problem by dynamic programming. *European Journal of Operational Research, Vol. 174, pp: 1220-1246.*

Edgar, T. E., Himmelblau, D. M. & Lasdon, L. S. (2001). Optimization of chemical processes, *2nd edition, McGraw-Hill Chemical Engineering Series, ISBN 0-07-118977-7.*

Juels, A. & Wattenberg, M. (1994). Stochastic Hillclimbing as a Baseline method for evaluating Genetic Algorithms. *Dept. of Computer Science and Dept. of Mathematics, University of California at Berkeley.*

Kosmidis, V. D., Perkins, J. D. & Pistikopoulos, E. N. (2005). A mixed integer optimization formulation for the well scheduling problem on petroleum fields. *Journal of Computers and Chemical Engineering, Vol. 29, pp. 1523-1541.*

MathWorks Inc. (2011). R2011a Documentation – Optimization Toolbox *[Online]: Available from: http://www.mathworks.com/help/toolbox/optim/ug/fmincon.html. [Accessed: 20.05.2011].*

PVTsim. (2008). PVT simulation program developed for reservoir engineers, flow assurance specialists, PVT lab engineers and process engineers *[online]. Available at: <http://www.pvtsim.com> [Accessed: 20 January 2011].*

Rashid, K. (2010). Optimal allocation procedure for gas-lifted optimization. *Industrial and Engineering Chemistry Research, Vol. 49, pp: 2286-2294.*

Ray, T. & Sarker, R. (2007). Genetic algorithm for solving a gas lift optimization problem. *Journal of Petroleum Science and Engineering, Vol. 59, pp: 84-96.*

Saepudin, D., Soewono, E., Sidarto, K. A., Gunawan, A. Y., Siregar, S. & Sukarno, P. (2007). An investigation on gas lift performance curve in an oil producing well. *International Journal of Mathematics and Mathematical Science, Vol. 2007, Article ID 81519, pp. 1-15, DOI: 10.1155/2007/81519.*

Sharma, R., Fjalestad, K. & Glemmestad, B. (2011). Modeling and Control of Gas lifted oil field with five oil well. *Paper submitted to SIMS conference, SIMS 2011, 29 Sept, Sweden.*

Souza, J. N. M., Medeiros, J. L., Costa, A. L. H. & Nunes, G. C. (2010). Modeling, simulation and optmization of continuous gas lift systems for deepwater offshore petroleum production. *Journal of Petroleum Science and Engineering, Vol. 72, pp: 277-289.*

Sukarno, P., Saepudin, S. D., Soewono, E., Sidarto, K. A. & Gunawan, A. Y. (2009). Optimization of gas injection allocation in a dual gas lift well system. *Journal of Energy Resources Technology, Vol. 131, Issue 3, DOI: 10.1115/1.3185345.*

Zhong, H., Li, Y. & Liu, Y. (2004). *An approach for optimization of gas-lift allocation to a group of wells and oil field*. Mathematics China, Vol. 1, No. 1, pp: 1-11.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Abstract

Model-Based Optimization of Economical Grade Changes for the Borealis Borstar® Polyethylene Plant

Per-Ola Larsson[1], *Johan Åkesson*[1], *Niclas Carlsson*[2], *Niklas Andersson*[3]
[1]*Dept. of Automatic Control, Lund University*, [2]*Borealis AB, Stenungsund,*
[3]*Dept. of Chemical Engineering, Lund University*

Economical grade changes are considered for a Borealis Borstar® polyethylene plant model. Figure 1 shows a schematic diagram of the plant, which incorporates two slurry-phase reactors, one gas-phase reactor and a recycle area with three distillation columns. The model is constructed in the Modelica language and the JModelica.org platform is used for optimization. The designed cost function expresses the economical profit during a grade change and is formulated using on-grade intervals for seven polymer quality variables such as melt index, density and reactor split factors reflecting polymer bi-modality. Additionally, incentives to produce polymer with quality variables on grade target values, not only inside grade intervals, are added together with a preparatory time interval prior defined transition time. In total, twelve inflows and three purge flows are used at optimization. Two optimal grade changes show the effect of using a cost function that regards plant economy. The resulting trajectories can be divided into three phases with distinguishing features, and the synchronization of inflows and usage of recycle area off-gas flows are important in the optimized grade changes. Figure 2 shows melt index and density during two optimized transitions while Figure 3 shows when the polymer is considered on- and off-grade, the profit and the plant production rate. A full version of the article can be found in the Ph.D. thesis [1].

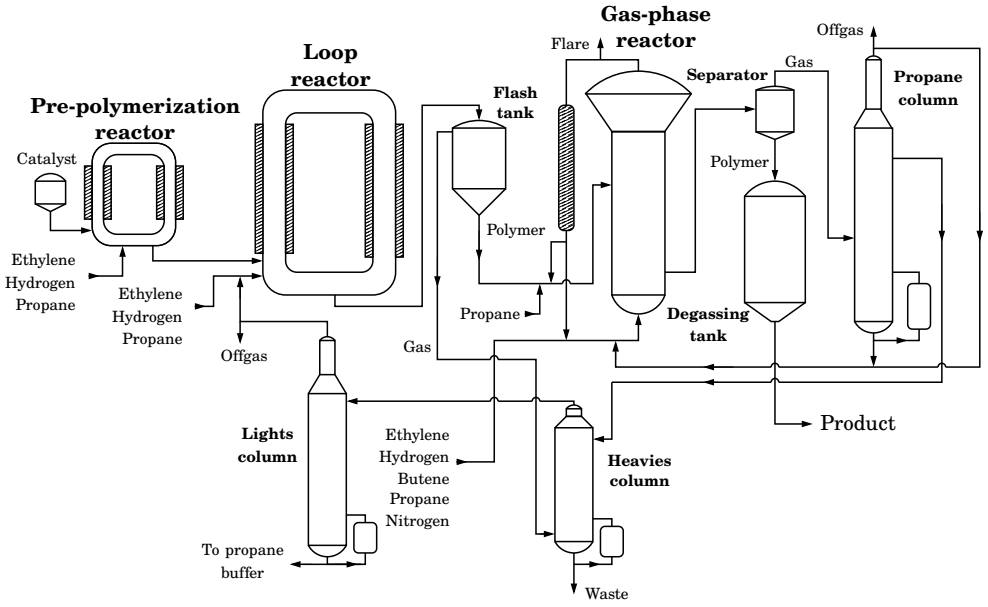


Figure 1 Overview of the plant PE3 at Borealis AB including three Borealis Borstar® reactors and a recovery area with three distillation columns.

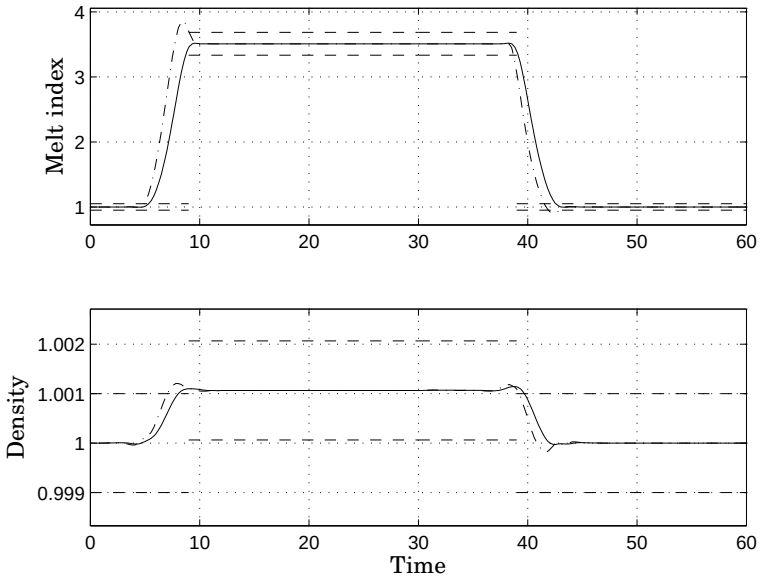


Figure 2 *Upper:* Instantaneous and bed average polymer melt index. *Lower:* Instantaneous and bed average polymer density.

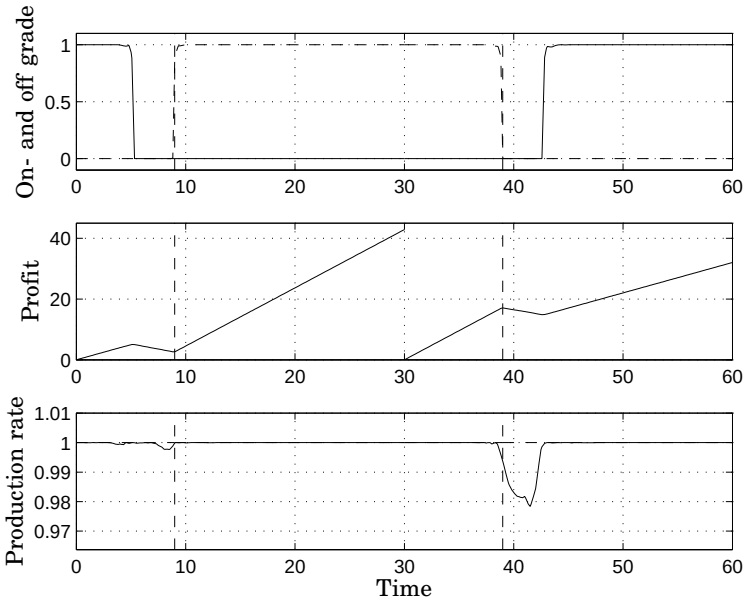


Figure 3 *Upper:* Functions determining if production is on- or off-grade. *Middle*: Cumulative profit. *Lower:* Production rate.

References

[1] Per-Ola Larsson. *Optimization of Low-Level Controllers and High-Level Polymer Grade Changes*. PhD thesis, Department of Automatic Control, Lund University, Sweden, November 2011.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Modeling Operating Modes during Plant Life Cycle

Sten Bay Jørgensen[1] and Morten Lind[2]
[1]CAPEC, Department of Chemical and Biochemical Engineering
[2]Department of Electrical Engineering
Technical University of Denmark, Denmark

Abstract

Modelling process plants during normal operation requires a set a basic assumptions to define the desired functionalities which lead to fullfillment of the operational goal(-s) for the plant. However during during start-up and shut down as well as during batch operation an ensemble of interrelated modes are required to cover the whole operational window of a processs plant including intermediary operating modes. Development of such an model ensemble for a plant would constitute a systematic way of defining the possible plant operating modes and thus provide a platform for also defining a set of candidate control structures. The present contribution focuses on development of a model ensemble for a plant with an illustartive example for a bioreactor.

Starting from a functional model a process plant may be conceptually designed and qualitative operating models may be developed to cover the different regions within the plant operating window, including transitions between operating regions. Subsequently qualitative functional models may be developed when the means for achieving the desired functionality are sufficiently specified during the design process. Quantitative mathematical models of plant physics can be used for detailed design and optimization. However the qualitative functional models already provide a systematic framework based on the notion of means-end abstraction hierarchies. Thereby functional modeling provides a scientific basis for managing complexity. A functional modelling framework has been implemented to facilitate model development and application in a computer environment. Defining means-end causal relations makes it possible to perform qualitative causal reasoning within a functional modelling framework. Thus such a framework renders it possible to develop potentially feasible control structures. This ability is based on goal reasoning and development of goal trees from causal relations. These capabilities of functional models extend the application potential of functional modelling significantly beyond that of conventional mathematical modeling representing quantitative physical phenomena.

The example case is a continuously operating bioreactor for manufacturing single cell protein from methane where also the bioreactor start-up is illustrated with switching between operating modes and their associated control structures as seen in a multiloop control configuration.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Teaching Control Principles to Industry Practitioners

Kristian Soltesz * **Charlotta Johnsson** * **Tore Hägglund** *

* *Dept. of Automatic Control, Lund University, Sweden.*
(e-mail: {kristian, charlotta, tore}@control.lth.se)

Abstract: This paper addresses the need of continued education of process industry practitioners such as operators and instrumentation engineers. The process industry regulatory control tuning situation of today is reviewed. Areas of potential improvement are identified. A course, aimed at fulfilling these needs is presented. Especially, useful laboratory experiments are outlined. The suggested course was given within PICLU – a regional collaboration between academia and process industry in Scandinavia.

1. INTRODUCTION

Automatic control is a subject rich in both mathematics and practical considerations. To device adequately working control systems, it is therefore important to have a broad competence span. Teaching automatic control to an audience without a strong mathematical background can be challenging, for students and instructors alike. Even the more basic theory of the commonly occurring Proportional Integrating and Derivating (PID) controller is based on concepts of ordinary differential equations, linearization, Laplace transforms and matrix algebra Åström and Hägglund (2006).

Likewise, an audience without practical experience generally has difficulties estimating the skills and effort required to implement a control system in a non-ideal world Kheir et al. (1996).

Part of the activity of PICLU, the Process Industrial Centre at Lund University pic (2010), is to provide technology transfer to regional process industry. As part of this mission, the Department of Automatic Control at Lund University is giving a series of courses aimed at different categories of industry professionals. The first course in this series was held in the spring of 2010 and aimed at practitioners such as instrumentation and process engineers.

The main purpose of this paper [1] is to draw attention to a situation, where much is to be earned. In addition, it is the hope of the authors to inspire to take similar pedagogic initiatives and reach out to audiences, which are generally forgotten, mainly due to lacking mathematical background. For this purpose a course format, which was found to work well for instructors and participants alike, is presented.

2. AUDIENCE

The audience consisted of industrial professionals, working in close connection to process industry processes. An illustrative way of introducing the background of the audience is through the functional model of Skogestad, Skogestad (2004), shown in Figure 1. In his model, Skogestad decomposed a generic process industrial plant into a vertical functional hierarchy. Each level is defined through its complexity and time scale of operation. The mentioned audience is employed within what Skogestad refers to as the 'Control layer', decomposed into 'Supervisory' and 'Regulatory' control. Practically this means that they work in close connection to physical processes, and have extensive hand-on experience.

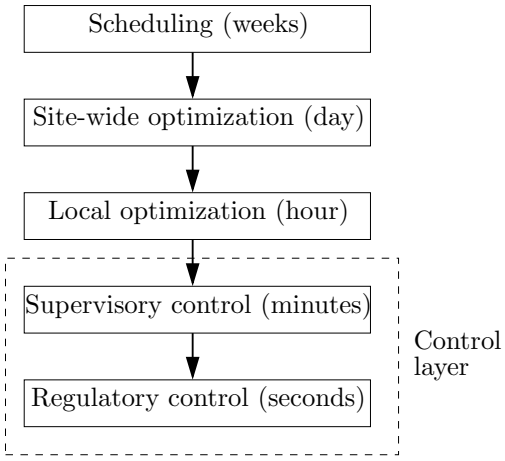


Fig. 1. Skogestad's functional model of process industry facility.

Parts of the audience have an academic background, however, not often in control systems. Some have started their careers as process operators and transcended from working in the 'Regulatory control' layer, to a more conceptually focused position.

Unlike what is commonly found among students in academia, the audience has a strong practical background and good practical intuition. They are generally motivated to learn new concepts, directly applicable in their professional work. However, they are not used to the format of university education (lectures, exercises, laboratory sessions) and have a limited theoretical background in control systems.

[1] Parts of this paper have been previously presented at the SEFI Annual Conference 2011, Lisboa, Portugal.

One aim of the course would hence be to exploit the intuition and motivation of the participants, without being limited by the format in which control systems are traditionally taught at an introductory university level. Before formulating the goals of such a course in greater detail, motivations for giving it will be presented.

3. THE IMPORTANCE OF CONTINUED EDUCATION

3.1 The Tuning Situation

PID control Åström and Hägglund (2006) is a technology well over 50 years old. Still, today, over 95% of all regulatory control loops in process industry are PID. Of the PID loops over 90% are PI. Studies, e.g. Panagopoulos (2000), have shown that adding derivative action would increase performance in many cases. However, it has often been omitted due to difficulties of tuning.

Although more advanced control strategies such as MPC Garca et al. (1989) are emerging to some extent, the base level controllers in an MPC solution are still typically PIDs.

As with any control technology, the PID controller needs to be tuned to function adequately. Even though most industrial PIDs are compensating stable, slow, reasonably damped processes with mainly monotonous step responses Hägglund (2008), several surveys witness of surprisingly poor performance. A survey by Ender in 1993 Ender (1993) on regulatory PIDs concludes that:

- $> 30\%$ operate in manual
- $> 30\%$ increase short term variability
- $\approx 25\%$ use factory default parameters

A similar survey by Bialkowski in 2002 gave the following numbers for PID controllers within process industry:

- 50% work well
- 25% ineffective
- 25% dysfunctional

A plausible reason for these figures is the cost of properly modeling and tuning a PID control loop. Hiring a consultant for the task is USD $250 - 1000$ in work costs alone, according to a survey by Honeywell Desborough and Miller (2002). Many companies have this competence in-house, but the holders of it are generally occupied above the 'Control layer' of Figure 1.

Providing operators and instrumentation engineers with the knowledge needed to conduct model-, rather than intuition-based tuning could contribute to improve the situation significantly.

3.2 Reliance, Disuse and Misuse

Handling undesired behavior in process industry control loops often involves switching the loop to manual mode. If the control system is critically malfunctioning, this is well motivated. However, switching to manual mode when the control is functional or not switching when it is dysfunctional, should be avoided.

In Dzindolet et al. (2003), the concepts of misuse and disuse are defined in the context of control reliance. Disuse is the under-utilization of functional control, while misuse occurs when the operator overly relies on control. Psychological experiments in the paper show that disuse and misuse decreases significantly if the operator is given a rational explanation to the behavior of the control system. The main conclusions in the paper is that optimizing a plant alone, is of limited value, if the operators are not updated on the underlying principles. The *interaction* between automated aid and human operator must be considered.

3.3 Increased Efficiency through Awareness

Another reason for continued education is to develop the ability to identify 'low hanging fruit'. Some control systems can be significantly improved by retuning or introducing an extension such as derivative action, a feed forward link or cascade structure. By learning to identify these situations, an individual can contribute significantly to the increased efficiency of the control system.

3.4 Personal Motivation

There is an additional motivation for continued education, which differs slightly from the ones already mentioned. The individual employee participating in the education, will generally feel recognized by his/her company. This fact, in combination with the aspects discussed above, may contribute to a more positive atmosphere, where own initiatives for improvements are closer at hand.

4. COURSE LAYOUT

4.1 Goals

A challenge in automatic control eduction is how to include practical experiments in an otherwise mathematically oriented curriculum Åström and Lundh (1992). The challenge faced here is the opposite. Practitioners generally have many hours of on-site experience. Rather than providing a complete control course, the proposed course aimed at fulfilling the following goals:

- Give a thorough understanding of the simple control loop.
- Become familiar with process types common in industry.
- Recap manual tuning of the PID controller and introduce alternatives.
- Go through the anatomy of the PID controller and handle practical implementation aspects.
- Introduce more advanced control structures such as cascades and feed forward links.
- Discuss the influence of sensor and actuator placement and characteristics.

Based on these goals, and motivated by experience from teaching undergraduate control courses, a course outline was assembled.

4.2 Methodology

The design of the course was influenced by the concepts of 'Zone of proximal development' Vygotsky (1978) and

'deep versus superficial learning' Marton and Säljö (1976) by Vygotsky and Säljö, respectively. A brief review of these pedagogical ideas, and their influence on the course, are given below. Similar concepts are thoroughly handled in Biggs and Tang (2007).

Zone of Proximal Development The ability to acquire new knowledge and skills is strongly coupled to what one presently knows. Vygotsky studied this basic idea more closely and introduced the 'zone of proximal development', being a set of yet unacquired knowledge or skills, lying close to what is already familiar to the learner. Based on Vygotskys studies, it was natural to make laboratory exercises a central part of the course, since the participants were themselves practically oriented. Further, effort was spent to identify topics of significance, which were both within the zone of proximal development of the audience and of practical use in their professional lives.

Deep versus Superficial Learning Säljö makes a clear distinction between deep and superficial learning. Deep learning is more persistent and more easily extendable. However, it requires more of the learning process. Exemplifications, learning by solving problems and learning by teaching each other are known methods to achieve a depth of learning. Lectures and text books are rich in information and provide good referencing material, while they risk to result in more superficial learning. To address this, the course was given a practical problem focus. In the interest of time, some material was presented by means of traditional lectures.

4.3 Structure

In order to relate to the practical background of the audience, all teaching was strongly coupled to laboratory exercises. One hour lectures were followed by hands-on sessions in the lab, where the theoretical results were applied to a physical plant.

There was also a course book, Hägglund (2008), covering the material on a conceptual level. The book was not used extensively during the course itself, but was given to the participants to keep for future reference. In addition to taking notes, participants were strongly encouraged to print plots of experimental result, which was possible to do in an uncomplicated way due to support in the lab user interface.

5. THE LABORATORY EQUIPMENT

In this section, the laboratory process is introduced. A physical overview of the process is followed by a presentation of its dynamics. Finally, the choice of the particular process for the course is given.

5.1 Process Overview

The equipment chosen for the course was a cascaded double tank. The process was developed at the Department of Automatic Control, Lund University, and is used regularly in the basic undergraduate, nonlinear, predictive and process control courses. An earlier version of the process is described in Åström and Östberg (1986). An operational

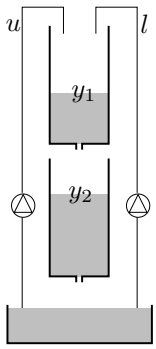
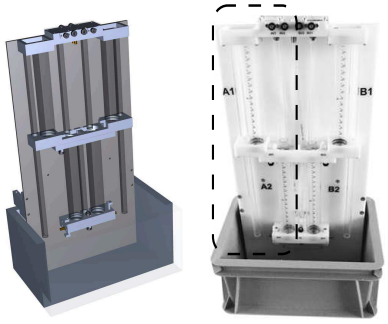


Fig. 2. Operational sketch of double tank process.



(a) CAD drawing. (b) Photograph.

Fig. 3. CAD drawing and photograph of laboratory process. Region corresponding to operational sketch is marked by dashed line in photo.

sketch of the process is shown in Figure 2. Figures 3(a) and 3(b) show a CAD drawing and photograph of the physical process. Note that only the leftmost half, indicated by dashed lines in the photo, was used.

5.2 Dynamics

Open Loop Dynamics Deriving the nonlinear tank dynamics based on Bernoulli (1738) and linearizing them around a stationary point is part of the introductory undergraduate control course at Lund University, giving a good connection between theory and practice. They are given here for completeness.

$$\begin{aligned}\frac{dy_1}{dt} &= -\frac{a_1}{A_1}\sqrt{2g\beta h_1} + \frac{\alpha}{A_1}\theta(u+l)(u+l),\\ \frac{dy_2}{dt} &= \frac{a_1}{A_1}\sqrt{2\beta g h_1} - \frac{a_2}{A_2}\sqrt{2g\beta y_2},\end{aligned} \tag{1}$$

where u is the input flow and l an input disturbance. The water level of the upper tank is y_1, while y_2 is the level of the lower tank. A_k are the tank cross sections and a_k are the cross sections of the holes connecting the tanks. The acceleration of gravity is denoted g. α_k, β_k are unit conversion constants. Finally, θ is the Heaviside step function, manifesting that the tanks cannot be emptied by means of the pumps.

Actuation Linearization Two cheap centrifugal bilge pumps actuate the input u and 'load disturbance' l, respectively. The system input $u$is pump voltage, while the input of the model (1) is flow q. The pump dynamics from u to q are approximately $\sqrt{\cdot}$, with stochastic deviations

caused by mode jumps. For pedagogical reasons it was decided to hide this nonlinearity by closing a flow PI loop over each pump and a corresponding upstream Venturi flow sensor, as shown in Figure 4.

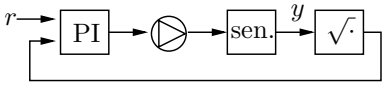


Fig. 4. Flow control loop layout ('sen.' denotes 'sensor').

As shown in Figure 5, the nonlinear voltage to flow characteristics were replaced by linear flow reference to flow ones. The PI loop of Figure 4 was tuned a magnitude faster than the open loop tank dynamics, hiding its dynamical behavior.

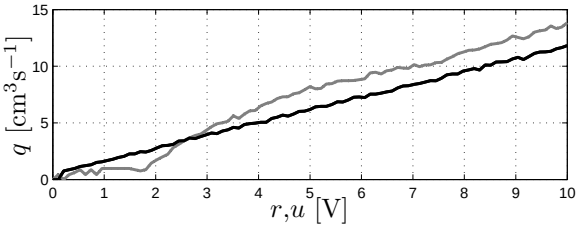


Fig. 5. Open and closed loop pump characteristics. Flow q is plotted against pump voltage u (gray) and PI reference r (black).

5.3 Interface

The process has an on-board micro controller, handling sensor A/D conversions, actuator D/A conversions, execution of the pump linearization control loop and serial port communication with a PC.

Using serial port (or USB to serial) enables the process to be used with all major operating systems.

The PC side interface can be implemented in various ways. A Java interface with graphical windows similar to those in a process industry control room is used in the basic course and shown in Figure 6. Other courses facilitate a Matlab/Simulink interface through locally developed communication blocks. Real time simulation is enabled by the TrueTime real time kernel Cervin and Årzén (2009).

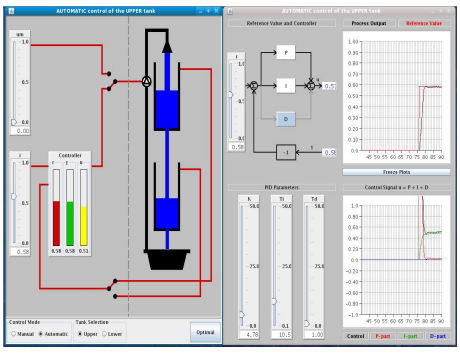


Fig. 6. Java GUI used in the basic undergraduate control course.

The main advantage of using a tailored high level interface, such as the Java one, lies in its flexibility in terms of graphical user interface (GUI). On the other hand, it is less transparent and hence structural changes may require substantial efforts.

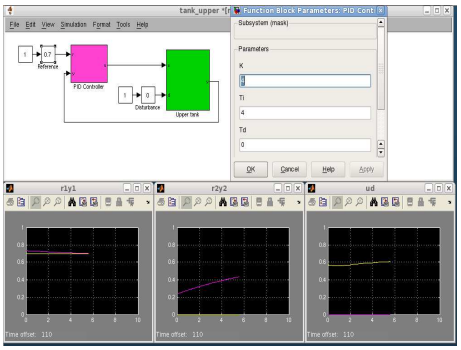


Fig. 7. Simulink GUI. (The PID parameter window is opened by double clicking on the PID block.)

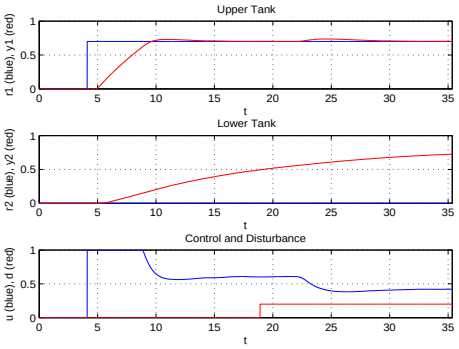


Fig. 8. Comprehensive printout of signals from the latest experiment.

While the Simulink model is not as flexible in terms of GUI, it is straight forward to introduce structural changes. Also, Matlab scripts can be used to set parameters, run simulations and plot data. These features are frequently exploited by students in advanced undergraduate courses at the department.

For the particular course, the GUI was implemented in Simulink. An effort was made to abstract away all technical detail and only present what was necessary to illustrate a given concept. Each experiment was associated with a tailored model, which was opened simply by writing the model name at the Matlab prompt. One of these interfaces are shown in Figure 7. Experiments could be started or stopped at any time. By typing a simple command a paper printout, as the one shown in Figure 8 could be obtained.

5.4 Motivation of Process Choice

The process was chosen due to several facts, in coherence with guidelines from Bencomo (2004), Feisel and Rose (2005) and Balchen et al. (1981). Appealing features include:

- Intuitive, but not trivial, dynamics
- Suitable time scales
- Visual and audible feedback
- Easy to generate load disturbances and measurement noise
- Relevant in process industry (buffer tank)
- Relatively cheap

The dynamics enable the demonstration of concepts such as model based controller tuning, disturbance feed forward, cascaded control and gain scheduling.

In the undergraduate curriculum additional features of the process are explored in greater detail:

- Nonlinear dynamics
- Asymmetric actuation
- Sampled system (zero order hold AD/DA, anti aliasing, etc.)
- Embedded micro controller and real time communications
- Easily extendable to MIMO (four tanks on physical process)
- Model uncertainties in terms of structure and parameters

These features can be bought to attention also during a course for industry practitioners. However, they are not as essential for the practitioner as the topics of the experiments, presented in the following section.

6. SUGGESTED EXPERIMENTS

Below follows a brief description, together with objective and learning outcome, of a set of experiments. Each experiment demonstrates a concept, significant for the process industry practitioner. Together they form the laboratory part of the suggested course.

6.1 Intuition Based Tuning

The first laboratory session was devoted to getting familiar with the process by means of self-designed open loop experiments. In addition, participants were asked to design controllers for the levels in the upper and lower tank, respectively.

The objective was to get the participants familiar with the equipment and dynamics, rather than producing a well tuned loop. Before moving on, the participants should have gained an intuitive understanding that a two-capacitive process is harder to control than a single capacitance. They will have experienced the need of systematic tuning methods, when process dynamics are slow. Those not having a recent experience with PID tuning would in addition obtain a conceptual understanding of the PID parameters and their influence on loop performance.

Experimenting with P, PI and PID controllers, the participants would see that pure P control leaves a static error, and that derivative action can increase performance when controlling the lower tank level, but not the upper. These results are all covered in the lectures. Finally, by choosing the amount of low pass filtering on the measurement signal, the influence of noise on the control signal, when using derivative action, was studied.

6.2 Model-based Tuning

Generally the participants managed to tune an acceptable PI loop for the upper tank within minutes. However, the slow dynamics of the lower tank posed a harder challenge. The process was harder to control per se, and the dominating time constant of ≈ 30 s rendered tedious experiments necessary.

Having learnt the step response method of Ziegler and Nichols Ziegler and Nichols (1942) and a more recent alternative Åström and Hägglund (2006), the participants were encouraged to make step response experiments and identify first order plus time delay model. These simple first order model with time constant T, delay L and static gain K were obtained visually from the input–output data plots.

The objective of this exercise was to demonstrate the practical use of simple model-based tuning methods. It also provided a natural opportunity to emphasize the importance of tuning with respect to load disturbances, rather than reference tracking, since most industrial processes operate with constant reference. Most participant found that model based tuning provides a good starting point for further manual tuning, especially if process dynamics are slow (as in the case of the lower tank).

6.3 Disturbance Feed Forward

Feed forward from a measurable disturbance to control signal, is a simple technique, which can reduce the influence of measurable load disturbances significantly. Ideally, one would use a dynamic link, based on the disturbance path model. However, in many cases, even a static link can provide significant performance improvement. In this experiment, the pump generating the signal l in Figure 2 was used to generate a step input disturbance. The participants are to choose the feed forward gain F (initially $F = 0$) in Figure 9 and investigate its influence when controlling the level of the upper tank. After having conducted this experiment, it should lie closer at hand to add feed forward sensors and compensators, where beneficial.

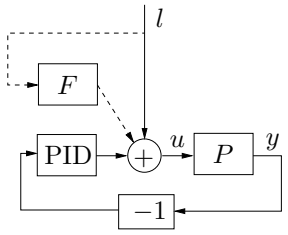


Fig. 9. Block diagram of the disturbance feed forward experiment.

6.4 Cascaded Control

In this experiment, a cascade control solution, consisting of two PID loops was investigated. The setup is shown in Figure 10. Rather than controlling the lower tank level, using only y_2, a cascaded solution, using both y_1 and y_2 was used. The inner loop, closed from y_1, had a dominating time constant considerably faster than that of the outer loop. By tuning a tight inner loop (using the switch), the participants were able to treat the inner loops as a static gain, while later tuning the outer loop.

The primary objective of this exercise was to demonstrate that more advanced control structures, such as cascades, can bring performance improvements at a low cost. It also

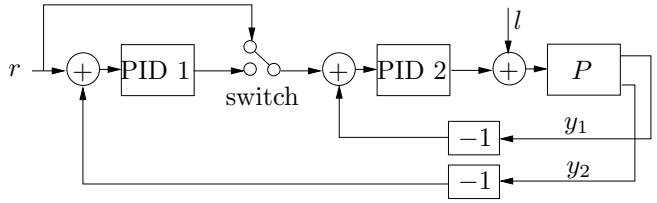


Fig. 10. Block diagram of the cascaded control experiment.

provides background for a continued discussion on sensor and actuator placement.

The participants should be able to identify when cascaded control is useful and be able to tune two cascaded PID loops, starting with the inner one. By experimenting with disturbance steps, l, the participants will find that the cascade suppressed them more efficiently than does a single loop PID solution, using only y_2.

7. OUTCOME

All along the course the instructors had discussions with the participants about the course material, both from a content perspective and from a teaching/learning perspective. This was done in order to assure that the course contained relevant material for the participants (e.g. not to hard and not to easy) and to assure that the participants appreciated the way the course was given (e.g. how much time to be spent on theory and how much time to be spent on the laboratory exercises). Relevant and possible adjustments were done during the execution of the course. At the end of the course, the participants were also asked to fill out an evaluation form with feedback to the instructors about the course. Getting feedback from the participants at the end of the course is important, not for this course itself, but because this course (as well as other courses) will be given to industry practitioners again. The four instructors involved in the course had a meeting shortly after the end of the course in which they gave their own reflections on the course and worked through the feedback from the participants.

7.1 Instructor Evaluation

The four instructors involved were satisfied with the course. Some reflections from the instructors: 'the questions asked by the audience were different from the ones we usually get from engineering students in the way that they were often more practical oriented and less theoretical', and 'The course was designed not to contain too much material, this left room for reflections – I feel the students went away with a good understanding of the material'.

7.2 Participant Evaluation

Some reflections from the participants: 'large and complex area described in an easy and understandable way', 'advance quicker in the start of the course and spend more time on the harder problems at the end', 'good atmosphere, good practical exercises', 'suitable level and useful material'. In the questionnaire the participants should formulate what they had expected of the course and how well the course matched their expectations; 94% felt that the course matched their expectations in a good or very good way, 6% said ok.

8. CONCLUSIONS

The professional role and background of the audience were presented in Section 2. Section 3 investigated the needs of further education of process industry practitioners. The goals and structure of a course aimed at providing continued education are given in Section 4. Large parts of the suggested course are based on laboratory experiments, reviewed in Sections 5, 6. Experience from the course were given in Section 7.

9. ACKNOWLEDGEMENT

This work has been done within PICLU, the Process Industrial Centre at Lund University, Sweden. A thank should go to all participants of the industrial control course as well as all members of PICLU. We would further like to thank Research Engineers Rolf Braun and Anders Blomdell at the Department of Automatic Control, Lund University. Rolf has constructed the hardware and Anders have programmed the embedded computer.

REFERENCES

(2010). Process industrial centre at lund university. `http://www.pic.lu.se`.

Åström, K.J. and Hägglund, T. (2006). *Advanced PID Control*. ISA.

Åström, K.J. and Lundh, M. (1992). Lund control program combines theory with hands-on experience. *IEEE Control Systems Magazine*, 12(3), 22–30.

Åström, K.J. and Östberg, A.B. (1986). A teaching laboratory for process control. *IEEE Control Systems Magazine*, 6, 37–42.

Balchen, J.G., Handlykken, M., and Tysso, A. (1981). The need for better laboratory experiments in control engineering education. In *Proc. 8th IFAC World Congress*.

Bencomo, S.D. (2004). Control learning: present and future. *Annual Reviews in Control*, 28(1), 115 – 136.

Bernoulli, D. (1738). *Hydrodynamica*.

Biggs, J. and Tang, C. (2007). *Teaching for Quality Learning at University*.

Cervin, A. and Årzén, K.E. (2009). TrueTime: Simulation tool for performance analysis of real-time embedded systems. In G. Nicolescu and P.J. Mosterman (eds.), *Model-Based Design for Embedded Systems*. CRC Press.

Desborough, L. and Miller, R. (2002). Increasing customer value of industrial control performance monitoring – Honeywell's experience. *AiChe Symposium Series*, 326, 169–189.

Dzindolet, M.T., Peterson, S.A., Pomranky, R.A., Pierce, L.G., and Beck, H.P. (2003). The role of trust in automation reliance. *International Journal of Human-Computer Studies*, 58(6), 697 – 718.

Ender, D.B. (1993). Process control performance: Not as good as you think. *Control Engineering*, 40, 180.

Feisel, L.D. and Rose, A.J. (2005). The role of the laboratory in undergraduate engineering education. *Journal of Engineering Education*.

Garca, C.E., Prett, D.M., and Morari, M. (1989). Model predictive control: Theory and practice–a survey. *Automatica*, 25(3), 335 – 348.

Hägglund, T. (2008). *Praktisk processreglering*. Studentlitteratur, Lund, Sweden.

Kheir, N.A., strm, K.J., Auslander, D., Cheok, K.C., Franklin, G.F., Masten, M., and Rabins, M. (1996). Control systems engineering education. *Automatica*, 32(2), 147 – 166.

Marton, F. and Säljö, R. (1976). On qualitative differences in learning - i: Outcome and process. *British Journal of Educational Psychology*, 46, 4–11.

Panagopoulos, H. (2000). *PID-Control. Design, Extension, Application*. Ph.D. thesis, Department of Automatic Control, Lund Institute of Technology, Sweden.

Skogestad, S. (2004). Control structure design for complete chemical plants. *Computers & Chemical Engineering*, 28(1-2), 219 – 234. Escape 12.

Vygotsky, L.S. (1978). *Mind in Society: Development of Higher Psychological Processes*.

Ziegler, J.G. and Nichols, N.B. (1942). Optimum settings for automatic controllers. *Transaction fo the ASME*, 759–768.

INDUSTRIAL APPLICATIONS OF PREDICTIVE FUNCTIONAL CONTROL

J.Richalet

1) Back to Basic

Elementary presentation of PFC
Internal model .Reference trajectory. Feed forward action. Tuning by the desired Closed Loop Time Response. Future Manipulated variable projected on a Polynomial Basis with no follow-up error on polynomial set points.

2) Implementation

The target is to bring PID users a control technique that could be easily implemented in a classical industrial PLC and able to control processes where PId is not efficient: taking into account long time delays, constraints on manipulated and process variables, feed forward actions, non stationary set points.etc..

3) Internal model

The internal model can be a classical discrete difference equation or a convolution representation.

4) Procedure

Training in Technical schools and Fachhoschule of industrial technical staff.

5) Recent industrial applications

a) Level control in steel continuous casting :control with constraints and feed forward. Implementation of a complex algebra (real /imaginary variable) to eradicate the harmonic and critical **bulging** effect.
b) Pharmaceutical and chemical industry
Cascaded control of batch reactor with non stationary set points, to be followed with **no** lag error like in a Mullin crystallization.

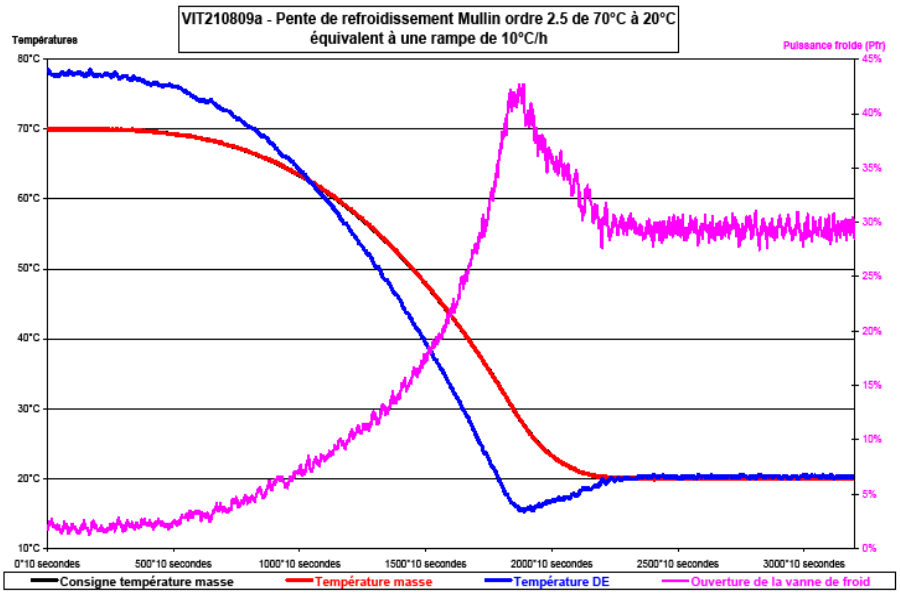


Fig 1 Crystallization : Cubic setpoint (red) followed with “no error”

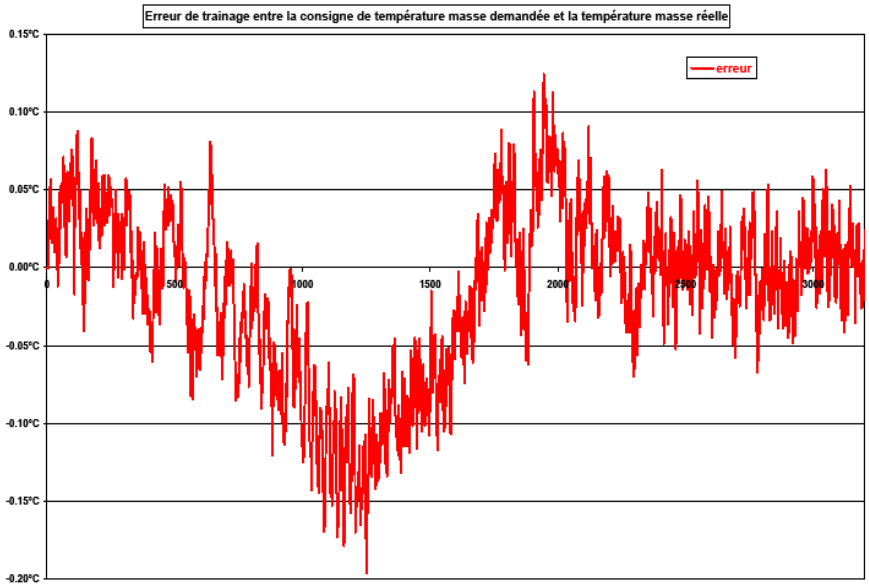


Fig 2 Error with a +/-0.4°channel from 70° to 20°

4° Conclusion :

- On line global estimator of model mismatch and external disturbance
- HVAC control of buildings
- Training of teachers in technical schools directly connected to industry

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

The SIMC method for smooth PID controller tuning

Sigurd Skogestad and Chriss Grimholt

Abstract The SIMC method for PID controller tuning (Skogestad 2003) has already found widespread industrial usage in Norway. This chapter gives an updated overview of the method, mainly from a user's point of view. The basis for the SIMC method is a first-order plus time delay model, and we present a new effective method to obtain the model from a simple closed-loop experiment. An important advantage of the SIMC rule is that there is a single tuning parameter (τ_c) that gives a good balance between the PID parameters (K_c, τ_I, τ_D), and which can be adjusted to get a desired trade-off between performance ("tight" control) and robustness ("smooth" control). Compared to the original paper of Skogestad (2003), the choice of the tuning parameter τ_c is discussed in more detail, and lower and upper limits are presented for tight and smooth tuning, respectively. Finally, the optimality of the SIMC PI rules is studied by comparing the performance (IAE) versus robustness (M_s) trade-off with the Pareto-optimal curve. The difference is small which leads to the conclusion that the SIMC rules are close to optimal. The only exception is for pure time delay processes, so we introduce the "improved" SIMC rule to improve the performance for this case.

Chapter for PID book (planned: Springer, 2011, Editor: R. Vilanova)
This version: September 7, 2011

Sigurd Skogestad
Department of Chemical Engineering, Norwegian University of Science and Technology (NTNU), Trondheim, e-mail: skoge@ntnu.no

Chriss Grimholt
Department of Chemical Engineering, Norwegian University of Science and Technology (NTNU), Trondheim

1 Introduction

Although the proportional-integral-derivative (PID) controller has only three parameters, it is not easy, without a systematic procedure, to find good values (settings) for them. In fact, a visit to a process plant will usually show that a large number of the PID controllers are poorly tuned. The tuning rules presented in this chapter have developed mainly as a result of teaching this material, where there are several objectives:

1. The tuning rules should be well motivated, and preferably model-based and analytically derived.
2. They should be simple and easy to memorize.
3. They should work well on a wide range of processes.

In this paper the simple two-step SIMC procedure (Skogestad 2003) that satisfies these objectives is summarized:

Step 1. Obtain a first- or second-order plus delay model.
Step 2. Derive model-based controller settings. PI-settings result if we start from a first-order model, whereas PID-settings result from a second-order model.

The SIMC method is based on classical ideas presented earlier by Ziegler and Nichols (1942), the IMC PID-tuning paper by Rivera *et al.* (1986), and the closely related direct synthesis tuning rules in the book by Smith and Corripio (1985). The Ziegler-Nichols settings result in a very good disturbance response for integrating processes, but are otherwise known to result in rather aggressive settings (Tyreus and Luyben 1992) (Astrom and Hagglund 1995), and also give poor performance for processes with a dominant delay. On the other hand, the analytically derived IMC-settings of Rivera *et al.* (1986) are known to result in poor disturbance response for integrating processes (Chien and Fruehauf 1990), (Horn *et al.* 1996), but are robust and generally give very good responses for setpoint changes. The SIMC tuning rule presented in this chapter works well for both integrating and pure time delay processes, and for both setpoints and load disturbances.

This chapter provides a summary of the original SIMC method and provides some new results on obtaining the model from closed-loop data, and on the Pareto-optimality of the SIMC method. There is some room for improvement for delay-dominant processes, and at the end of the chapter "improved" SIMC rules are presented.

Notation. The notation is summarized in Figure 1. Here u is the manipulated input (controller output), d the disturbance, y the controlled output, and y_s the setpoint (reference) for the controlled output. $g(s) = \frac{\Delta y}{\Delta u}$ denotes the process transfer function and $c(s)$ is the feedback part of the controller. Note that all the variables u, d and y are deviations from the initial steady state, but the Δ used to indicate deviation variables is usually omitted. Similarly, the Laplace variable s is often omitted to simplify notation. The settings given in this chapter are for the series (cascade, "interacting") form PID controller:

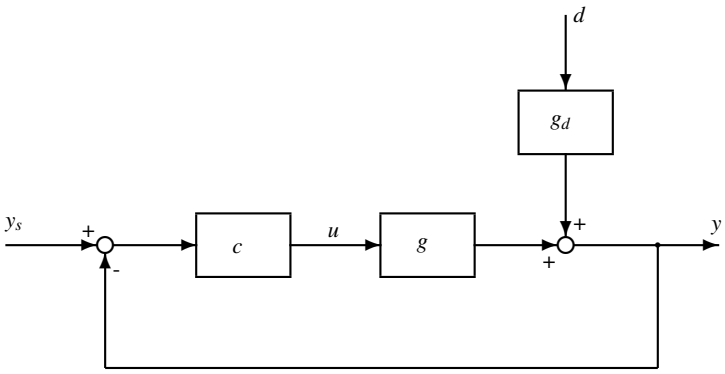


Fig. 1 Block diagram of feedback control system.
In this chapter we consider an input ("load") disturbance ($g_d = g$).

$$\text{Series PID}: \quad c(s) = K_c \cdot \left(\frac{\tau_I s + 1}{\tau_I s}\right) \cdot (\tau_D s + 1) = \frac{K_c}{\tau_I s}\left(\tau_I \tau_D s^2 + (\tau_I + \tau_D)s + 1\right) \tag{1}$$

where K_c is the controller gain, τ_I the integral time, and τ_D the derivative time. The reason for using the series form is that the PID rules with derivative action are then much simpler. The corresponding settings for the ideal (parallel form) PID controller are easily obtained using (30).

Simulations. The following practical PID controller (series form) is used in the simulations:

$$u(s) = K_c \left(\frac{\tau_I s + 1}{\tau_I s}\right)\left(y_s(s) - \frac{\tau_D s + 1}{(\tau_D/N)s + 1} y(s)\right) \tag{2}$$

with $N = 10$. Note that we in order to avoid "derivative kick" do not differentiate the setpoint in (2). In most cases we use PI-control, i.e. $\tau_D = 0$, and the above implementation issues and differences between series and ideal form do not apply.

2 Model approximation (Step 1)

The first step in the SIMC design procedure is to obtain an approximate first- or second-order time delay model on the form

$$g_1(s) = \frac{k}{\tau_1 s + 1} e^{-\theta s} = \frac{k'}{s + 1/\tau_1} e^{-\theta s} \tag{3}$$

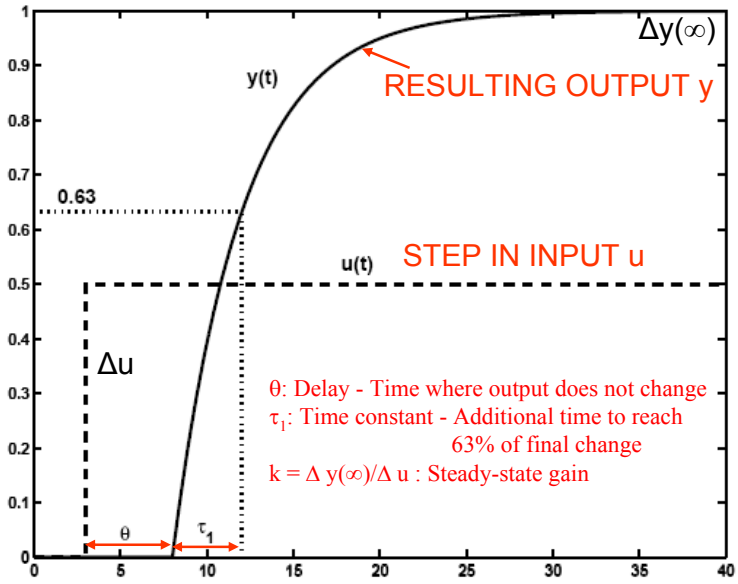


Fig. 2 Open-loop step response experiment to obtain parameters k, τ_1 and θ in first-order model (3)

$$g_2(s) = \frac{k}{(\tau_1 s + 1)(\tau_2 s + 1)} e^{-\theta s} \tag{4}$$

Thus, we need to estimate the following model information

- Plant gain, k
- Dominant lag time constant, τ_1
- (Effective) time delay (dead time), θ
- Optional: Second-order lag time constant, τ_2 (for dominant second-order process for which $\tau_2 > \theta$, approximately)

Such data may be obtained in many ways, three of which are discussed below.

1. From open-loop step response
2. From closed-loop setpoint response with P-controller
3. From detailed model: Approximation of effective delay using the half rule

2.1 Model from open-loop step response

In practice, the model parameters for a first-order model are commonly obtained from a step response experiment as shown in Figure 2. From a theoretical point of view this may not be the most effective method, but it has the advantage of being very simple to use and interpret.

For plants with a large time constant τ_1, one has to wait a long time for the process to settle. Fortunately, it is generally not necessary to run the experiment for longer than about 10 times the effective delay (θ). At this time, one may simply stop the experiment and either extend the response "by hand" towards settling, or approximate it as an integrating process (see Figure 3),

$$\frac{ke^{-\theta s}}{\tau_1 s+1} \approx \frac{k'e^{-\theta s}}{s} \tag{5}$$

where

- Slope, $k' \stackrel{\text{def}}{=} k/\tau_1$

is the slope of the integrating response. The reason is that for lag-dominant processes, i.e. for $\tau_1 > 8\theta$ approximately, the individual values of the time constant
r ratio k'
d below.

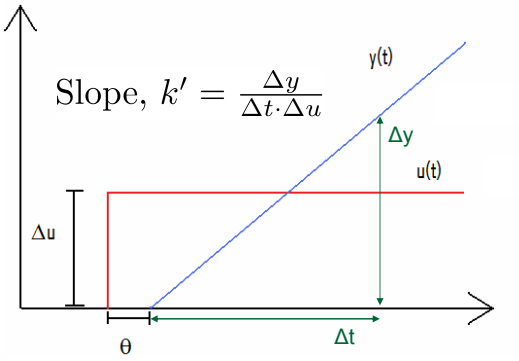


Fig. 3 Open-loop step response experiment to obtain parameters k' and θ in integrating model (5).

2.2 Model from closed-loop setpoint response

In some cases, open-loop responses may be difficult to obtain, and using closed-loop data may be more effective. The most famous closed-loop experiment is the

Ziegler-Nichols where the system is brought to sustained oscillations by use of a P-only controller. One disadvantage with the method is that the system is brought to its instability limit. Another disadvantage is that it does not work for a simple second-order process. Finally, only two pieces of information are used (the controller gain K_u and the ultimate period P_u), so the method cannot possibly work on a wide range of first-order plus delay processes, which we know are described by three parameters (k, τ_1, θ).

Yuwana and Seborg (1982), and more recently Shamsuzzoha and Skogestad (2010), proposed a modification to the Ziegler-Nichols closed-loop experiment, which does not suffer from these three disadvantages. Instead of bringing the system to its limit of stability, one uses a P-controller with a gain that is about half this value, such that the resulting overshoot (D) to a step change in the setpoint is about 30% (that is, D is about 0.3).

We here describe the procedure proposed by Shamsuzzoha and Skogestad (2010)

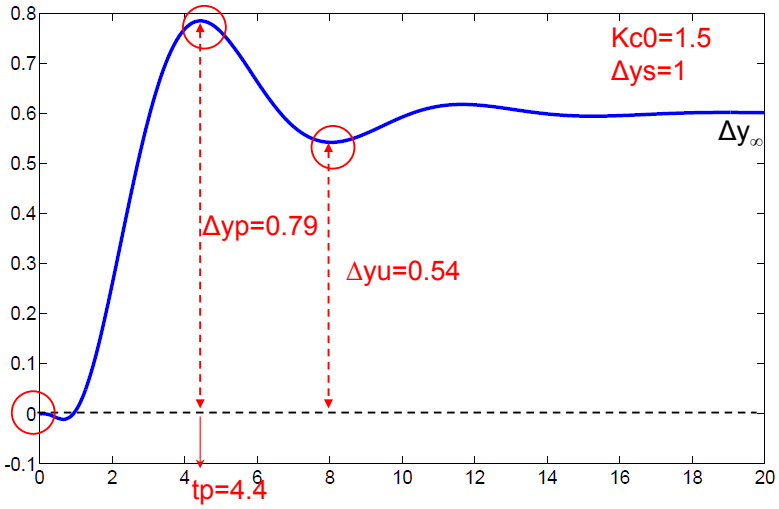


Fig. 4 Extracting information from closed-loop setpoint response with P-only controller.

- Controller gain used in experiment, K_{c0}
- Setpoint change, Δy_s.
- Time from setpoint change to reach first (maximum) peak, t_p.
- Corresponding maximum output change, Δy_p.
- Output change at first undershoot, Δy_u.

This seems to be the information that is most easy (and robust) to observe directly, without having to record and analyze all the data before finding the parameters. Also note that one may stop the experiment already at the first undershoot.

The undershoot Δy_u is used to estimate the steady-state output change (at infinite time)(Shamsuzzoha and Skogestad 2010),

$$\Delta y_\infty = 0.45(\Delta y_p + \Delta y_u) \tag{6}$$

Alternatively, if one has time to wait for the experiment to settle, one may record Δy_∞ instead of Δy_u.

From this information one computes the relative overshoot and the absolute value of the relative steady-state offset, defined by:

- Overshoot, $D = \frac{\Delta y_p - \Delta y_\infty}{\Delta y_\infty}$.
- Steady-state offset, $B = \left| \frac{\Delta y_s - \Delta y_\infty}{\Delta y_\infty} \right|$.

Shamsuzzoha and Skogestad (2010) use this information to obtain directly the PI settings. Alternatively, we may use a two-step procedure, where we first from K_{c0}, D, B and t_p obtain estimates for the parameters in a first-order plus delay model (see the Appendix for details). We compute the parameters

$$A = 1.152D^2 - 1.607D + 1$$

$$r = 2A/B$$

and we obtain the following first-order plus delay model parameters from the closed-loop setpoint response (Figure 4):

$$k = 1/(K_{c0}B) \tag{7}$$

$$\theta = t_p \cdot (0.309 + 0.209e^{-0.61r}) \tag{8}$$

$$\tau_1 = r\theta \tag{9}$$

These values may subsequently be used with any tuning method, for example, the SIMC PI rules. The closed-loop method may also be used for an unstable process, provided it can be approximated reasonably well by a stable first-order process. The extension to unstable processes is the reason for taking the absolute value when obtaining the steady-state offset B.

Example E2(Skogestad 2003). *For the process*

$$g_0(s) = \frac{(-0.3s+1)(0.08s+1)}{(2s+1)(1s+1)(0.4s+1)(0.2s+1)(0.05s+1)^3}$$

the closed-loop setpoint response with P-only controller with gain $K_{c0} = 1.5$ is shown in Figure 4. The following data is obtained from the closed-loop response

$$K_{c0} = 1.5, \Delta y_s = 1, \Delta y_p = 0.79, t_p = 4.4, \Delta y_u = 0.54$$

and we compute

$$\Delta y_\infty = 0.5985, D = 0.32, B = 0.67, A = 0.6038, r = 1.80$$

which using (7) - (9) gives the following first-order with delay model approximation,

$$k = 0.994, \theta = 1.67, \tau_1 = 3.00 \tag{10}$$

This gives a good approximation of the open-loop step response, as can seen by comparing the curves for g_0 and g_{cl} in Figure 5. The approximation is certainly not the best possible, but it should be noted that the objective is to use the model for tuning, and the resulting difference in the tuning, and thus closed-loop response, may be smaller than it appears by comparing the open-loop responses.

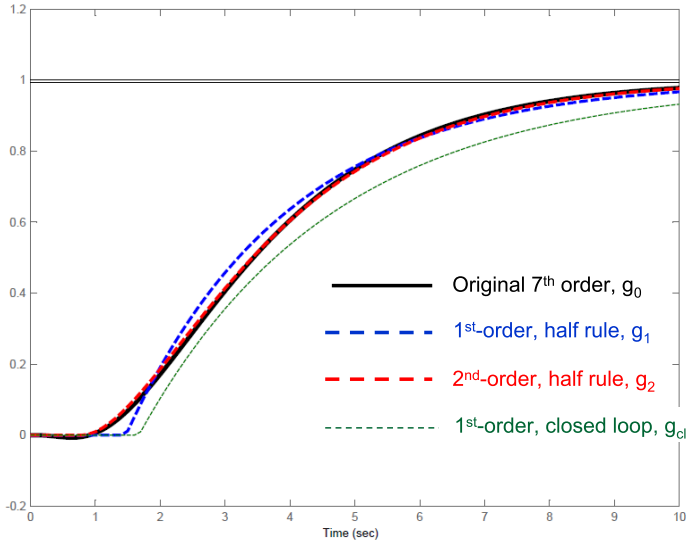


Fig. 5 Open-loop response to step change in input u for process E2, $g_0(s) = \frac{(-0.3s+1)(0.08s+1)}{(2s+1)(1s+1)(0.4s+1)(0.2s+1)(0.05s+1)^3}$ (solid line), and comparison with various approximations.

2.3 Approximation of detailed model using half rule

Assume that we have a given detailed transfer function model in the form

$$g_0(s) = \frac{\prod_j (-T_{j0}^{\text{inv}} s + 1)}{\prod_i (\tau_{i0} s + 1)} e^{-\theta_0 s} \tag{11}$$

where all the given parameters are positive and the time constants are ordered according to their magnitudes. To approximate this with a first or second-order time delay model, (3) or (4), Skogestad (2003) recommends that the "effective delay" θ is taken as the "true" delay θ_0, plus the inverse response (negative numerator) time constant(s) T^{inv}, plus half of the largest neglected time constant (half rule), plus all smaller time constant τ_{i0}. The "other half" of the largest neglected time constant is added to get at larger time constant τ_1 (or τ_2 for a second-order model).

Half rule: The largest neglected (denominator) time constant (lag) is distributed evenly to the effective delay (θ) and the smallest retained time constant (τ_1 or τ_2).

In summary, for a model in the form (11), to obtain a first-order model (3) we use

$$\tau_1 = \tau_{10} + \frac{\tau_{20}}{2}; \quad \theta = \theta_0 + \frac{\tau_{20}}{2} + \sum_{i \geq 3} \tau_{i0} + \sum_j T_{j0}^{\text{inv}} + \frac{h}{2} \tag{12}$$

and, to obtain a second-order model (4), we use

$$\tau_1 = \tau_{10}; \quad \tau_2 = \tau_{20} + \frac{\tau_{30}}{2}; \quad \theta = \theta_0 + \frac{\tau_{30}}{2} + \sum_{i \geq 4} \tau_{i0} + \sum_j T_{j0}^{\text{inv}} + \frac{h}{2} \tag{13}$$

where h is the sampling period (for cases with digital implementation).

Example E1. *Using the half rule, the process*

$$g_0(s) = \frac{1}{(s+1)(0.2s+1)}$$

is approximated as a first-order time delay process, $g(s) = ke^{-\theta s+1}/(\tau_1 s+1)$*, with* $k = 1, \theta = 0.2/2 = 0.1$ *and* $\tau_1 = 1 + 0.2/2 = 1.1$.

Example E2 (continued). *Using the half rule, the process*

$$g_0(s) = \frac{(-0.3s+1)(0.08s+1)}{(2s+1)(1s+1)(0.4s+1)(0.2s+1)(0.05s+1)^3}$$

is approximated as a first-order time delay process (3) with

$$\tau_1 = 2 + 1/2 = 2.5$$

$$\theta = 1/2 + 0.4 + 0.2 + 3 \cdot 0.05 + 0.3 - 0.08 = 1.47$$

or a second-order time delay process (4) with

$$\tau_1 = 2$$

$$\tau_2 = 1 + 0.4/2 = 1.2$$

$$\theta = 0.4/2 + 0.2 + 3 \cdot 0.05 + 0.3 - 0.08 = 0.77$$

The small positive numerator time constant $T_0 = 0.08$ was subtracted from the effective time delay according to rule T3 (see below). Both approximations, and in particular the second-order model, are very good as can be seen by from the open-loop step responses in Figure 5. Note that with the SIMC tuning rules, a first-order model yields a PI-controller, whereas a second-order model yields a PID controller.

Comment: In this case, we have $\tau_2 > \theta (1.2 > 0.77)$ for the second-order model, and the use of PID control is expected to yield a significant performance improvement compared to PI control (see below for details). However, adding derivative action has disadvantages, such as increased input usage and increased noise sensitivity.

2.4 Approximation of positive numerator time constants

A process model can also contain positive numerator time constants T_0 as the following process:

$$g(s) = g_0(s) \frac{T_0 s + 1}{\tau_0 s + 1} \tag{14}$$

Skogestad (2003) propose to cancel out the numerator time constant T_0 against a "neighboring" lag time constant τ_0 by the following rules: [1]

$$\frac{T_0 s + 1}{\tau_0 s + 1} \approx \begin{cases} T_0/\tau_0 & \text{for } T_0 \geq \tau_0 \geq \tau_c \quad \text{(Rule T1)} \\ T_0/\tau_c & \text{for } T_0 \geq \tau_c \geq \tau_0 \quad \text{(Rule T1a)} \\ 1 & \text{for } \tau_c \geq T_0 \geq \tau_0 \quad \text{(Rule T1b)} \\ T_0/\tau_0 & \text{for } \tau_0 \geq T_0 \geq 5\tau_c \quad \text{(Rule T2)} \\ \frac{(\tilde{\tau}_0/\tau_0)}{(\tilde{\tau}_0 - T_0)s+1} & \text{for } \tilde{\tau}_0 \stackrel{\text{def}}{=} \min(\tau_0, 5\tau_c) \geq T_0 \quad \text{(Rule T3)} \end{cases} \tag{15}$$

Here τ_c is the desired closed-loop time constant, which appears as the tuning parameter in the SIMC PID rules. Because the tuning parameter is normally chosen after obtaining the effective time delay (the recommended value for "tight control" is $\tau_c = \theta$), one may not know this value before the model is approximated. Therefore, one may initially have to guess the value τ_c and iterate.

We normally select τ_0 as the closest *larger* denominator time constant ($\tau_0 > T_0$) and use Rules T2 or T3. Note that an integrating process corresponds to a process with an infinitely large time constant, $\tau_0 = \infty$. For example, for an integrating-pole-

[1] The rules are slightly generalized compared to Skogestad (2003) by replacing θ (effective time delay in final model) by τ_c (desired closed-loop time constant). This makes the rules applicable also to cases where τ_c is selected to be different from θ.

zero (IPZ) process on the form $k'\frac{e^{-\theta s}}{s}\frac{Ts+1}{\tau_2 s+1}$, we get $\frac{Ts+1}{s} \approx T$ (Rule T2 with $\tau_0 = \infty > T$). However, if T is smaller than τ_2 then we may use the approximation $\frac{Ts+1}{\tau_2 s+1} \approx \frac{T}{\tau_2}$ (Rule T2 with $\tau_2 > T > 5\theta$). Rule T3 would apply if T was even smaller.

However, if there exists no larger τ_0, or if there is smaller denominator time constant "close to" T_0, then we select τ_0 as the closest *smaller* denominator time constant ($\tau_0 < T_0$) and use rules T1, T1a or T1b. To define "close to" more precisely, let τ_{0a} (large) and τ_{0b} (small) denote the two neighboring denominator constants to T_0. Then, we select $\tau_0 = \tau_{0b}$ (small) if $T_0/\tau_{0b} < \tau_{0a}/T_0$ and $T_0/\tau_{0b} < 1.6$ (both conditions must be satisfied).

Derivations of the above rules and additional examples are given in (Skogestad 2003).

3 SIMC PI and PID tuning rules (step 2)

In step 2, we use the model parameters $(k, \theta, \tau_1, \tau_2)$ to tune the PID controller. We here derive the SIMC rules and apply them to some typical processes.

3.1 Derivation of SIMC rules

The SIMC rules may be derived using the method of direct synthesis for setpoints (Smith and Corripio 1985), or equivalently the Internal Model Control approach for setpoints (Rivera *et al.* 1986). For the system in Figure 1, the closed-loop setpoint response is

$$\frac{y}{y_s} = \frac{g(s)c(s)}{g(s)c(s)+1} \tag{16}$$

where we have assumed that the measurement of the output y is perfect. The idea of direct synthesis is to specify the desired closed-loop response and solve for the corresponding controller. From (16) we get

$$c(s) = \frac{1}{g(s)} \frac{1}{\frac{1}{(y/y_s)_{\text{desired}}} - 1} \tag{17}$$

We here consider the second-order time delay model $g(s)$ in (4), and specify that we, following the delay, desire a "smooth" first-order response with time constant τ_c

$$\left(\frac{y}{y_s}\right)_{\text{desired}} = \frac{1}{\tau_c s+1} e^{-\theta s} \tag{18}$$

The delay θ is kept in the "desired" response because it is unavoidable. Substituting (18) and (4) into (17) gives a "Smith Predictor" controller (Smith 1957):

$$c(s) = \frac{(\tau_1 s+1)(\tau_2 s+1)}{k} \frac{1}{(\tau_c s+1-e^{-\theta s})} \tag{19}$$

τ_c is the desired closed-loop time constant, and is the sole tuning parameter for the controller. To derive PID settings, we introduce in (19) a first-order Taylor series approximation of the delay, $e^{-\theta s} \approx 1-\theta s$. This gives

$$c(s) = \frac{(\tau_1 s+1)(\tau_2 s+1)}{k} \frac{1}{(\tau_c+\theta)s} \tag{20}$$

which is a series form PID-controller (1) with (Smith and Corripio 1985) (Rivera *et al.* 1986)

$$K_c = \frac{1}{k}\frac{\tau_1}{\tau_c+\theta} = \frac{1}{k'}\frac{1}{\tau_c+\theta}; \quad \tau_I = \tau_1; \quad \tau_D = \tau_2 \tag{21}$$

These settings are derived by considering the setpoint response. However, it is well known that for lag dominant processes with $\tau_1 \gg \theta$ (e.g. integrating processes), the choice $\tau_I = \tau_1$ results in a long settling time for *input ("load") disturbances* (Chien and Fruehauf 1990). To improve the load disturbance response, one may reduce the integral time, but not by too much, because otherwise we get slow oscillations and robustness problems. Skogestad (2003) suggests that a good trade-off between disturbance response and robustness is obtained by selecting the integral time such that we just avoid the slow oscillations, which with the controller gain given in (21) corresponds to

$$\tau_I = 4(\tau_c+\theta) \tag{22}$$

3.2 Summary of SIMC rules (original)

For a **first-order model**

$$g_1(s) = \frac{k}{(\tau_1 s+1)} e^{-\theta s} \tag{23}$$

the SIMC method results in a PI controller with settings

$$K_c = \frac{1}{k}\frac{\tau_1}{\tau_c+\theta} = \frac{1}{k'}\frac{1}{\tau_c+\theta} \tag{24}$$

$$\tau_I = \min\{\tau_1, 4(\tau_c+\theta)\} \tag{25}$$

The desired first-order *closed-loop* time constant τ_c is the only tuning parameter.

For a **second-order model**

$$g_2(s) = \frac{k}{(\tau_1 s+1)(\tau_2 s+1)} e^{-\theta s} \tag{26}$$

the SIMC method results in a PID controller with settings (cascade form)

$$K_c = \frac{1}{k}\frac{\tau_1}{\tau_c+\theta} = \frac{1}{k'}\frac{1}{\tau_c+\theta} \tag{27}$$

$$\tau_I = \min\{\tau_1, 4(\tau_c+\theta)\} \tag{28}$$

$$\tau_D = \tau_2 \tag{29}$$

Again, the desired first-order *closed-loop* time constant τ_c is the only tuning parameter. These PID settings are for the cascade (series) form in (1). The corresponding settings for the ideal (parallel form) PID controller are easily obtained using (30).

PID-control (with derivative action) is primarily recommended for processes with dominant second order-dynamics, defined as having $\tau_2 > \theta$, approximately. We note that the derivative time is then selected so as to cancel the second-largest process time constant.

In Table 1 we summarize the resulting tunings for a few special cases, including the pure time delay process, integrating process, and double integrating process. The double integrating process corresponds to a second-order process with $\tau_2 = \infty$ and direct application of the rules actually yield a PD controller, so in Table 1 integral action has been added to eliminate the offset for input disturbances.

The choice of the tuning parameter τ_c is discussed in more detail below. If the objective is to have "tight control" (good output performance) subject to having good robustness, then the recommendation is to choose τ_c equal to the effective time delay, $\tau_c = \theta$. The same recommendation for τ_c applies to both PI- and PID-control, but the actual values will differ, because the effective delay θ in a first-order model (PI control) will be larger than that in a second-order model (PID control) of a given process.

Example E2 (further continued). *We want to derive PI- and PID-settings for the process*

$$g_0(s) = \frac{(-0.3s+1)(0.08s+1)}{(2s+1)(1s+1)(0.4s+1)(0.2s+1)(0.05s+1)^3}$$

using the SIMC tuning rules with the "default" recommendation $\tau_c = \theta$. From the closed-loop setpoint response, we obtained in a previous example a first-order model with parameters $k = 0.994, \theta = 1.67, \tau_1 = 3.00$ (10). The resulting SIMC PI-settings with $\tau_c = \theta = 1.67$ are

$$K_c = 0.904, \tau_I = 3$$

Process	$g(s)$	K_c	τ_I	$\tau_D^{(5)}$
First-order, eq.(3)	$k\frac{e^{-\theta s}}{(\tau_1 s+1)}$	$\frac{1}{k}\frac{\tau_1}{\tau_c+\theta}$	$\min\{\tau_1, 4(\tau_c+\theta)\}$	-
Second-order, eq.(4)	$k\frac{e^{-\theta s}}{(\tau_1 s+1)(\tau_2 s+1)}$	$\frac{1}{k}\frac{\tau_1}{\tau_c+\theta}$	$\min\{\tau_1, 4(\tau_c+\theta)\}$	τ_2
Pure time delay$^{(1)}$	$ke^{-\theta s}$	0	0 $^{(*)}$	-
Integrating$^{(2)}$	$k'\frac{e^{-\theta s}}{s}$	$\frac{1}{k'}\cdot\frac{1}{(\tau_c+\theta)}$	$4(\tau_c+\theta)$	-
Integrating with lag	$k'\frac{e^{-\theta s}}{s(\tau_2 s+1)}$	$\frac{1}{k'}\cdot\frac{1}{(\tau_c+\theta)}$	$4(\tau_c+\theta)$	τ_2
Double integrating$^{(3)}$	$k''\frac{e^{-\theta s}}{s^2}$	$\frac{1}{k''}\cdot\frac{1}{4(\tau_c+\theta)^2}$	$4(\tau_c+\theta)$	$4(\tau_c+\theta)$
IPZ process$^{(4)}$	$k'\frac{e^{-\theta s}}{s}\frac{Ts+1}{\tau_2 s+1}$	$\frac{1}{k'T}\cdot\frac{\tau_2}{\tau_c+\theta}$	$\min\{\tau_2, 4(\tau_c+\theta)\}$	-

Table 1 SIMC PID-settings (27)-(29) for some special cases of (4) (with τ_c as a tuning parameter).
(1) The pure time delay process is a special case of a first-order process with $\tau_1 = 0$.
(2) The integrating process is a special case of a first-order process with $\tau_1 \to \infty$.
(3) For the double integrating process, integral action has been added according to eq.(22).
(4) For the integrating-pole-zero (IPZ) process we assume $T > \tau_2$. Then $(Ts+1)/s \approx T$ (rule T2) and the PI-settings follow.
(5) The derivative time is for the cascade form PID controller in eq.(1).
(*) Pure integral controller $c(s) = \frac{K_I}{s}$ with $K_I = \frac{K_c}{\tau_I} = \frac{1}{k(\tau_c+\theta)}$.

From the full-order model $g_0(s)$ and the half rule, we obtained in a previous example a first-order model with parameters $k = 1, \theta = 1.47, \tau_1 = 2.5$. The resulting SIMC PI-settings with $\tau_c = \theta = 1.47$ are

$$K_c = 0.850, \tau_I = 2.5$$

From the full-order model $g_0(s)$ and the half rule, we obtained a second-order model with parameters $k = 1, \theta = 0.77, \tau_1 = 2, \tau_2 = 1.2$. The resulting SIMC PID-settings with $\tau_c = \theta = 0.77$ are

$$\text{Cascade PID}: \quad K_c = 1.299, \tau_I = 2, \tau_D = 1.2$$

The corresponding settings with the more common ideal (parallel form) PID controller are obtained by computing $f = 1 + \tau_D/\tau_I = 1.60$ and we have

$$\text{Ideal PID}: \quad K_c' = K_c f = 1.69, \tau_I' = \tau_I f = 3.2, \tau_D' = \tau_D/f = 0.75 \qquad (30)$$

The closed-loop responses for the three controllers to a setpoint change at $t = 0$ and an input (load) disturbance at $t = 10$ is shown in Figure 6. The responses for the two PI controllers are very similar, as expected. The PID controller shows better output performance (upper plot), especially for the disturbance, but it may not be sufficient to outweigh the increased input usage (lower plot) and increased sensitivity to noise (not shown in plot).

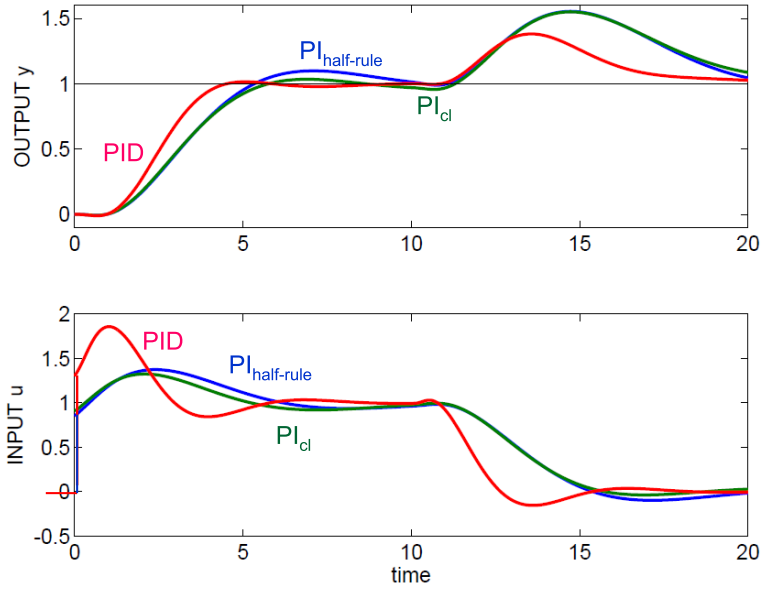


Fig. 6 Closed-loop responses for process E2 using SIMC PI- and PID-tunings with $\tau_c = \theta$. Setpoint change at $t = 0$ and input (load) disturbance at $t = 10$. For the PID controller, D-action is only on the feedback signal, i.e., not on the setpoint y_s.

4 Choice of tuning parameter τ_c

The value of the desired closed-loop time constant τ_c can be chosen freely, but from (27) we must have $-\theta < \tau_c < \infty$ to get a positive and nonzero controller gain. The optimal value of τ_c is determined by a trade-off between:

1. **Output performance (tight control):** Fast speed of response and good disturbance rejection (favored by a small value of τ_c). This "tightness" can be quantified by the magnitude of the setpoint error, $|y(t) - y_s(t)|$, which should be as small as possible. Here, one may consider different "norms" of the error, for example, the maximum deviation (∞-norm), the integrated square deviation (2-norm) and the integrated absolute error (IAE) (1-norm),

$$\text{IAE} = \int_0^\infty |y(t) - y_s(t)| dt$$

2. **Robustness (smooth control):** Good robustness, small input changes and small noise sensitivity (favored by a large value of τ_c). The "smoothness" is here quantified by the peak value $M_s \geq 1$ of the frequency-dependent sensitivity function, $S = 1/(1+gc)$. In terms of robustness, $1/M_s$ is the closest distance of the loop transfer function gc to the critical (-1)-point in the Nyquist diagram, so M_s

should be as small as possible. Notice that $M_s < 1.7$ guarantees gain margin (GM)> 2.43 and phase margin (PM)$> 34.2^o$ (Rivera *et al.* 1986).

In general, we have a multiobjective optimization problem, so there is no value of τ_c which is "optimal". We will consider in more detail the two limiting cases of "tight" and "smooth" control, and also consider in some detail the required input usage.

4.1 Tight control

With tight control, the primary objective is to keep the output close to its setpoint, but there should be some minimum requirement in terms of robustness and smoothness. A good trade-off is obtained by choosing τ_c equal to the time delay:

Tuning parameter τ_c. SIMC-recommendation for "tight control", or more precisely "tighest possible subject to maintaining smooth control":

$$\tau_c = \theta \tag{31}$$

The choice $\tau_c = \theta$ gives a reasonably fast response with moderate input usage and a good robustness with M_s about 1.6 to 1.7. More specifically, the robustness margins with the SIMC PID-settings in (27)-(29) and $\tau_c = \theta$, when applied to first- or second-order time delay processes, are always between the values given by the two columns in Table 2. The values in the left column in Table 2 apply to a case with a relatively small lag time constant (so $\tau_I = \tau_1$), and the somewhat less robust values in the right column apply to an integrating process (so $\tau_I = 4(\tau_c + \theta) = 8\theta$). For the integrating process, we reduce the integral time relative to the original value of $\tau_I = \tau_1$ to get better output performance for load disturbances, and not surprisingly we have to "pay" for this in terms of less robustness.

Process $g(s)$	$\frac{k}{\tau_1 s+1} e^{-\theta s}$	$\frac{k'}{s} e^{-\theta s}$
Controller gain, K_c ($\tau_c = \theta$)	$\frac{0.5}{k} \frac{\tau_1}{\theta}$	$\frac{0.5}{k'} \frac{1}{\theta}$
Integral time, τ_I	τ_1	8θ
Gain margin (GM)	3.14	2.96
Phase margin (PM)	61.4^o	46.9^o
Allowed time delay error, $\Delta\theta/\theta$	2.14	1.59
Sensitivity peak, M_s	1.59	1.70
Complementary sensitivity peak, M_t	1.00	1.30
Phase crossover frequency, $\omega_{180} \cdot \theta$	1.57	1.49
Gain crossover frequency, $\omega_c \cdot \theta$	0.50	0.51

Table 2 "Tight" settings: Robustness margins for first-order and integrating time delay process for SIMC-rules (24)-(25) with $\tau_c = \theta$. The same margins apply to a second-order process (4) if we choose $\tau_D = \tau_2$ in (29).

To be more specific, for processes with a relatively small time constant where we use $\tau_I = \tau_1$ (left column), the system always has a gain margin GM=3.14 and phase margin PM=61.4^o, which is much better than the typical minimum requirements GM$>$ 1.7 and PM$>$ 30^o (Seborg *et al.* 1989). The sensitivity and complementary sensitivity peaks are $M_s = 1.59$ and $M_t = 1.00$ (here small values are desired with a typical upper bound of 2). The maximum allowed time delay error is $\Delta\theta/\theta = \text{PM}\ [\text{rad}]/(w_c \cdot \theta)$, which in this case gives $\Delta\theta/\theta = 2.14$ (i.e., the system goes unstable if the time delay is increased from θ to $(1+2.14)\theta = 3.14\theta$).

For an integrating processes (right column) and $\tau_I = 8\theta$, the suggested "tight" settings give GM=2.96, PM=46.9^o, $M_s = 1.70$ and $M_t = 1.30$, and the maximum allowed time delay error is $\Delta\theta = 1.59\theta$.

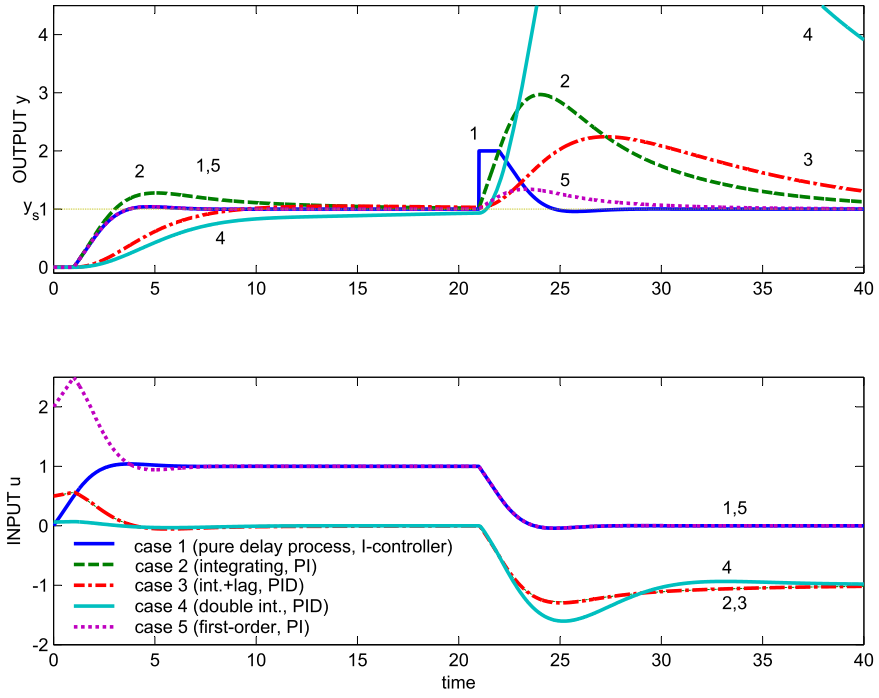


Fig. 7 Responses using SIMC settings for the five time delay processes ($\tau_c = \theta$).
Unit setpoint change at $t = 0$; Unit load disturbance at $t = 20$.
Simulations are without derivative action on the setpoint.
Parameter values: $\theta = 1, k = 1, k' = 1, k'' = 1$.

The simulated time responses to setpoint changes and disturbances with SIMC-settings are shown for five cases in Figure 7 (Skogestad 2003). Even though these are for the "tight" settings ($\tau_c = \theta$), the responses are all smooth. This means that it is certainly possible to get even tighter responses by choosing a smaller value,

for example $\tau_c = 0.5\theta$, but for most process control applications this is not recommended because of less robustness, larger input usage and more sensitivity to noise. It may seem from Figure 7 that the SIMC PID-controller does not work well for the double integrating process (curve 4), but this is a difficult process to control and the response to a unit input disturbance will be large for any robust controller.

4.2 Smooth control

Even though the recommended "tight" settings ($\tau_c = \theta$) gives responses that are reasonably smooth, they may still be unnecessary aggressive compared to the required performance objectives, especially if the effective delay θ is small. For example, for the limiting case with $\theta = 0$ (no delay), we get with $\tau_c = \theta$ an infinite controller gain, which is clearly not realistic. Thus, in practice one often uses a "smoother" tuning, that is, $\tau_c > \theta$.

However, τ_c should not be too large, because otherwise the output y will go out of bound when there are disturbances d. The question is: *How slow (smooth) can we tune the controller and still get acceptable control?* This issue is addressed in the paper by Skogestad (2006) on "tuning for smooth PID control with acceptable disturbance rejection", where the following lower bound on the controller gain is derived (for both PI- and PID-control).

Controller gain. SIMC-recommendation for "smooth control", or more precisely "smoothest possible subject to acceptable disturbance rejection":

$$|K_c| > |K_{c,min}| = \frac{|\Delta u_0|}{|\Delta y_{max}|} \tag{32}$$

where
Δy_{max} = maximum allowed deviation in the output y
Δu_0 = required input change to reject the disturbance(s) d.

Substituting $K_{c,min}$ into (24) or (27) one can obtain the corresponding value $\tau_{c,max}$, and we end up with a region of recommended values for the tuning parameter τ_c:

$$\tau_{c,min} \text{ ("tight")} < \tau_c < \tau_{c,max} \text{ ("smooth")} \tag{33}$$

where

$$\tau_{c,min} = \theta, \quad \tau_{c,max} = \frac{1}{K_{c,min}} \cdot \frac{\tau_1}{k} - \theta \tag{34}$$

The final choice of τ_c is an engineering decision. A small value for τ_c ("tight control" of y) is typically desired for control of active constraints, because tight control reduces the required backoff (safety margin to the constraint). On the other hand, tight control will require larger input changes which may disturb the rest of the pro-

cess. For example, for liquid level there is usually no reason to control the level tightly, so a large value of τ_c ("smooth control") is desired.

Details on the derivation of (32) and $\tau_{c,max}$ are given in (Skogestad 2006), but let us here give a simplified version. Consider disturbance rejection and assume we use a P-only controller with gain K_c. The input change (in deviation from the nominal value) is then $\Delta u = -K_c \Delta y$ or

$$|\Delta u| = |K_c| \cdot |\Delta y|$$

Assume that the required input change to reject a disturbance is Δu_0. For example, if we have a disturbance Δd_I at the input, then $\Delta u_0 = -\Delta d_I$. The smallest controller gain that can generate the required input change Δu_0 is obtained when we have the largest output change ($|\Delta y| = |\Delta y_{max}|$), and we get

$$|\Delta u_0| = |K_{c,min}| \cdot |\Delta y_{max}|$$

and (32) follows.

4.3 Input usage

The magnitude of the dynamic input change can be an important issue when tuning the controller, that is, when selecting the value for τ_c. The transfer function from the disturbance d to the input u is given by (see Figure 1):

$$u(s) = -\frac{g_d c}{1 + gc} d(s)$$

With integral action in the controller (e.g., PI or PID control), the steady-state input change to a step disturbance d is independent of the controller and is given by $u(t = \infty) = -\frac{k_d}{k} d$ where k_d is the steady-state disturbance gain and k is the steady-state process gain. We assume that we can reject the expected disturbances at steady-state, that is, we assume $|u(t = \infty)| = |\frac{k_d}{k} d| \leq |u_{max}|$ where $|u_{max}|$ and $|d|$ is the magnitude of the disturbance change, is the maximum allowed input change, because otherwise the process is not "controllable" (with any controller). However, the dynamic input change $u(t)$ will depend on the controller tuning, and we will consider the initial change (at $t = 0^+$) just after a step disturbance d.

We consider two important disturbances, namely an input "load" disturbance d_u (corresponding to $g_d = g$), and an output disturbance d_y (corresponding to $g_d = 1$). Note that an output disturbance has an immediate effect on the output y. A physical example is a process where we add another stream (output disturbance) just before the measurement y. Mathematically, an output disturbance is equivalent to a setpoint change (with $y_s = -d_y$)

For an *input ("load") disturbances* d_u, input usage is not an important issue for SIMC-tuning, even dynamically. This is because the SIMC controller gives a

closed-loop transfer function $\frac{y}{y_s} = \frac{gc}{1+gc}$ with little or no overshoot, see (16) and (18), and since $\frac{u}{d_u} = -\frac{gc}{1+gc}$, we get for d_u a corresponding input response with little overshoot. This is illustrated by the input changes for a load disturbance ($t = 20$) in Figure 7.

On the other hand, for an *output disturbances* d_y ($g_d = 1$), or equivalently for a *setpoint change* $y_s = -d_y$, input usage may be an important issue for tuning. The steady-state input change to a step setpoint change y_s is $u(t = \infty) = \frac{1}{k} y_s$. However, with PI-control the input will initially jump to the value $u(t = 0^+) = K_c y_s$, as illustrated for the setpoint change in Figure 7 (e.g., see the first-order process, case 5). This initial change is larger than the steady-state change if $K_c k > 1$, which is usually the case, except for delay-dominant processes. If we assume that the allowed input change is $u_{\max}$, then to avoid input saturation we must select τ_c such that (SIMC PI control):

$$|u(t = 0^+)| = |K_c y_s| = |\frac{\tau_1}{\tau_c + \theta} \frac{1}{k} y_s| \leq |u_{\max}| \tag{35}$$

Note that u and y_s are deviation variables. Consider, for example, a first-order process with $\tau_1 = 8$ and $\theta = 1$. With the choice $\tau_c = \theta$, the initial input change is $\tau_1/(\tau_c + \theta) = 4$ times the steady-state input change y_s/k. If such a large dynamic input change is not feasible then one would need to use "smoother" control with a larger value for τ_c in order to satisfy (35).[2]

With PID control, the derivative action will cause even larger input changes for output disturbances and this may be one reason for reducing or even avoiding derivative action. It is also the reason why we to avoid "derivative kick", recommend that the setpoint is not differentiated, see (2).

5 Optimality of SIMC PI rules

How good are the SIMC PI rules, that is, how much room is there for improvements? To study this, we compare the SIMC PI performance, with τ_c as a parameter, to the "Pareto-optimal" PI-controller. Pareto-optimality applies to multiobjective problems, and means that no further improvement can be made in objective 1 (output performance in our case) without sacrificing objective 2 (robustness and input usage in our case).

We choose to quantify robustness and input usage in terms of the sensitivity peak M_s. We also considered other "robustness" measures, for example, the relative delay margin as suggested by Foley *et al.* (2005), but we choose to use M_s. One reason is that we found that the M_s-value correlates well with the input usage as given by its total variation (TV), which agrees with the findings of Foley *et al.* (2005). Such a

[2] It may seem from (35) that "slow" processes, which have a large time constant τ_1, will always require "slow" control (large τ_c) in order to avoid excessive input changes. However, this is usually not the case because such processes often have a corresponding large gain k, such that the value $k' = k/\tau_1$ may be sufficiently large to satisfy (35) even with $\tau_c = \theta$.

correlation is reasonable since a large M_s-value corresponds to an oscillatory system with large input variations.

We choose to quantify performance in terms of the integrated absolute error in response to a setpoint change (IAE_{ys}) and to an input "load" disturbance (IAE_d). The setpoint performance is often referred to as the "servo" behavior and the disturbance (in this case the input "load" disturbance) performance is often referred to as "regulator" behavior. It may be argued that a two-degree of freedom controller ("feedforward action") may be used to improve the response for setpoints, but note that a setpoint change is equivalent to an output disturbance (with $g_d = 1$ in Figure 1) which can only be counteracted by feedback. Thus, both setpoint changes (output disturbances) and input disturbances should be included when evaluating performance, and to get a good balance between the two, we weigh them about equally by defining the following performance cost

$$J(c) = 0.5 \left[\frac{\text{IAE}_{ys}(c)}{\text{IAE}^o_{ys}} + \frac{\text{IAE}_d(c)}{\text{IAE}^o_d} \right] \tag{36}$$

where the reference values, IAE^o_{ys} and IAE^o_d, are for IAE-optimal PI-controllers (with $M_s = 1.59$) for a setpoint change and input disturbance, respectively. We could have used the truly optimal IAE-value as the reference when computing J (without the restriction $M_s = 1.59$), but this would not have changed the results much because the IAE-value is anyway quite close to its minimum at $M_s = 1.59$. Table 3 gives the tunings and reference values obtained using IAE-optimal PI-controllers (with $M_s = 1.59$) for four different processes, and Table 4 gives the tunings, costs J and M_s-values for the SIMC PI-controller (with $\tau_c = \theta$). Importantly, the weighted cost J is independent of the process gain k and the disturbance magnitude, and also of the unit used for time. Note that two different optimal PI-controllers are used to obtain the two reference values, whereas a single controller c is used to find $\text{IAE}_{ys}(c)$ and $\text{IAE}_d(c)$ when evaluating the weighted IAE-cost $J(c)$.

	Setpoint			Input disturbance			Optimal combined (minimize J)					
Process	K_c	τ_I	IAE^o_{ys}	K_c	τ_I	IAE^o_d	K_c	τ_I	IAE_{ys}	IAE_d	J	M_s
e^{-s}	0.20	0.32	1.607	0.20	0.32	1.607	0.20	0.32	1.607	1.607	1	1.59
$\frac{e^{-s}}{s+1}$	0.54	1.10	2.083	0.50	1.0	2.036	0.54	1.10	2.083	2.041	1.00	1.59
$\frac{e^{-s}}{8s+1}$	4.0	8	2.169	3.34	3.7	1.135	3.46	4.0	3.111	1.158	1.23	1.59
$\frac{e^{-s}}{s}$	0.50	∞	2.169	0.40	5.8	15.09	0.41	6.3	4.314	15.4	1.51	1.59

IAE_{ys} is for a unit setpoint change. IAE_d is for a unit input disturbance.

Table 3 Optimal PI-controllers ($M_s = 1.59$) and corresponding IAE-values for four processes.

Figure 8 shows the trade-off between performance (J) and robustness (M_s) for the SIMC PI-controller (blue solid curve) and the Pareto-optimal controller (dashed black curve) for four different processes: pure time delay ($\tau_1/\theta = 0$), small time constant ($\tau_1/\theta = 1$), intermediate time constant ($\tau_1/\theta = 8$), and integrating process ($\tau_1/\theta = \infty$). The curve for the SIMC controller was generated by varying the tuning

	SIMC PI ($\tau_c = \theta$)						Improved SIMC PI ($\tau_c = \theta$)					
Process	K_c	τ_I	IAE_{ys}	IAE_d	J	M_s	K_c	τ_I	IAE_{ys}	IAE_d	J	M_s
e^{-s}	0	0 $^{(*)}$	2.17	2.17	1.35	1.59	0.17	0.33	1.95	1.95	1.21	1.45
$\frac{e^{-s}}{s+1}$	0.5	1	2.17	2.04	1.15	1.59	0.67	1.33	1.99	1.99	1.09	1.69
$\frac{e^{-s}}{8s+1}$	4	8	2.17	2.00	1.38	1.59	4.17	8	2.14	1.92	1.34	1.62
$\frac{e^{-s}}{s}$	0.5	8	3.92	16	1.43	1.70	0.5	8	3.92	16	1.43	1.70

$^{(*)}$ Pure integral controller with $K_I = K_c/\tau_I = 0.5$.

Table 4 SIMC PI-controllers ($\tau_c = \theta$) and corresponding J- and M_s-values for four processes.

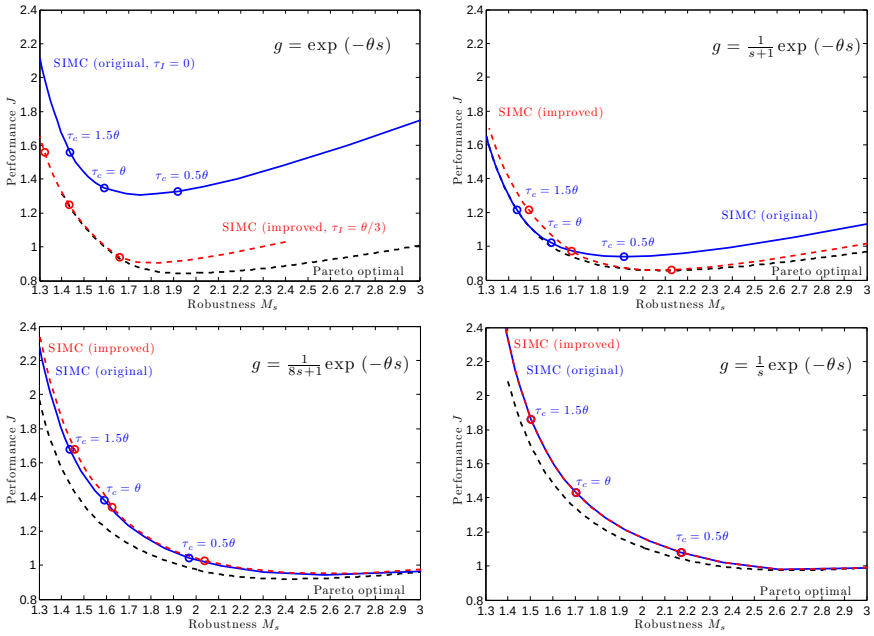


Fig. 8 Check of optimality of SIMC PI tuning rules for four processes.

parameter τ_c from a large to a small value. The controllers corresponding to the choices $\tau_c = 1.5\theta$ (smoother), $\tau_c = \theta$ (recommended) and $\tau_c = 0.5\theta$ (aggressive) are shown by circles. The Pareto-optimal curve was generated by finding for each value of M_s, the optimal PI-controller c, with the smallest IAE-value $J(c)$. Except for the pure time delay process, the differences between the J-values for SIMC (blue solid curve) and optimal (dashed black curve) are small (within 10%), which shows that the SIMC PI-rules are close to optimal.

Note that we have a real trade-off between performance (J) and robustness (M_s) only when there is a negative slope between these variables (in the left region in the figures in Figure 8). We never want to be in the region with a zero or positive slope (to the right in the figures), because here we can improve both performance (J) and robustness (M_s) at the same time with another choice for the tuning parameter (using a larger value for τ_c). Another important observation from Figure 8 is then that the

SIMC-recommendation $\tau_c = \theta$ for "tight" control (as given by middle of the three circles) in all cases is located in the desired trade-off region with a negative slope, well before we reach the minimum. Also, the recommended choice give a fairly constant M_s-value in the region 1.59 to 1.7. From this we conclude that, except for the time delay process, there is little room to improve on the SIMC PI rules, at least when performance and robustness are as defined above (J and M_s).

The IAE-cost J in (36) is based on equal weighting of servo (output disturbance) and regulator (input disturbance) performance. The existence of a trade-off between servo and regulator performance, can be quantified by considering how much larger the (Pareto) optimal cost J_{opt} (dashed black line) is than 1 at the reference robustness, $M_s = 1.59$, see also Table 3. For a pure time delay-process, we have that $J_{opt} = 1$ for $M_s = 1.59$ and there is no trade-off. The reason is that the setpoint and output disturbance responses are the same. On the other hand, for the other extreme of an integrating process, we have a clear trade off since the optimal PI-controller has $J_{opt} = 1.51$ (the SIMC PI-controller with $M_s = 1.59$ is close to this with J about 1.6). The existence of the servo/regulator trade-off for an integrating process, implies that one for a given robustness (M_s-value) can find PI-settings with significantly better regulator (load disturbance) performance or better servo (setpoint) performance, but not both at the same time. To be able to shift the trade-off, one may introduce an extra parameter in the PID rules (Alcantara *et al.* 2010), in addition to τ_c. For the SIMC method, this extra servo/regulator trade-off parameter could be c in the following expression for the integral time,

$$\tau_I = \min(\tau_1, c(\tau_c + \theta)) \tag{37}$$

where $c = 4$ gives the original SIMC-rule. A larger value if c improves the setpoint performance, and a smaller value, e.g. $c = 2$, improves the input disturbance performance (Haugen 2010). However, introducing an extra parameter adds complexity and the potential benefit does not seem sufficiently large. Nevertheless, one may consider choosing another (lower) fixed value for c. There are two reasons why we recommend keeping the SIMC-value of $c = 4$. First, it is close to the Pareto-optimal PI controller (as seen from Figure 8), so we cannot get a significant improvement with our performance objective J. Second, with a smaller value for c, say $c = 2.5$, the recommended choice $\tau_c = \theta$ becomes less robust (with a higher M_s), so one would need to recommend a different value for τ_c for an integrating process, say $\tau_c = 1.5\theta$, which would add complexity. In summary, we find that the value $c = 4$ in the original SIMC rule provides a well-balanced servo/regulator trade-off.

6 Improved SIMC tuning rules

For a pure time delay process, we see from Figure 8 that the IAE-value (J) for the SIMC controller is about 40% higher than the minimum with the same robustness (M_s). This is further illustrated by the closed-loop simulations in Figure 9 where we

see that the SIMC PI-controller (denoted SIMC-original in the figure) gives a nice and smooth response. However, the response is somewhat sluggish initially, because it is actually a pure I-controller (with $K_c = 0, \tau_I = 0$ and $K_I = K_c/\tau_I = 0.5$). On the other hand, the IAE-optimal PI-controller (with minimum J for $M_s = 1.59$) has K_c about 0.2 and τ_I about 0.32 (and $K_I = 0.62$). In fact, the optimal PI-controller for a pure time delay process (dashed black line in Figure 8), has an almost fixed integral time of approximately $\theta/3$ for all values of M_s between 1.4 and 1.7.

Based on this fact, we propose a simple change to the SIMC-rules, namely to replace τ_1 by $\tau_1 + \theta/3$ in the rules (PI control), which markedly improved the responses for a pure time delay process. It is important that the change is simple because "simplicity" was one of the main objectives when originally deriving the SIMC rules.

A similar change, but with $\theta/2$ rather than $\theta/3$, was originally proposed by Rivera *et al.* (1986) for their "improved PI" tuning rule, and the effectiveness of this modification is also clear from the paper of Foley *et al.* (2005). However, as seen in Figure 9, the response with this IMC PI controller also settles rather slowly towards the setpoint, indicating that the integral time $\theta/2$ is too large. The proposed value $\theta/3$ gives a faster settling and is also closer to the original SIMC-rule (which is zero for a time delay process). The conclusion is that we recommend to replace τ_1 by $\tau_1 + \theta/3$ in the SIMC rules to get the improved SIMC rules:

Improved SIMC PI-rule for first-order with delay process.

$$K_c = \frac{1}{k}\frac{\tau_1 + \frac{\theta}{3}}{\tau_c + \theta} \tag{38}$$

$$\tau_I = \min\{\tau_1 + \frac{\theta}{3}, 4(\tau_c + \theta)\} \tag{39}$$

The improvement of this rule for a pure time delay processes is clear from the red curves in Figures 9 and 8 (upper left); for small M_s-values the improved SIMC-controller is almost identical to the Pareto-optimal, which confirms that $\tau_I = \theta/3$ is close to optimal for a pure time delay process. For the process with a small time constant ($\tau_1 = \theta$), the improved SIMC rule (red curve in upper right plot in Figure 8) is slightly better than the "original" SIMC rule (blue curve) for higher M_s-values (where we get better performance) but slightly worse for lower M_s-values. For the two processes with a large time constant ($\tau_1 = 8\theta$ and $\tau_1 = \infty$) there are, as expected, almost no difference between the original and improved SIMC rules.

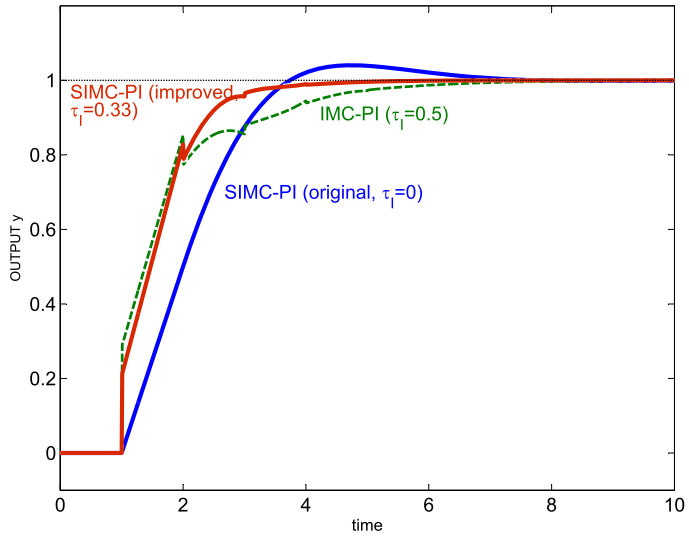


Fig. 9 Closed-loop setpoint responses for pure time delay process ($\theta = 1, k = 1, \tau_1 = 0$) with PI-control. All three controllers have the same robustness ($M_s = 1.59$).
For a pure time delay process, the setpoint and disturbance responses are identical, and the input and output are identical.
IMC PI: $K_c = 0.29$ and $\tau_I = 0.5$ ($K_I = K_c/\tau_I = 0.58$).
SIMC PI original ($\tau_c = \theta$): $K_c = 0$ and $\tau_I = 0$ ($K_I = 0.5$).
SIMC PI improved ($\tau_c = 0.61\theta$): $K_c = 0.207$ and $\tau_I = 0.333$ ($K_I = 0.62$).

7 Discussion

7.1 Measurement noise

Measurement noise has not been considered in this chapter, but it is an important consideration in many cases, especially if the proportional gain K_c is large, or, for cases with derivative action, if the derivative gain $K_c\tau_D$ is large. However, since the magnitude of the measurement noise varies a lot in applications, it is difficult to give general rules about when measurement noise may be a problem. In general, robust designs (with small M_s) are insensitive to measurement noise. Therefore, the SIMC rules with the recommended choice $\tau_c = \theta$, are less sensitive to measurement noise than most other published settings method, including the Ziegler-Nichols-settings. If actual implementation shows that the sensitivity to measurement noise is too large, then the following modifications may be attempted:

1. Filter the measurement signal, for example, by sending it through a first-order filter $1/(\tau_F s + 1)$; see also (2). With the proposed SIMC-settings one can typically increase the filter time constant τ_F up to almost 0.5θ, without a large affect on performance and robustness.

2. If derivative action is used, one may try to remove it, and obtain a first-order model before deriving the SIMC PI-settings.
3. If derivative action has been removed and filtering the measurement signal is not sufficient, then the controller needs to be detuned by selecting a larger value for τ_c.

7.2 Retuning for integrating processes

Integrating processes,

$$g(s) = k' \frac{e^{-\theta s}}{s}$$

are common in industry, but control performance is often poor because of incorrect controller settings. When encountering oscillations, the intuition of the operators is to reduce the controller gain. If the oscillations are relatively slow, then this is the exactly opposite of what one should do for an integrating process. The product of the controller gain K_c and the integral time τ_I must be *larger* than $4/k'$ to avoid slow oscillations (Skogestad 2003). One solution is to simply use proportional control (with $\tau_I = \infty$), but this is often not desirable. Here we show how to easily retune the controller to just avoid the oscillations without actually having to derive a model. This approach has been applied with success to industrial examples.

Consider a PI controller with (initial) settings K_{c0} and τ_{I0} which results in "slow" oscillations with period P_0 (larger than $3 \cdot \tau_{I0}$, approximately). Then we likely have a close-to integrating process for which the product of the controller gain and integral time ($K_{c0}\tau_{I0}$) is too low. To avoid oscillations with the new settings K_c and τ_I we must require (Skogestad 2003):

$$\frac{K_c \tau_I}{K_{c0}\tau_{I0}} \geq \frac{1}{\pi^2} \cdot \left(\frac{P_0}{\tau_{i0}}\right)^2 \tag{40}$$

Here $1/\pi^2 \approx 0.10$, so we have the **rule**:

- To avoid "slow" oscillations the product of the controller gain and integral time should be increased by a factor $f \approx 0.1(P_0/\tau_{I0})^2$.

7.3 Controllability

The effective delay θ is easily obtained using the proposed half rule. Since the effective delay is the main limiting factor in terms of control performance, its value gives invaluable insight about the inherent controllability of the process.

From the settings in (27)-(29), a PI-controller results from a first-order model, and a PID-controller from a second-order model. With the effective delay computed

using the half rule in (12)-(13), it then follows that PI-control performance is limited by (half of) the magnitude of the second-largest time constant τ_2, whereas PID-control performance is limited by (half of) the magnitude of the third-largest time constant, τ_3.

8 Conclusions and Future Perspectives

This chapter has summarized the SIMC two-step procedure for deriving PID settings for typical process control applications.

Step 1. The real process is approximated by a first-order with delay model (for PI control) or a second-order model (for PID control). To obtain the model, the simplest approach is probably to use an open-loop step experiment (Figure 3), but if this is difficult for some reasons, then one may alternatively use a closed-loop setpoint response with P-controller (Figure 4). If the starting point is a detailed model, then the half rule may be used to obtain the effective delay θ, see (12)-(13).

Step 2. For a first-order model (with parameters k, τ_1 and θ) the following SIMC PI-settings are suggested (original SIMC rule):

$$K_c = \frac{1}{k}\frac{\tau_1}{\tau_c + \theta}; \quad \tau_I = \min\{\tau_1, 4(\tau_c + \theta)\}$$

where the closed-loop response time τ_c is the tuning parameter. For a dominant second-order process (for which $\tau_2 > \theta$, approximately), one needs to add derivative action with

$$\text{Series} - \text{form PID}: \quad \tau_D = \tau_2$$

To improve the performance for delay-dominant processes, one may replace τ_1 by $\tau_1 + \frac{\theta}{3}$ and use the "improved" SIMC PI-rules in (38)-(39). A more careful analysis needs to be done to check if a similar improvement can be used with a PID controller.

Note that although the same formulas are used to obtain K_c and τ_I for both PI- and PID-control, the actual values will differ since the effective delay θ is smaller for a second-order model. The tuning parameter τ_c should be chosen to get the desired trade-off between fast response (small IAE) on the one side, and smooth input usage and robustness (small M_s) on the other side. The recommended choice $\tau_c = \theta$ gives robust (M_s about 1.6 to 1.7) and somewhat conservative settings when compared with most other tuning rules, and if it is desirable to get faster control one may consider reducing τ_c to about $\theta/2$ (see Figure 8). More commonly, one may want to have "smoother" control with $\tau_c > \theta$ and a smaller controller gain K_c. However,

the controller gain must be larger than the value given in (32) to achieve a minimum level of disturbance rejection.

Comparing the performance of the SIMC-rules with the optimal for a given robustness (M_s value) shows that the SIMC-rules are close to the Pareto-optimal settings (Figure 8). This means that the room for improving the SIMC PI-rules is limited, at least for the first-order plus delay processes considered in this chapter, and with a good trade-off between rejecting input and output (setpoint) disturbances.

However, it should be noticed that the SIMC rules apply to processes that can be reasonably well approximated by first or second order plus delay models. This applies to most process control applications, including some unstable plants, but it obviously does not apply in general, for example, for some of the unstable or oscillating processes found in mechanical systems. For such processes, it would be interesting to study the validity and extension of the SIMC rules or similar analytic model-based PID tuning rules. It is also interesting to establish for which processes the PID controller is a suitable controller and for which processes it is not.

References

Alcantara, S., C. Pedret and R. Vilanova (2010). On the model matching approach to PID design: Analytical perspective for robust servo/regulator tradeoff tuning. *J. Process Control* **20**, 596–608.

Astrom, K.J. and T. Hagglund (1995). *PID controllers: Theory, design and tuning*. 2nd ed.. ISA-Instrument Society of America.

Chien, I.L. and P.S. Fruehauf (1990). Consider IMC tuning to improve controller performance. *Chemical Engineering Progress* pp. 33–41.

Foley, M.W., N.R Ramharack and B.R Copeland (2005). Comparison of PI controller tuning methods. *Ind. Eng. Chem. Res.* **44**(17), 6741–6750.

Grimholt, C. (2010). Verification and improvements of the SIMC method for PI control. Technical report. 5th year project. Department of Chemical Engineering. Norwegian University of Science and Technology, Trondheim. http://www.nt.ntnu.no/users/skoge/diplom/prosjekt10/grimholt.

Haugen, F. (2010). Comparing PI tuning methods in a real benchmark temperature control system. *Modeling, Identifiation and Control* **31**, 79–91.

Horn, I.G., J.R. Arulandu, J. Gombas, J.G. VanAntwerp and R.D. Braatz (1996). Improved filter design in Internal Model Control. *Ind. Eng. Chem. Res.* **35**(10), 3437–3441.

Rivera, D.E., M. Morari and S. Skogestad (1986). Internal model control. 4. PID controller design. *Ind. Eng. Chem. Res.* **25**(1), 252–265.

Seborg, D.E., T.F. Edgar and D.A. Mellichamp (1989). *Process Dynamics and Control*. John Wiley & Sons.

Shamsuzzoha, M. and S. Skogestad (2010). The setpoint overshoot method: A simple and fast method for closed-loop PID tuning. *J. Process Control* **20**, 1220–1234.

Skogestad, S. (2003). Simple analytic rules for model reduction and PID controller tuning. *J. Process Control* **13**, 291–309.

Skogestad, S. (2006). Tuning for smooth PID control with acceptable disturbance rejection. *Ind. Eng. Chem. Res.* **45**, 7817–7822.

Smith, C.A. and A.B. Corripio (1985). *Principles and Practice of Automatic Process Control*. John Wiley & Sons.

Smith, O.J. (1957). Closer control of loops with dead time. *Chem. Eng. Prog.* **53**, 217.

Tyreus, B.D. and W.L. Luyben (1992). Tuning PI controllers for integrator/dead time processes. *Ind. Eng. Chem. Res.* pp. 2628–2631.

Yuwana, M. and D.E. Seborg (1982). A new method for on-line controller tuning. *AIChE Journal* **28**, 434–440.

Ziegler, J.G. and N.B. Nichols (1942). Optimum settings for automatic controllers. *Trans. of the A.S.M.E.* **64**, 759–768.

Appendix

Estimation of parameters τ_1 and θ from closed-loop step response.

Shamsuzzoha and Skogestad (2010) discuss at the end of their paper a two-step closed-loop procedure, where the first step is to use closed-loop data and some expressions to obtain the parameters k, τ_1 and θ. We use this approach but have modified the expressions. Our expressson for k in (7) is given by their equation (35) by noting that $B = |(1-b)/b|$ where $b = \Delta y_\infty / \Delta y_s$. However, our expressions for θ and τ_1 in (8)-(9) differ somewhat from their equations (36) and (37). The reason is that their equations (36) and (37) are not consistent in terms of the time delay estimate, because the expression for τ_1 in (36) is based on $\theta = 0.43 t_p$, whereas (37) uses $\theta = 0.305 t_p$. To correct for this, we first note from (19) in their paper (noting that $\tau_1 = \tau_I$ for the delay-dominant case), that τ_1 and θ are related by

$$\tau_1 = r\theta$$

where $r = 2A/B$, which is our expression in (9). Here, Shamsuzzoha and Skogestad (2010) recommend to use $\theta = 0.44 t_p$ for $\tau_1 < 8\theta$ and $\theta = 0.305 t_p$ for $\tau_1 > 8\theta$. However, to get better accuracy and a smooth transition, we fitted simulation data for θ/t_p as a function of τ_1/θ for a wide range of processes with an overshoot of 0.3, and obtained the correlation (Grimholt 2010)

$$\theta = t_p \cdot (0.309 + 0.209 e^{-0.61(\tau_1/\theta)})$$

as given in (8). Note here that $(0.309 + 0.209 e^{-0.61(\tau_1/\theta)})$ is 0.518 for $r = \tau_1/\theta = 0$, and 0.309 for $r = \infty$.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Wannabe-MPC for Large Systems Based on Multiple Iterative PI Controllers

Pasi Airikka*, Mats Friman*

**Metso Corporation, Tampere, P.O Box 237, FIN-33101*
FINLAND (Tel: +358-40-5872730; e-mail: firstname.lastname@metso.com).

Abstract: PID controllers are indisputably the most common controller type encountered in industrial process control. Despite their tremendous usage in single-variable process control applications, they have more than that. Multiple PID controllers can be connected to work as a multivariable controller or, as it is shown in this paper, they can be totally re-organised to work together for establishing a controller structure that mimics the model predictive controller. Having the characteristics of the MPC, the proposed structure does not require a solver for optimization problems but only a bunch of PID controllers that exist in any automation platform.

Keywords: PID control, model predictive, iterative, process.

1. INTRODUCTION

For years, distributed control systems or programmable logic controllers have contained a functionality of a PID controller as a work horse for process control. The tremendous distribution the PID controllers have made them an ad hoc basic level solution for industrial control applications (Åström, 1995). And when proven useful, an optimization control layer of model predictive controllers may have been laid upon the basic layer of the PID controllers.

The model predictive controllers such as Model Predictive Heuristic Control (Richalet et. al, 1978), Dynamic Matrix Control (Cutler & Ramaker, 1980), Generalized Predictive Control (Clarke et. al, 1987) and Non-linear Model Predictive Control (Bequette, 2007) can be categorized as members of a Model Predictive Control (MPC). The MPC controllers with different variants are nowadays diverse but, at the same time, well-accepted in industrial applications.

The MPC controllers require special attention on their design and tuning. There have been studies on different tuning strategies for MPC controllers (e.g Rani & Unbehauen, 2007). Yet, their design and implementation still takes a lot of effort with respect to time and money. Therefore, they are not put into practise in industrial applications unless there is evidence on their contribution and benefits with pay offs.

This paper proposes a novel control design for creating a multivariable wannabe-MPC controller based on usage of conventional PI controllers alone. Instead of applying optimization solvers and numerous tuning parameters, a structure consisting of Multiple Iterative PI controllers (MIPI) is introduced. The proposed rather simple MIPI controller can be applied for controlling large, even complex processes without requiring extensive computational capacity for solving optimization problems. An application solution might be found e.g in steam network optimization for co-generation power plants where MPC-based control strategy has been traditionally used for optimizing steam balance and delivery with respect to changing loads due to power generating steam turbines and steam consuming processes such as paper or pulp mill, or district heating process (Airikka & Mäkilä, 2009).

2. MULTIPLE PI CONTROLLERS

The idea of the Multiple Iterative PI Controller (MIPI) is to organise several PI controllers for one control loop having a manipulated variable (MV) and a control variable (CV) as given in figure 1 (Friman, 2008).

Following the idea of a general model predictive controller, there is a target CV trajectory for the MIPI controller and its predicted CV trajectory following the target the best it can. However, instead of applying any optimization solver on a pre-defined cost function with prediction and control horizons, there are simply a bunch of PI controllers working on the optimization.

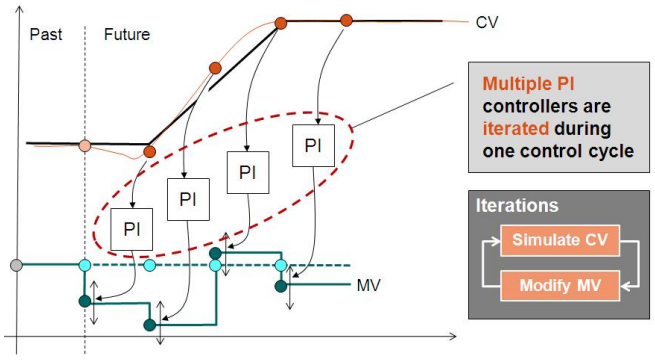


Figure 1. Multiple PI controllers: general principle of organizing several PI controllers for one control loop.

The predicted CV trajectory is controlled towards the target CV trajectory with multiple PI controllers working on an iterative manner. There are selected time instants (which can be e.g consequtive time instants) with their predicted CV trajectory values. Each selected CV trajectory instant is assigned with a PI controller. The PI controller receives the target and the predicted trajectory values at the pre-selected time instant and these values serve as inputs for calculating the control error. As an output, the PI controller returns the needed MV value.

The calculation of each PI controller is repeated until a pre-defined stopping criterion has been reached. The criterion for ending iterations can be e.g an absolute value of the control error or a number of iterations. The authors suggest combining both above mentioned criteria in order to limit the computation time used for each controller.

Having iterations, the MV trajectory is available. But as with a MPC, only the first MV trajectory value representing the current controller output is put into action and the rest are simply ignored. Then, at the next control cycle, the sampe described procedure is to be repeated again.

As described, the MV trajectory with its values at given time instants (figure 1) are evaluated using the MIPI controller structure. When using a MPC, those values would be available only through optimization. Hence, the optimization problem has been converted here to a control task for multiple PI controllers being iterated.

The calculation procedure of the MIPI controller for unit control loop with one MV and CV is given in figure 2.

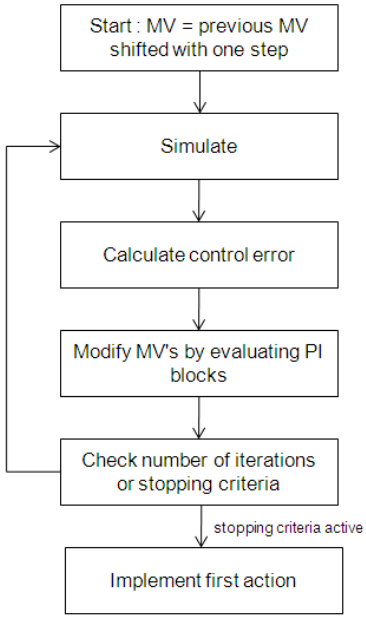


Figure 2. Calculation procedure of Multiple Iterative PI controller (MIPI) in terms of consequtive function blocks.

3. PRACTICALITIES

There are several practical issues to be considered when designing a MIPI controller and implementing it. These practicalities are treated in this chapter.

3.1 Extention to multivariable control

The proposed structure and procedure of the MIPI controller (fig. 1-2) applies to a single variable control loop having one MV and one CV only. The corresponding control loop layout is illustrated in figure 3. The layout contains only one MIPI controller, which, however, includes several iteratives PI controllers as decribed in the previous chapter.

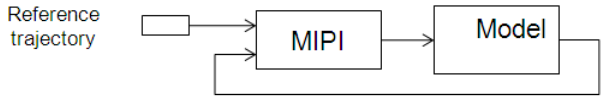


Figure 3. Single-Input-Single-Output control loop (1 MV + 1 CV) with MIPI controller.

However, the MIPI controller is naturally expandable to multivariable control with unlimited dimensions in MV and CV. The control loop layout for a multivariable (2 x 2) design is shown in figure 4 where there are two MIPI controllers, one for each MV and CV pair. The illustrated multivariable structure can be expanded to cover all the dimensions necessary for a succesful multivariable MPC-like control

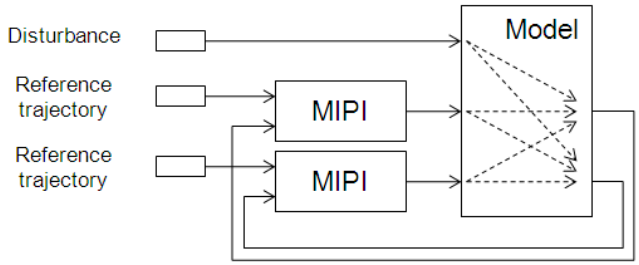


Figure 4. Multivariable (2 MV x 2 CV) control loop with MIPI controllers.

3.2 Non-measured disturbances and model mismatches

A control loop typically has disturbances that are to be attenuated in order to keep the CV at its reference value. These disturbances lay the foundation for the existence of the feedback control loop. The disturbances can be measured, unmeasuremable, modelled or unmodelled.

Ideally, in case of having no disturbances and perfert modeling, the resulted control performance is also ideally perfect. However, facing uncertainties such as unmodelled disturbances and model mismatch these need to be considered in the design phase of any controller.

It is proposed by the authors to introduce a target trajectory that would be generated from the reference trajectory and the current control error at a time. The reference trajectory is the ultimate target but instead of taking that to the MIPI controller, its filtered correspondent target trajectory would be given as a setpoint for the MIPI controller. Figure 5 illustrates the principle.

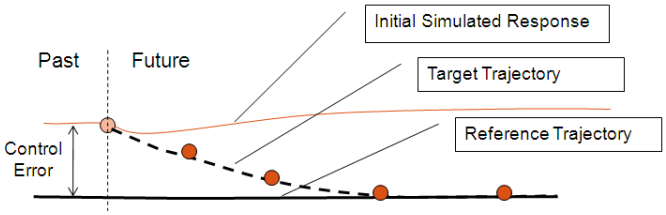


Figure 5. Generation of target trajectory which given as a setpoint to the MIPI controller instead of the ultimate reference trajectory.

3.3 Coupling different instants of MV and CV

As shown in figure 1, each iterative PI controller is connected between selected instants of MV and CV trajectory. The selected time instants should be selected in such a manner that a change in the MV value of the chosen time instant will have an impact on the coupled CV value (figure 2). Basically, this leads to a requirement of coupling MV and CV so that there is a time window of the process dead time in-between, at the minimum.

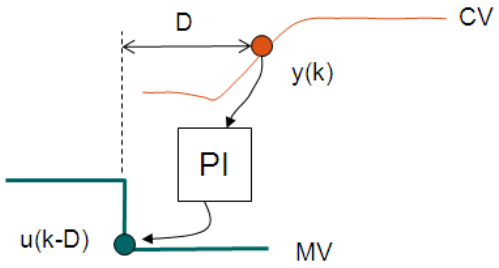


Figure 6. Iterative PI controller connection in MIPI controller structure.

Consequently, coupling requires determination of process dead times between different MVs and CVs. Typically, this is not an issue as there should be a process model with dead times available anyhow for predicting CV trajectory based on the MVs and the modelled disturbances. However, if for any reason, the dead times are missing, they should be estimated using delay-estimation techniques.

3.4 Input constraints

A wanna-be MPC must be able to treat with input constraints as any MPC would do. Therefore, handling of input constraints could be considered using the principle illustrated in figure 7. Each iterative PI controller is given the input constraints for limiting its own contribution for the CV trajectory. When violating the constraint, the iterative PI controller output is shared with the neighbour (previous or next) iterative PI controller by adding the overflowing part of the output to the output of another controller.

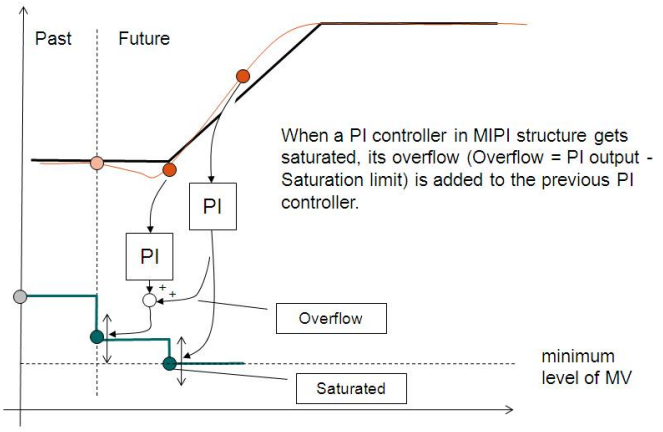


Figure 7. Handling input contraints by setting constraints to iterative PI controllers and sharing overflows with previous iterative PI controllers.

In addition to the constraints of input (MV) amplitudes, constraints for input rate changes can be treated as well. This is achieved by setting rate limiters for each iterative PI controller output.

3.5 Output constraints

Similarly to the requirement of being capable of handling input constraints, the output constraints must be considered as well. The way they are included in the MIPI controller design is rather simple (figure 8). Assuming that the MV trajectory constraints are given in terms minimum and maximum limits y_{min} and y_{max}, they can be considered in each iterative PI controller separately. If the CV trajectory at any time instant remains within the limits, the particular iterative PI controller is given a control error of zero. Otherwise, the control error between the target and the predicted CV trajectory is taken to the iterative PI controller as such.

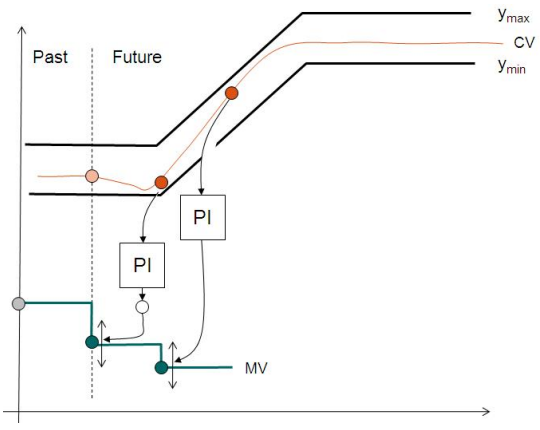


Figure 8. Handling input contraints by setting constraints to iterative PI controllers and sharing overflows with previous iterative PI controllers.

4. DESIGN OF MIPI CONTROLLER

The MIPI controller requires a process model as any other model predictive controller would do. The model can be of multivariable type, linear or non-linear, transfer-function or state-space spaced. The process model is inevitably needed for predicting the CV trajectories but also for coupling the time instants of CV and MV trajectories.

4.1 Design parameters and constraints

Design of the MIPI controller includes selection of control horizon (number of control samples) and sampling time as tuning parameters. For these, it is recommended to follow the same guidelines as in the desing of MPC controllers. For iterative calculations, the maximum number of iterations and a stopping criterion are to be set. These are, however, internal parameters which can be fixed regardless of the solution. Input and output constraints are, in turn, to be considered in each project separately to follow the given control performance specifications.

4.2 Iterative PI controller tuning

In addition, the iterative PI controllers need their tuning parameters (proportional gain, integral time) at the minimum. It is suggested that all the iterative PI controllers for one and the same CV are given the same tuning parameters. Preferrably, they could be tuned using an autotuning procedure. It is also noticeable that the iterative PI controllers can be employed without any dead time compensation elements as they controlling CV and MV trajectory instants which have their interrelated dead time considered by the coupling strategy.

4.3 Filter design for target trajectory

Propably the most crucial design entity is filter design for generating the CV target trajectory from the reference trajectory. Introducing slow filtering or ramped outputs of the reference trajectory, the MIPI controller performance can be affected to a great extent. This paper does not give any guidelines for practical choices but only adresses the underlying importance of the filter design.

4.4 No weighting matrices

Differentiating from the MPC controllers, no weighting matrices with control error and control action penalties are needed. Lack of the numerous tuning parameters drastically reduces the tuning effort. Yet, some of the tuning work is simply converted to filter design for the CV target trajectories.

4.5 Stability criterion

Stability criteria for the proposed MIPI control has not been established yet and, therefore, meanwhile, it remains an open issue to be investigated in further work.

5. SIMULATION EXAMPLE

The proposed MIPI control method is tested using a multivariable transfer function model of the distillation process separating methanol and water as given in (Wood & Berry, 1973):

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \dfrac{12.8e^{-s}}{16.7s+1} & \dfrac{-18.9e^{-3s}}{21s+1} \\ \dfrac{6.6e^{-7s}}{10.9s+1} & \dfrac{-19.4e^{-3s}}{14.4s+1} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} \dfrac{3.8e^{-8s}}{14.9s+1} \\ \dfrac{4.9e^{-3s}}{13.2s+1} \end{bmatrix} d \qquad (1)$$

where distillate composition y_1 (%) and bottom composition y_2 (%) are CVs being controlled by MVs reflux flow u_1 (lb/min) and steam flow u_2 (lb/min). The feed flow rate d acts as an unmeasured disturbance.

Figure 9 illustrates the CVs and MVs when the process (1) is being controlled by the MIPI controller as described in this paper and the process is assumed to be perfectly modelled. The upper figure shows both CVs with respect to time whereas the lower figure plots both MVs, accordingly.

MIPI controller has a sampling time h = 1 min and a prediction horizon of 10 minutes. The number of iterative PI controllers used for one CV/MV pair is 10 and the number of iterations in computations is 11. The iterative PI controllers have parameters k_p = 0.0134, t_i = 0.1 for CV1/MV1 and k_p = -0.0077, t_i = 0.1 for CV2/MV2. The time constants for filtering the reference trajectory are set to 40 mins (CV1/MV1) and 60 mins (CV2/MV2).

The simulation has included four upsets. First, at time t = 20, a known disturbance d has entered the system. Second, at time t = 120, the same disturbance d has entered but now being unmeasured. Third, setpoint changes on both CVs have made at t = 220 and, finally, at t = 320, the setpoint changes have taken place unexpectedly.

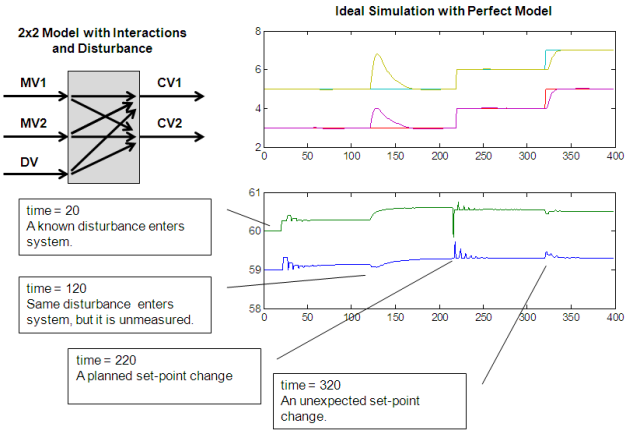


Figure 9. Simulated CV and MV responses of the distillation process (Wood & Berry, 1973) controlled by the MIPI controller. The process model is ideal with no model mismatch.

Figure 10 shows the same CV and MV responses of the MIPI controlled distillation process with the same process upsets but this time with model mismatch. The modelling uncertanties are simulated by using time constants half of the given model (1) and static gains and dead times 1.5 larger than in the given model (1).

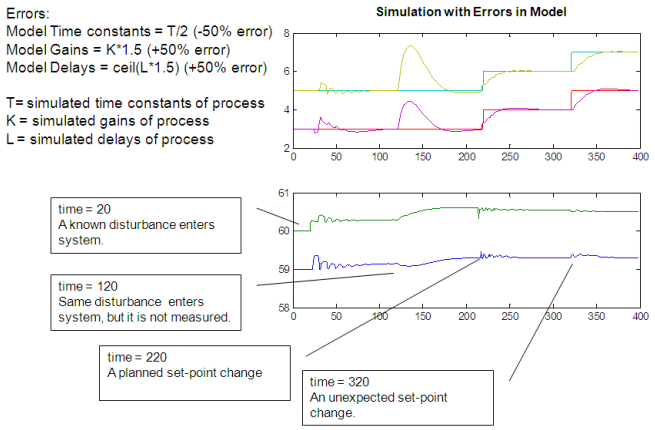


Figure 10. Simulated CV and MV responses of the distillation process (Wood & Berry, 1973) controlled by the MIPI controller. The process model is not perferct having mismatches in terms of dynamics and static gains.

6. CONLUSION

A novel method for a multivariable wannabe-MPC controller was presented in this paper. The proposed method relies on conventional PI controllers which are linked to each other for working iteratively to find the optimal MV trajectory for the given CV target trajectory. The proposed method does not require any optimization algorithm or solver, neither tuning matrices with weighted penalties on control error and control output.

Especially, finding an optimal solution for a non-linear control problem can be far from trivial. In the proposed MIPI control, there is not such an issue. Instead, the optimization task is converted to multiple tasks for the iterative PI controllers to remove the error between the target and predicted CV trajectory.

The proposed MIPI control is especially applicable to large, complex systems where the computational burden increases exponentially with increasing dimensions of MVs and CVs. The proposed MIPI control is also capable of dealing with input and output constraints.

The soft spot of the MIPI control might be in its stability and robustness which have not been studied with sufficient accuracy at the time of writing this paper. Thefefore, there remains future work for establishing stability criteria and robustness conditions with respect to modelling uncertainties, different model types and process disturbances.

REFERENCES

Airikka, P., J. Mäkilä (2009), Energy Management System for CHP plants, *IFAC conference on Control Methodologies and Technology for Energy Efficiency (CMTEE 2009)*, March 29-31, Vilamoura, Portugal.

Bequette B.W. (2007). Non-Linear Model Predictive Control: A Personal Retrospective, *The Canadian Journal of Chemical Engineering*, Vol. 85, August.

Clarke, D. W., C. Mohtadi and P.S. Tuffs (1987). Generalized predictive control: Part I. the basic algorithm, Part II. Extensions and interpretations. *Automatica*, 23, pp. 137-160.

Cutler, C. R., B.L. Ramaker (1980). Dynamic matrix control - A computer control algorithm. *Joint American Control Conference*, San Francisco.

Friman, M. (2008). Internal discussions and communication within Metso Corporation.

Rani K.Y., H. Unbehauen (1997). Study of Predictive Controller Tuning Methods, *Automatica*, Vol. 33, No. 12, pp. 2243-2248.

Richalet, J.A., A. Rault, J.L. Testud, J. Papon (1978). Model predictive heuristic control: Application to industrial processes., *Automatica*, 14, pp. 413-428.

Wood, R.K., M.W Berry (1973). Terminal composition control of a binary distillation column, *Chemical Engineering Science*, 28, pp. 1707-1717.

Åström K.J, T. Hägglund (1995). *PID Controllers: Theory, Design and Tuning*, 2nd ed., Instrument Society of America, USA.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Regulatory layer selection through partial control

Ramprasad Yelchuru Sigurd Skogestad*

Department of Chemical Engineering,
Norwegian University of Science and Technology,
NTNU, Trondheim, Norway.
{e-mail : {skoge}@chemeng.ntnu.no}

Abstract:

Controlled variables selection based on economic objectives using self optimizing concepts are developed (Skogestad, 2000). In this paper, we extend the ideas of self optimizing control to find optimal controlled variables in the regulatory layer. Regulatory layer is often designed to facilitate stable operation, to regulate and to keep the operation in the linear operating range. We can quantify these objectives with the use of state drift criterion from the nominal operating point. Self optimizing control (Skogestad, 2000} is to arrive at optimal controlled variables based on economics in the secondary layer, where as the focus here is to identify the controlled variables in the regulatory layer based on state drift in the presence of disturbances. Using partial control analysis (Shinnar, 2000) and MIQP methods (Yelchuru et al., 2010), we propose quantitative methods to find optimal controlled variables in the partially controlled system.

Regulatory layer selection with n_u steady state degrees of freedom using partial control analysis will require us to solve, $\sum_{i=1}^{n_{u_s}} {}^{n_{u_s}} C_i = 2^{n_{u_s}} - 1$, partially controlled systems with their optimal controlled variables to find the optimal regulatory layer with 1, 2, … loops closed. Even though the number of partial controlled systems increases with n_u, the problem is tractable as the regulatory layer selection is an offline method. These methods can be used to obtain minimum regulatory layer with an acceptable state drift in the presence of disturbances. The developed framework is evaluated on a distillation column case study with 41 stages.

Shinnar, R., (1981). Chemical reactor modelling for purposes of controller design, Chemical Engineering Communications, 9, 73--99

Skogestad, S. (2000). Plantwide control: The search for the self-optimizing control structure, Journal of Process Control, 10, 487--507.

Yelchuru, R., Skogestad, S., Manum, H. (2010). MIQP formulation for controlled variable selection in self optimizing control, DYCOPS, July 7-9, Brussels, 61--66.

Fault detection and diagnosis system for the drying section of a board machine

Zakharov A., Jämsä-Jounela S.-L.

Aalto University, School of Chemical Technology, Process Control and Automation Research Group, P.O.Box 16100, FI-00076 Aalto, Finland (e-mail: alexey.zakharov@aalto.fi)

Extended Abstract

Identifying a reliable and accurate process model from the data is a key step in designing and implementation of Fault Detection and Diagnosis (FDD) systems. Improved model accuracy allows the decrease in detection thresholds while the false alarms probability is kept constant. This allows the FDD method to detect faults quickly and it enables recognition of small magnitude faults. Furthermore, the most FDD methods are based on linear models and are, therefore, unsuitable for highly nonlinear processes. Simultaneously, the non-linear methods suffer from the overfitting effect which degrades the accuracy of the models. In addition, the parameters of the non-linear models often do not have any physical interpretation (black box models). To conclude, there is a need to develop non-linear fault detection and diagnosis methods providing maximum accuracy and transparency of the process models, which are involved to the method implementation.

The proposed method identifies a number of static nonlinear parity equations from the process data and then utilizes the residuals of these equations for fault detection and diagnosis tasks. The changes in the residuals are detected using the cumulative sum (CUSUM) method. Diagnosis is performed utilizing the structured residuals approach. To this end, an incidence matrix describing the faults vs. the residuals is developed. The method incorporates some common properties of the systems in order to avoid overfitting and to achieve good accuracy of the model. In contrast to other fault diagnosis methods, the proposed approach requires only the process knowledge which is typically available.

The applicability of the method to complex flow networks controlled by valves is tested using the drying section of the industrial board machine. The drying section is used to evaporate water from the board by heating. Since only insignificant amount of moisture is removed after the drying section the drying is typically controlled to achieve the specified moisture of the final product. In the multi-cylinder drying, which is the most common drying method, the process contains a number of consequent steam drying cylinders combined into several drying groups. In fact, a complex steam-water network with a number of recirculation streams is needed to support the desired conditions in the cylinders and to provide the acceptable energy efficiency of the drying process.

The FDD in the drying section must be aimed at the most frequent faults and problems which can cause the most valuable losses of productivity and deviations in quality of the final product. The key problems to focus on are the leakages and blockages of valves and pipes in the steam-water network of the drying section. Small magnitude faults may cause energy losses but a large fault could even cause inability of the control system to maintain the required quality and may cause a plant shutdown.

A number of nonlinear parity equations based on mass balances have been identified for the steam-water network of the drying section. The successful validation of the developed equations has proven the ability of the method to deal with fault detection and diagnosis of the non-liner processes.

Keywords: Fault detection and diagnosis, non-linear model identification, structured residuals, industrial application, paper making

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

MPC techniques in fault tolerant control design

Florin Stoican * **Sorin Olaru** ** **Morten Hovd** *

* *Norwegian University of Science and Technology (NTNU) - Engineering Cybernetics Department, Trondheim, Norway*
** *SUPELEC Systems Sciences (E3S) - Automatic Control Department, Gif sur Yvette, France*

Abstract: In the present paper we provide a robust approach for fault tolerant control (FTC) schemes using the methodology detailed in Seron et al. [2008], Olaru et al. [2010]. We guarantee the detection and isolation of a fault through a set-separation condition (FDI mechanism) and use this condition further in the reconfiguration control (RC) mechanism in order to stabilize the closed-loop system and respect performance criteria.

Keywords: Fault tolerant control; MPC; Invariant sets.

1. INTRODUCTION

Nowadays the use of redundant sensors in applications is becoming increasingly more common. In modern control applications there are strict requirements on the stability and performance criteria. There are safety-critical systems in which this behavior is not merely inconvenient but can become catastrophic (well known examples of malfunctioning in aircraft incidents are discussed in Maciejowski and Jones [2003]). As a consequence, a great deal of effort has been put into developing closed-loop systems which can tolerate faults, while maintaining desirable performance and stability properties [Zhang and Jiang, 2008]. Any fault tolerant control (FTC) scheme relies on two basic mechanisms: the fault detection and isolation (FDI) block and the control reconfiguration (RC) block. The solutions employed usually implement *active* FTC schemes which react to a detected fault and reconfigure the control actions so that stability and performance can be satisfied. Arguably, the most important aspect of a FTC scheme is the interaction between the FDI and RC mechanisms. In this sense, it is desirable to adapt the control to the requirements of the FDI mechanism.

Usually [Zhang and Jiang, 2008], the detection and reconfiguration parts of a fault tolerant control (FTC) scheme are treated separately thus neglecting reciprocal influences and substandard behavior (e.g. missed faults). The proposed scheme, based on set theoretic methods, integrates all the FTC components, and analyzes their interactions, to create an overall system with guaranteed fault tolerance properties. We enhance upon previous versions of this scheme (see Seron et al. [2008], Olaru et al. [2010]) by explicitly constraining the relevant signals into feasible domains which permit an exact FDI mechanism. Firstly, we consider that the feedback gain is fix and that the state reference is a free parameter. In these conditions we describe the constraints upon the state reference which permit exact fault detection and use this information in the construction of a reference governor. Secondly, we consider the converse case, where the state reference is fix and the feedback gain is variable. We obtain similarly, a set of constraints bounding the tracking error and use them into a receding horizon optimization procedure in order to deliver a feedback action guaranteeing exact fault detection. In both cases we consider and discuss several selection policies and make use of set theoretic elements to minimize the numerical difficulties (in particular set invariance and the "tube-MPC" construction of Mayne et al. [2005]).

The rest of the paper is organized as follows. In Section 2 the multisensor scheme used in the paper is described and some set theoretic issues are presented. The FDI mechanism is briefly described in Section 3 and the various aspects of the RC mechanism are presented in Section 4. Finally, some conclusions are drawn in Section 5.

Notation

Let $x_{[c_1,c_2]} = \left[x(k+c_1)^T \ \dots \ x(k+c_2)^T\right]^T$ with $c_1, c_2 \in \mathbb{N}^+$ denote a column vector of elements whose index increases monotonically and where $k \in \mathbb{N}^+$ denotes the current instant of time. Whenever $c_1 = c_2 = c$ the shorthand notation $x_{[c]}$ may be employed. Notation x^+ (x^-) denotes the successor (predecessor) element to the current value of x (i.e., $x_{[1]}$ and $x_{[-1]}$).

The Minkowski sum of two sets, A and B is denoted as $A \oplus B = \{x : \ x = a + b, \ a \in A, b \in B\}$.

2. PRELIMINARIES

2.1 Plant description

As a benchmark for the FTC scheme we use a multisensor scheme (depicted in Figure 1) composed from the following components: plant, sensor/estimator and control mechanism.

The plant P has a linear discrete-time state space model:

$$x^+ = Ax + Bu + Ew \tag{1}$$

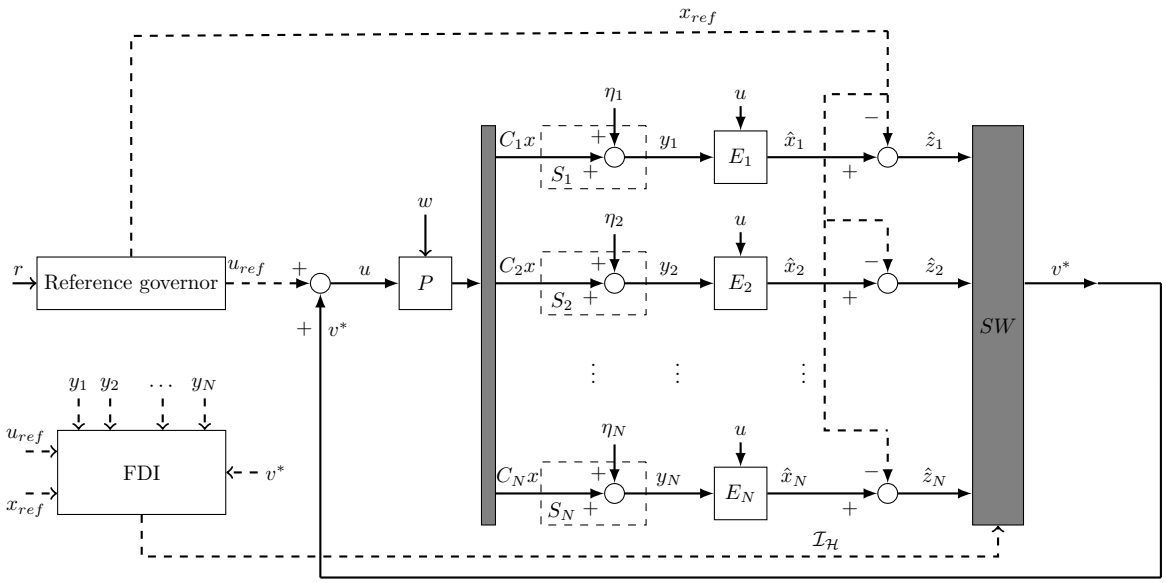


Fig. 1. Multisensor fault tolerant control scheme

where $x \in \mathbb{R}^n$ is the system state, $u \in \mathbb{R}^m$ is the input, and $w \in W \subset \mathbb{R}^r$ is the bounded process disturbance (by the bounded polyhedral set W).

The control problem is to design a closed-loop control scheme such that the state of the plant (1) tracks a reference signal x_{ref} which obeys the nominal dynamics

$$x_{ref}^+ = Ax_{ref} + Bu_{ref} \tag{2}$$

where u_{ref} is a known (stabilizing in case that A is not stable) signal.

The plant tracking error is given by the difference between the state (1) and its respective reference signal (2):

$$z^+ = x^+ - x_{ref}^+ = Az + B\underbrace{(u - u_{ref})}_{v} + Ew. \tag{3}$$

The state of the plant (1) is assumed to not be directly accessible. As such, a bank of sensors S_i with $i \in \mathcal{I} \triangleq \{1 \dots N\}$, measuring linear combinations of the system state, $C_i x \in \mathbb{R}^{p_i}$, are considered. The associated output signals y_i are affected by bounded measurement noises $\eta_i \in N_i \subset \mathbb{R}^{p_i}$, with N_i bounded polyhedral sets:

$$y_i = C_i x + \eta_i. \tag{4}$$

Assumption 1. The pairs (A, C_i) associated to the i^{th} sensing channels are observable. ♦

To each sensor output y_i we associate a state estimator E_i (see Figure 1). The corresponding state estimation $\hat{x}_i$ will be constructed to provide an adequate dynamic behavior:

$$\hat{x}_i^+ = A\hat{x}_i + Bu + L_i\left(y_i - C_i\hat{x}_i\right). \tag{5}$$

The gains L_i are chosen such that matrices $A_{L_i} = A - L_iC_i$ are strictly stable (always possible by Assumption 1).

By virtue of (1) and (5), the estimation error

$$\tilde{x}_i \triangleq x - \hat{x}_i \tag{6}$$

associated to the i^{th} sensor satisfies the relation:

$$\tilde{x}_i^+ = x^+ - \hat{x}_i^+ = A_{L_i}\tilde{x}_i + Ew - L_i\eta_i. \tag{7}$$

For further use we define the estimation tracking error as

$$\hat{z}_i \triangleq \hat{x}_i - x_{ref}. \tag{8}$$

The control mechanism (indicated as SW in Figure 1) can be any control law

$$u = u_{ref} + v^* \tag{9}$$

stabilizing the system. We will detail the control strategies in Section 4.

2.2 Set theoretic issues

In the rest of the paper we will apply a fault tolerant control (FTC) scheme in order to detect and isolate faults and subsequently, design a reconfiguration mechanism for the control action. In both operations we will make use of set theoretic elements, in the sense that we will translate the various events of interest (i.e., fault detection, sensor recovery, control design) into set operations (e.g., the fault detection will be equated with a set membership testing). This kind of approach was used before in the literature, see for example [Planchon and Lunze, 2008, Reppa and Tzes, 2008], but usually the sets defining the signals of interest are computed iteratively. In here we use a technique presented in Seron et al. [2008] and further developed in Olaru et al. [2010]. Instead of reactualizing the sets at each iteration we compute invariant sets for the signals of interest (defined in Definition 1). This minimizes the computations (during runtime only set membership tests will be executed) and permits the analysis of global stability.

We consider the following definitions in set invariance analysis (see for example Blanchini [1999]).

Definition 1. Consider the dynamics $x^+ = Ax + \delta$ with $\delta \in \Delta \subset \mathbb{R}^n$. The set $\Phi \in \mathbb{R}^n$ is a *robust positively invariant* (RPI) set for the given dynamics if $Ax + \delta \in \Phi$ for all $x \in \Phi$ and for all $\delta \in \Delta$, or, equivalently, if $A\Phi \oplus \Delta \subseteq \Phi$. ♦

Definition 2. The *minimal robust positively invariant* (mRPI) set for a given dynamic is the RPI set that is contained in every RPI set for the same dynamic system. ♦

The latter notion is useful because often we prefer sets as "tight" as possible (in order to have a better chance at fulfilling separation conditions for example). In general, it is not possible to compute an exact representation of the mRPI set, except under restrictive assumptions such as when matrix A is nilpotent [Mayne and Schroeder, 1997]. One then needs to resort to approximations, and different algorithms for the construction of RPI approximations can be found in the literature, see for example Raković et al. [2005] and Olaru et al. [2008]. However, those approaches, based on set iterations, focus on the quality of the approximation disregarding its complexity. On the other hand, ultimate bounds sets [Kofman et al., 2007] offer an alternative for the construction of RPI sets of low complexity at the price of increased conservativeness. Finally, it remains a personal choice to strike a balance between computation time and accuracy of approximation and chose a particular set construction technique.

3. FAULT DETECTION AND ISOLATION

The faults considered here are *abrupt* and *total*[1] sensor output outages. The failure is then represented by the following switch in the structure of the observation equation:

$$y_i = C_i x + \eta_i \underset{RECOVERY}{\overset{FAULT}{\rightleftarrows}} y_i = 0x + \eta_i. \tag{10}$$

The noise affecting the observation channel during the fault, $\eta_i^F \in N_i^F \subset \mathbb{R}^{p_i}$, with N_i^F a bounded polyhedral set, may be different from the one during the healthy functioning, η_i.

A signal called residual [Blanke et al., 2006], sensitive to fault occurrences and with a manageable dependence on the disturbances, will be defined for the detection of faults. We consider here a simple parity equation as our residual:

$$r_i = y_i - C_i x_{ref}. \tag{11}$$

Using the sensors' output as defined in (4) and in (10), respectively, it is possible to revisit (11) and obtain the "healthy" and "faulty" residuals:

$$r_i^H = C_i z + \eta_i,\ r_i^F = -C_i x_{ref} + \eta_i^F. \tag{12}$$

Considering the bounds upon the noises η_i, η_i^F we have the following set condition which permits to guarantee fault detection and isolation:

$$\underbrace{(\{C_i z\} \oplus N_i)}_{R_i^H(z)} \cap \underbrace{\left(\{-C_i x_{ref}\} \oplus N_i^F\right)}_{R_i^F(x_{ref})} = \emptyset. \tag{13}$$

We can partition the sensor indices into:

- $\mathcal{I}_H$, all the sensors acknowledged healthy (i.e. with healthy functioning (4) and estimation error (7) inside its invariant set $\tilde{S}_i$:

$$\mathcal{I}_H = \left\{i \in \mathcal{I}_H^- : r_i \in R_i^H(z)\right\} \cup \left\{i \in \mathcal{I}_F^- : \tilde{x}_i \in \tilde{S}_i,\ r_i \in R_i^H(z)\right\}$$

  where $\mathcal{I}_H^-$, $\mathcal{I}_F^-$ indicate the sets of healthy, respectively, faulty, sensors at the previous time instant.

- $\mathcal{I}_F$, all the sensors acknowledged faulty (i.e. with faulty functioning (10)):

$$\mathcal{I}_F = \left\{i \in \mathcal{I} : r_i \notin R_i^H(z)\right\}.$$

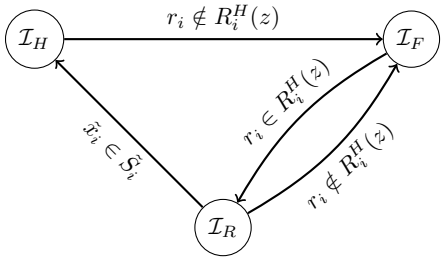


Fig. 2. Depiction of sets $\mathcal{I}_H$, $\mathcal{I}_R$ and $\mathcal{I}_F$ with their transitions and corresponding set membership testings.

Assuming that (13) is validated we conclude that the sets $\mathcal{I}_H$, $\mathcal{I}_R$ and $\mathcal{I}_F$ are disjoint and thus, permit an exact fault detection and isolation. More precisely, the detection and isolation of a fault is equivalent with transitions $\mathcal{I}_H \to \mathcal{I}_F$ and $\mathcal{I}_R \to \mathcal{I}_F$ whereas the eventual switch to healthy functioning from a previous faulty mode is described by transition $\mathcal{I}_F \to \mathcal{I}_R$. Note that the set $\mathcal{I}_R$ acts as a "quarantine" set: the fact that a sensor recovered its healthy functioning is not sufficient, there are signals whose transient behavior is not acceptable. To this end, we consider a sensor ready to participate again in the design of the feedback when its estimation error has suitable values $(\tilde{x}_i \in \tilde{S}_i)$ and describe this event by transition $\mathcal{I}_R \to \mathcal{I}_H$.

Remark 2. Note that condition $r_i \in R_i^H$ depends on how the tracking error z is bounded and that condition $\tilde{x}_i \in \tilde{S}_i$ is not verifiable in its actual form since the signal $\tilde{x}_i$ is not directly measurable. We direct the interested reader to Olaru et al. [2009], Stoican et al. [2010a] for a geometrical analysis providing necessary and sufficient conditions for the aforementioned inclusion and to Stoican et al. [2010b] for a timer/convergence time analysis. ♦

For exemplification we depict in Figure 2 the partition subsets together with their transitions and the corresponding set membership testings.

4. RECONFIGURATION OF THE CONTROL ACTION

The control action actuating the plant can be decomposed into the feedforward part (the reference control action u_{ref}) and the feedback counterpart (the control action v which uses the information provided by the sensors).

In order to make the scheme fault tolerant we need to take into account the results given by the FDI mechanism. Firstly, we will limit the sensors pooling for the design of the feedback action to only the healthy ones (the sensors with indices in $\mathcal{I}_H$). Secondly, we design the control actions such that we increase the range of values for which the set separation (13) is guaranteed.

Taking into account these requirements and the sensor partitioning under healthy/faulty functioning we are capable to provide several feasible control strategies. We consider in this paper two main directions:

[1] More complex scenarios can be considered (partial outage, smooth degradation) but they do not add significantly to the problem, only increase its complexity.

- the input reference is provided by a reference governor which keeps the state reference close to an ideal reference and inside a given feasible region;
- The feedback control action can be either fix or the result of a sliding horizon optimization problem.

4.1 Fixed gain feedback

We start with the design of a fix gain feedback. We choose here to select at each instant of time one of the estimation tracking errors (8) from the set of available sensors (with indices in $\mathcal{I}_H$) and to consider a stabilizing gain matrix K for the feedback action $v = -K\hat{z}_i$, $i \in \mathcal{I}_H$ such that a given cost function will be minimized:

$$l = \arg\min_{i \in \mathcal{I}_H} \mathcal{J}(\hat{z}_i). \tag{14}$$

where $l \in \mathcal{I}_H$ is the selected index.

Then, using (6), (9) can be reformulated as

$$u = u_{ref} + K(z - \tilde{x}_l). \tag{15}$$

Substituting (15) in (3), leads to

$$z^+ = (A + BK)z + Ew - BK\tilde{x}_l. \tag{16}$$

Remark 3. Equally well, we could have been used a more traditional sensor fusion technique (e.g., by taking as the control feedback action as a convex sum of the available estimation tracking error). However, this leads to complicated representations and the gain in the cost function minimization is reduced: it was shown in Seron et al. [2009] that the switching between available sensors leads to a "leveling" effect which mimics traditional fusion strategies. ♦

The main question which we may pose in this context is the choice of the cost function. A classical choice is to consider the current estimation errors and select the index of the one minimizing the cost function:

$$\begin{aligned} v* &= -K\hat{z}_l \\ l &= \arg\min_{i \in \mathcal{I}_H} \{||\hat{z}_i||_Q + ||v||_R\}, \end{aligned} \tag{17}$$

where Q and R are suitable chosen weight matrices.

Here, however, we propose to use a receding horizon technique, where the current selection in (14) takes into account future possible switches in order to minimize a cost function over a finite horizon.

We enumerate several approaches which, with increasing degree of flexibility, take explicitly in consideration the way the switch operates. To this end, we recall the dynamic equation describing each state estimation error (5) and substract the state reference (2) in order to obtain the dynamic equation for the plant estimated tracking error [2] (8) by each sensor-estimation pair:

$$\hat{z}_i^+ = A\hat{z}_i + Bv + L_iC_i\tilde{x}_i + L_i\eta_i. \tag{18}$$

With this notation we point to three receding horizon implementations with different flavors according to the choice of the objective function or the constraints to be fulfilled by the group of sensors.

"Individual merit" selection. Here the sensors are compared [3] with respect to their individual cost-to-go for the given initial conditions and the index with the best "individual merit" is selected for the feedback control action. This can be seen as an "elitist" type of multi-agent formulation.

$$\begin{aligned} v &= -K\hat{z}_{i^*} \\ i^* &= \arg\min_{i \in \mathbf{I}_H} \left\{ \sum_{j=0}^{\tau-1} \left(||\hat{z}_{i[j]}||_Q + ||v_{[j]}||_R \right) + ||\hat{z}_{i[\tau]}||_P \right\} \\ &\text{s.t.:} \\ \hat{z}^+_{i[j]} &= A\hat{z}_{i[j]} + Bv_{[j]}. \end{aligned} \tag{19}$$

"Relay race". Here switchings are allowed along the prediction horizon between the estimators which build the control action. The predictions are still performed in parallel, but the global cost can benefit from the changes of index along the prediction horizon. This can be seen as a multi-agent system in which the leader can change at each stage of the prediction horizon.

$$\begin{aligned} v &= -K\hat{z}_{i_0^*} \\ \{i_0^*, \ldots, i_{\tau-1}^*\} &= \arg\min_{i_j \in \mathbf{I}_H} \left\{ \sum_{j=0}^{\tau-1} \left(||\hat{z}_{i_j[j]}||_Q + ||v_{[j]}||_R \right) + \right. \\ &\qquad \left. ||\hat{z}_{i_\tau[\tau]}||_P \right\} \\ &\text{s.t.:} \\ \hat{z}^+_{i[j]} &= A\hat{z}_{i[j]} + Bv_{[j]},\ j = 0 \ldots \tau - 1. \end{aligned} \tag{20}$$

"Collaborative" scenario. Here the cost index allows switching during the prediction horizon and the terminal penalty is considered with respect to a combination of predicted estimation errors. This approach can be seen as a collaborative multi-agent decision: along the prediction horizon, all the agents apply the same control policy. The performance of the group in the given horizon is given by the summation of the performance of the best individual at each stage.

$$\begin{aligned} v &= -K\hat{z}_{i_0^*} \\ \{i_0^*, \ldots, i_{\tau-1}^*\} &= \arg\min_{i_j \in \mathbf{I}_H} \left\{ \sum_{j=0}^{\tau-1} \left(||\hat{z}_{i_j[j]}||_Q + ||v_{[j]}||_R \right) + \right. \\ &\qquad \left. ||\hat{z}^*_{[\tau]}||_P \right\} \\ &\text{s.t.:} \\ \hat{z}^+_{i[j]} &= A\hat{z}_{i[j]} + Bv_{[j]},\ j \in \{0 \ldots \tau - 1\} \\ \hat{z}^*_{[\tau]} &\in conv\left\{ \hat{z}_{i[\tau]} \right\}_{i \in \mathbf{I}_H}. \end{aligned} \tag{21}$$

Notice that the decision based on individual cost evaluation does not exploit the degrees of freedom offered by the prediction window. It can be reduced in fact to the comparison of cost indices for different estimations. The advantage of such a scheme lies in the simplicity of its implementation. On the other hand, the second and third schemes propose optimization problems which belong to the class of mixed integer programming problems and the combinatorial complexity of their discrete decisions grows with the prediction horizon.

[2] Assuming of course healthy functioning for sensor output y_i which is granted as long as $i \in \mathcal{I}_H$.
[3] Note that we discarded the noises from relation (18) to simplify the formulation of the problem.

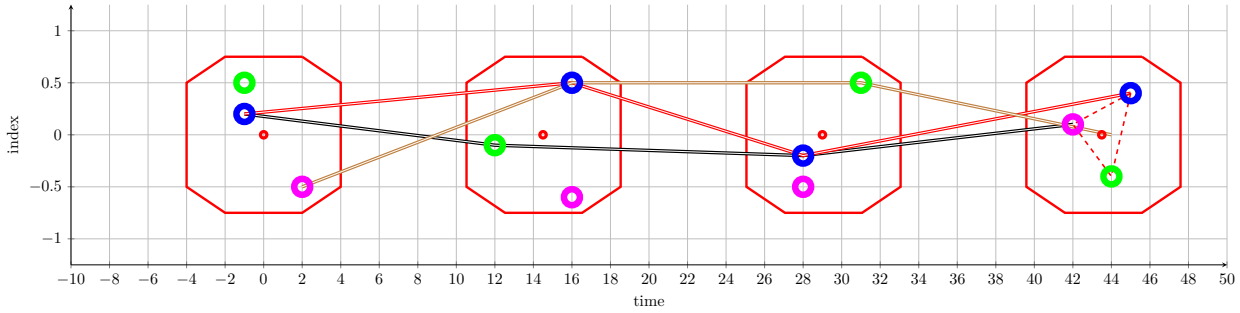


Fig. 3. Depiction of "individual merit", "relay race" and "collaborative" selection strategies.

The tuning rules are not mature and they have been seldom been tested. With the development of the cooperative MPC techniques, such approaches can present a certain interest as a future research direction in the FTC-MPC.

A "proof of concept" depiction of the strategies described above is shown in Figure 3. The color-coded line segments describe the selection made at each instance by a specific selection strategy from the set of available sensors.

4.2 Reference governor

Assuming a fixed gain as described in the previous section it is now possible to construct a robust positively invariant (RPI) set S_z associated to dynamics (16) and to use it in (13) in order to obtain a feasible domain for x_{ref}:

$$\mathbb{D}_{x_{ref}} = \{x_{ref} : R_i^F(x_{ref}) \oplus R_i^H(S_z) = \emptyset,\ i \in \mathcal{I}\}. \quad (22)$$

Note that any value $x_{ref} \in \mathbb{D}_{x_{ref}}$ will respect by its very definition the set separation (13) which guarantees exact fault detection and isolation.

We can now use this constraint in a reference governor setting. Supposing an ideal reference r we are able to provide through a reference governor respecting constraint (22) a feasible pair of state/input reference (u_{ref}, x_{ref}):

$$u^*_{ref[0,\tau-1]} = \underset{u_{ref[0,\tau-1]}}{\arg\min} \sum_{i=0}^{\tau-1} \left(||r_{[i]} - x_{ref[i]}||_{Q_r} + ||u_{ref[i]}||_{R_r}\right)$$

subject to

$$\begin{aligned} x^+_{ref[i]} &= Ax_{ref[i]} + Bu_{ref[i]} \\ x^+_{ref[i]} &\in \mathbb{D}_{x_{ref}} \end{aligned} \quad (23)$$

where τ is the prediction horizon and cost matrices Q_r, R_r are appropriately chosen. A "proof of concept" depiction of this mechanism is shown in Figure 4 where an ideal reference passes through an interdicted region and consequently, the reference governor provides a feasible reference which respects the boundaries of $\mathbb{D}_{x_{ref}}$.

Remark 4. One can observe that the set (22) is the complement of the union of N convex regions (i.e., $R_i^F(x_{ref}) \oplus R_i^H(S_z) = \emptyset$ is equivalent with $-C_i x_{ref} \notin R_i^H(S_z) \oplus (-N_i^F)$). As a consequence, the optimization problem has to be solved over a nonconvex set which imposes the use of mixed-integer techniques – Osiadacz [1990]. To alleviate the computational burden specific to these techniques one can reduce the number of auxiliary variables as in Stoican et al. [2011]. ♦

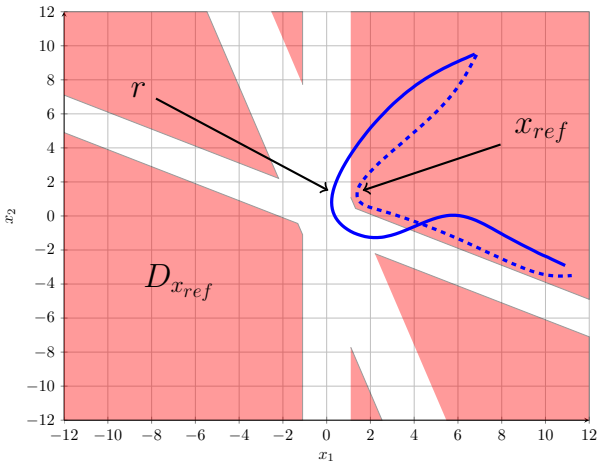


Fig. 4. Depiction of admissible region $\mathbb{D}_{ref}$ characterizing x_{ref}.

4.3 MPC techniques

Up to this point we assumed that the feedback control is given through a fixed gain matrix. This has the advantage of giving an easy to compute invariant set for the plant tracking error but is, on the other hand, limited in its reach by the fix structure. The use of MPC techniques for computing the feedback (and ultimately, the feedforward) control action(s) relaxes these constraints by providing a time-varying feedback control structure [Bitmead et al., 1990, Maciejowski, 2002].

Similarly with the construction (22) we assume now the converse: the reference state is bounded ($x_{ref} \in X_{ref}$ with $X_{ref} \subset \mathbb{R}^n$) and the tracking error z has to reside into a feasible (from the point of view of condition (13)) domain [4]:

$$\mathbb{D}_z = \{z : R_i^F(X_{ref}) \cap R_i^H(z) = \emptyset,\ i \in \mathcal{I}\}. \quad (24)$$

This permits to write the following optimization problem for the feedback:

$$v^*_{[0,\tau-1]} = \underset{v_{[0,\tau-1]}}{\arg\min} \left\{ \sum_{i=0}^{\tau-1} \left(||z_{[i]}||_Q + ||v_{[i]}||_R\right) + ||z_{[\tau]}||_P \right\} \quad (25)$$

subject to:

[4] A similar argument to the one in Remark 4 can be drawn for this set: the result will be nonconvex and consequently, mixed integer techniques will have to be used in any optimization problem involving it.

$$\begin{aligned} z^+_{[i]} &= Az_{[i]} + Bv_{[i]} + Ew_{[i]} \\ z^+_{[i]} &\in D_z \end{aligned}, \quad i = 0 \dots \tau - 1 \qquad (26)$$

where τ is the prediction horizon, and $Q \in \mathbb{R}^{n\times n}$, $P \in \mathbb{R}^{n\times n}$ and $R \in \mathbb{R}^{m\times m}$ are weighting matrices.

Although easy to write in a compact finite-time optimization formulation, the above relations suffer from a list of difficult to handle particularities. The foremost is that the plant tracking error z is not directly measurable and as such, its estimations must be used (based for example on the currently healthy sensors). Even so, the future values of z are set-valued by the presence of the plant noise w, leading practically to a robust MPC formulation. As a consequence, the optimization problem becomes difficult to solve in real-time (see Kerrigan and Maciejowski [2004]).

A tube predictive control philosophy [Mayne et al., 2006] can in this case be considered as an alternative. This approach presumes the construction of a "nominal" plant tracking error dynamics:

$$z^+_{nom} = Az_{nom} + Bv_{nom} \qquad (27)$$

where, due the absence of noise, the "nominal plant tracking error" is directly predictable. If additionally, we consider the nominal feedback control v_{nom} and take it as

$$v_{nom} \triangleq v + K\left(\hat{z}^* - z_{nom}\right) \qquad (28)$$

where $\bar{z} \triangleq z - z_{nom}$ and $\hat{z}^*$ denotes an estimation of true, unmeasurable, value. We are now able to describe the dynamic relation characterizing $\bar{z}$:

$$\begin{aligned} \bar{z}^+ &= A\left(z - z_{nom}\right) - BK\left(\hat{z}_l - z_{nom}\right) + v_{nom} - v_{nom} + Ew \\ &= A\bar{z} - BK\left(z - \tilde{x}^* - z_{nom}\right) + Ew \\ &= \left(A - BK\right)\bar{z} + BK\tilde{x}^* + Ew \end{aligned} \qquad (29)$$

where $\tilde{x}^* \triangleq z - \hat{z}^*$. If an addition we provide a set $\bar{S}_z$ such that $\bar{z} \in \bar{S}_z$ at any instant of time we have that

$$z \in \{z_{nom}\} \oplus \bar{S}_z. \qquad (30)$$

Remark 5. Note that, as was the case in Section 4.1, we have several possibilities of selection for the nominal dynamic z_{nom}. The choice will be influenced by the sought after degree of flexibility of the scheme together with the acceptable complexity for the representation of set $\bar{S}_z$ (e.g., in the case where we consider an arbitrary switch, we can safely construct $\bar{S}_z$ as an invariant set whch is numerically identical with S_z). ♦

With these elements it is straightforward to rewrite (25)–(26) into:

$$v^*_{nom[0,\tau-1]} = \underset{v_{nom[0,\tau-1]}}{\arg\min} \left\{ \sum_{i=0}^{\tau-1} \left(||z_{nom[i]}||_Q + ||v_{nom[i]}||_R \right) + ||z_{nom[\tau]}||_P \right\} \qquad (31)$$

subject to:

$$\begin{aligned} z^+_{nom[i]} &= Az_{nom[i]} + Bv_{nom[i]} \\ z^+_{nom[i]} &\in D_z \ominus \bar{S}_z \end{aligned}, \quad i = 0 \dots \tau - 1 \qquad (32)$$

with the same notations as before and making use of the fact that relation $z_{nom} \in D_z \ominus \bar{S}_z$ implies $z \in D_z$.

For illustration purposes a qualitative depiction of the "tube" behavior is given in Figure 5, it can be seen that the tracking error remains at all times around the nominal trajectory.

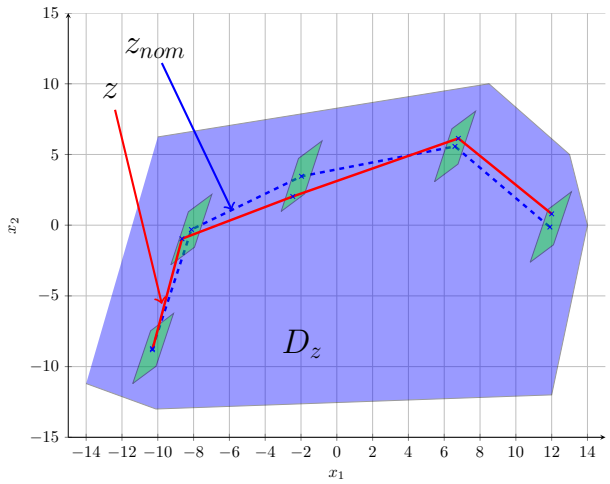


Fig. 5. Depiction of "individual merit", "relay race" and "collaborative" selection strategies.

Remark 6. If the set D_z is too tight then it may become impossible to respect condition (13). Then we may apply the same technique as in Section 4.2 where the state reference is considered to be also a decision variable and we can formulate an extended MPC optimization problem which provides both reference input u_{ref} and nominal feedback control v_{nom} such that condition (13) is verified. We will not repeat the optimization problems discussed anteriorly but note that the feasible set is now defined as

$$\mathbb{D}_{ref} = \{(x_{ref}, z) : \; R_i^F(x_{ref}) \oplus R_i^H(z) = \emptyset, \; i \in \mathcal{I}\} \qquad (33)$$

and as long as there are no additional constraints on the tracking error and the state reference, we have recursive stability and thus we guarantee that all times the optimization problem is feasible. ♦

5. CONCLUSIONS

In this paper we revisited a FTC scheme realized in a set theoretic framework. We used FDI-derived conditions to describe feasible regions for state reference/tracking error. We further used these constraints to design optimal feedforward (as the result of a reference governor) and feedback actions (using either fix gain or a MPC methodology). Lastly, we discussed different selection strategies from the available sensors.

REFERENCES

R.R. Bitmead, M. Gevers, and V. Wertz. *Adaptive Optimal Control: The Thinking Man\'s {G}{P}{C}.* Prentice Hall, 1990.

F. Blanchini. Set invariance in control–a survey. *Automatica*, 35(11):1747–1767, 1999.

M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki. *Diagnosis and fault-tolerant control.* 2006.

E.C. Kerrigan and J.M. Maciejowski. Feedback min-max model predictive control using a single linear program: robust stability and the explicit solution. *International Journal of Robust and Nonlinear Control*, 14(4):395–413, 2004.

Ernesto Kofman, Hernan Haimovich, and María M. Seron. A systematic method to obtain ultimate bounds for perturbed systems. *International Journal of Control*, 80(2):167–178, 2007.

J.M. Maciejowski. *Predictive control: with constraints*. Pearson education, 2002.

J.M. Maciejowski and C.N. Jones. MPC fault-tolerant flight control case study: Flight 1862. In *Proceedings of the 4th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, Washington, DC, USA, June 2003.

David Q. Mayne and W.R. Schroeder. Robust time-optimal control of constrained linear systems. *Automatica*, 33:2103–2118, 1997.

David Q. Mayne, María M. Seron, and Sasa V. Raković. Robust model predictive control of constrained linear systems with bounded disturbances. *Automatica*, 41(2): 219–224, 2005.

DQ Mayne, SV Rakovic, R. Findeisen, and F. Allgöwer. Robust output feedback model predictive control of constrained linear systems. *Automatica*, 42(7):1217–1222, 2006.

Sorin Olaru, José A. De Doná, and María M. Seron. Positive invariant sets for fault tolerant multisensor control schemes. In *Proc. of the 17th IFAC World Congress*, pages 1224–1229, Seoul, South Korea, 6-11 July 2008.

Sorin Olaru, Florin Stoican, José A. De Doná, and María M. Seron. Necessary and sufficient conditions for sensor recovery in a multisensor control scheme. In *Proc. of the 7th IFAC Symp. on Fault Detection, Supervision and Safety of Technical Processes*, pages 977–982, Barcelona, Spain, 30 June-3 July 2009.

Sorin Olaru, José A. De Doná, María M. Seron, and Florin Stoican. Positive invariant sets for fault tolerant multisensor control schemes. *International Journal of Control*, 83(12):2622–2640, 2010.

A.J. Osiadacz. Integer and combinatorial optimization, George L. Nemhauser and Laurence A. Wolsey, Wiley-Interscience Series in Discrete Mathematics and Optimization, New York, 1988, ISBN 0-471-82819-X, 763pp. *International Journal of Adaptive Control and Signal Processing*, 4(4):333–334, 1990.

P. Planchon and J. Lunze. Diagnosis of linear systems with structured uncertainties based on guaranteed state observation. *International Journal of Control Automation and Systems*, 6(3):306–319, June 2008.

Sasa V. Raković, Eric C. Kerrigan, Koustas I. Kouramas, and David Q. Mayne. Invariant approximations of the minimal robust positively invariant set. *IEEE Transactions on Automatic Control*, 50(3):406–410, 2005.

V. Reppa and A. Tzes. Fault detection based on orthotopic set membership identification for robot manipulators. In *Proc. of the 17th IFAC World Congress*, pages 7344–7349, Seoul, South Korea, 6-11 July 2008.

María M. Seron, Xiang W. Zhuo, José A. De Doná, and J.J. Martinez. Multisensor switching control strategy with fault tolerance guarantees. *Automatica*, 44(1):88–97, 2008. ISSN 0005-1098.

María M. Seron, José A. De Doná, and Sorin Olaru. Multisensor fault tolerant control allowing sensor healthy-to-faulty and faulty-to-healthy transitions. *under review*, 2009.

Florin Stoican, Sorin Olaru, José A. De Doná, and María M. Seron. Improvements in the sensor recovery mechanism for a multisensor control scheme. In *Proceedings of the 29th American Control Conference*, pages 4052–4057, Baltimore, Maryland, USA, 30 June-2 July 2010a.

Florin Stoican, Sorin Olaru, María M. Seron, and José A. De Doná. A fault tolerant control scheme based on sensor switching and dwell time. In *Proceedings of the 49th IEEE Conference on Decision and Control*, Atlanta, Georgia, USA, 15-17 December 2010b.

Florin Stoican, Ionela Prodan, and Sorin Olaru. On the hyperplanes arrangements in mixed-integer techniques. In *Proceedings of the 30th American Control Conference*, pages 1898–1903, San Francisco, California, USA, 29 June-1 July 2011.

Y. Zhang and J. Jiang. Bibliographical review on reconfigurable fault-tolerant control systems. *Automation and Remote Control*, 32(2):229–252, 2008.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Event-driven flow control of incompressible fluid

Sebastian Roll, Heinz A. Preisig

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

Abstract: This paper introduces a simple self-calibrating liquid flow rate control system. A level glass containing discrete optical level sensors is connected on the flow line upstream of the pump. The optical sensors in the level glass provide event information and define volumes that can be used to calibrate and control the pump. The system is continuous and is observed by a quantiser, making it overall a discrete-event dynamic system to be controlled by a corresponding controller.
The set-up was installed as a reflux system on a set of distillation columns where the distillate is condensed below the top of the column, providing a industrial-like physical arrangement of the different components. The control system is low-cost and therefore provides a reasonable alternative for flow control and flow measurement.

1. INTRODUCTION

Flow control is a common process operation. It normally involves a flow measurement which is used to control a pump of one or the other kind. The alternative is a feeding pump, which is an accurate volume transport device. A feeding pump requires a construction that makes the flow rate nearly independent of the pressure conditions on either side of the pump, which then makes it possible to calibrate the pump, thereby relating the control input to the pumping rate.

In this project we were seeking a cheap and simple alternative for both components. The flow measurement should not only be simple but also robust and the pump should have the basic properties of a feeding pump. The result was a self-calibrating liquid flow control system that uses a level glass for observing the condition of having the outflow of the pump match the supply of the fluid on the inflow. The level glass also serves the purpose of a volume standard being used for the on-line auto-calibration of the pump. The arrangement is used as a reflux pump and flow measurement system on a distillation column.

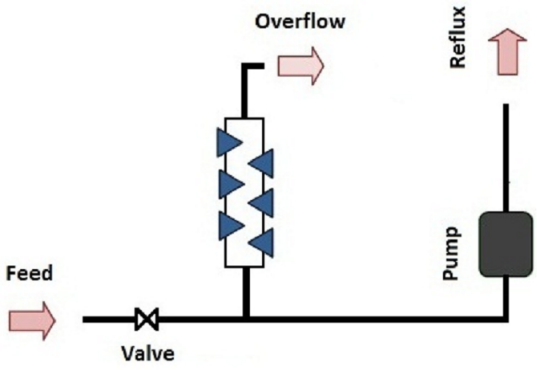


Fig. 1. Schematic of event-driven flow control

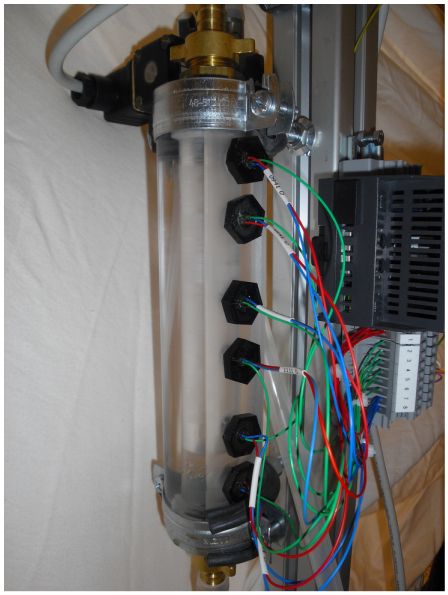


Fig. 2. Pipe with discrete optical liquid level sensors, used for flow control

2. PHYSICAL ARRANGEMENT

The physical arrangement consists of two main components: a level glass and a pump. The level glass is connected to the flow line before the pump via a generously wide connection so as to not delay the level change in the level glass. The pump then is put into the flow line for moving the fluid. The level glass is equipped with six discrete level detectors, which provide an event-signal whenever the level passes an optical sensor. Knowing the dimension of the level glass and the location of the sensors provides a set of known volumes, which can be used for the calibration and control of the pump. The dimension of the level glass is chosen to be in a reasonable relation to the expected range of flow rates in the line and, as a second consideration, the

size must be such that the detectors can be placed along the length reasonably easily. In addition, a solenoid tab is put into the line before the level class connection so as to interrupt the supply flow, enabling the on-line calibration of the pump.

The peristaltic pump provides the volume transport over a reasonable range of inflow/outflow pressure of the pump and generates a head that is suitable for our application, namely pumping the distillate that we condense on the side of the column, back to the top, thus providing the necessary reflux. Volume transporters are usually moving small constant volumes generated in the pump by toothed wheels or other mechanisms. This makes the flow pulsating, which has a negative effect on the accuracy of the flow rate estimates.

3. DISCRETE-EVENT DYNAMIC CONTROLLER

The overall system is a hybrid of a continuous system and a discrete-event observed system. With the measurement being event-based, the controller is event-based. The task is to design a discrete-event dynamic controller, a DED controller. For the design of the DED controller, we use the techniques discussed in Preisig [1996], Philips [2001], Philips et al. [2003]. The one-dimensional state space, being the level in the level glass, is quantised. The number and location of the discrete sensors determine the discretisation of the state space, dividing it into seven discrete states. Five of those states corresponds to the volume space between two sensors. In addition there are the two additional volume spaces, one above the top sensor and one below the bottom sensor. The sensors report via a process interface (digital input) an event of passing a sensor to the computing device and thus to the control algorithm.

3.1 Time event

The DED controller is designed to keep the level within the middle domain. If the supply flow increases, the level will increase and eventually when the level reaches the upper boarder of the middle domain a state event will be reported. At this point the pump rate will be increased. The aim is to bring the process back to the intermediate domain. If the increase of the pump rate was too small and it happens that the level settles somewhere in the current upper domain, no event will occur.

To get insight into the problem consider the difference between sampled systems and event-driven systems. In the sampled system, the signal value is read every sampling time instance. So, sampling is driven by a clock. In the case of an event-driven system, the 'reading' is done in the opposite direction, namely the value (state-boundary) is given and if the process passes it an event signal is dispatched. This makes the process being the 'clock'. If one does not have a event detector, one may generate a virtual event by using a kind of prediction in which one replaces the process with a stop watch that is set to a time that corresponds to the expectation that an unobservable event would occur. Thus one can view this mechanism as substituting an unobservable event with a stop-watch event, de facto a time event.

The unobservable event that takes place when the level does not change over a specified time, is replaced by a time event. For certain levels, here the events g, h, i and j in Figure 3, a time event is triggered if a level change event does not occurred within 50 seconds, with the action of altering the pump flow rate. Also a new timer is started and if the time event still does not trigger a level change event after the additional 50 seconds and the time event happens again, the pump flow rate is further changed in the same direction (increased if above and decreased if below the current limit). The time it takes for the time event to trigger depends on the ratio between level volume and flow rate. For a big level glass with large level volumes and with a low flow rate, the time event should not be triggered until a sufficient time has passed, as a level change event gives flow rate information and is therefore more valuable.

In our case we implement this behaviour for the main reason that we do not want to change the speed of the pump abruptly but smoothly. The speed of rotation of the pump is thus changed slowly. As an example, when the liquid level changes from "Level 3" to "Level 4", the level change event c causes the pump speed to be set to equal its current speed plus the estimated liquid flow in the level glass. If the estimate is accurate, the action will only stop the rising of liquid in the glass, and not lower it. The time event h will then increase the pump speed by 2.5% so that it reaches "Level 3" again.

3.2 Control automaton

Figure 3 shows the state diagram of the resulting flow control system. Events that are highlighted in red illustrate when pump flow rate adjustment actions are being taken. These highlighted events occur after a change in liquid level, or when a time-event occurs. The ideal state for the controller is the state labelled "Level 3", which is in the middle of the domains bounded by the six sensors.

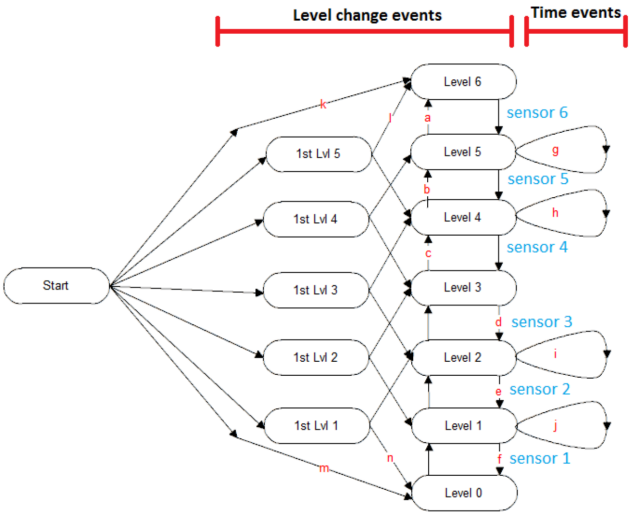


Fig. 3. State diagram

Figure 5 shows the events that are triggered by a change in level and also the events triggered when there is no change in level for a given time, which is relative to the normal

process dynamics rather substantial. The four sensors in the middle acts as soft constraints and the top and bottom sensors acts as hard constraints. On hard constraints the controller "panics" and shifts to maximum action, which is full speed on the top and zero on the bottom.

Symbol	Definition
$\dot{V}_{in}$	Flow rate going in to flow control system
$\dot{V}_p$	Flow rate produced by pump

Fig. 4. Symbol definition for Figure 5.

Letter	Event	Action	
a	Level 5 →Level 6	$\dot{V}_p = max$	Set PFR[1]to maximal allowed value
b	Level 4 → Level 5	$\dot{V}_p = \dot{V}_{in} * 1.10$	Set PFR to 10% above flow rate
c	Level 3 → Level 4	$\dot{V}_p = \dot{V}_{in}$	Set PFR equal to flow rate
d	Level 2 ← Level 3	$\dot{V}_p = \dot{V}_{in}$	Set PFR equal to flow rate
e	Level 1 ← Level 2	$\dot{V}_p = \dot{V}_{in} * 0.90$	Set PFR to 10% below flow rate
f	Level 0 ← Level 1	$\dot{V}_p = 0$	Set PFR to zero
g	dt=[50s,100s,150s]	$\dot{V}_p = \dot{V}_p * 1.05$	Increase PFR by 5%
h	dt=[50s,100s,150s]	$\dot{V}_p = \dot{V}_p * 1.025$	Increase PFR by 2.5%
i	dt=[50s,100s,150s]	$\dot{V}_p = \dot{V}_p * 0.975$	Decrease PFR by 2.5%
j	dt=[50s,100s,150s]	$\dot{V}_p = \dot{V}_p * 0.95$	Decrease PFR by 5%
m	Start → Level 0	$\dot{V}_p = 0$	Set PFR to zero
k	Start → Level 6	$\dot{V}_p = max$	Set PFR to maximal allowed value
n	1st Lvl 1→ Level 0	$\dot{V}_p = 0$	Set PFR to zero
l	1st Lvl 5→ Level 6	$\dot{V}_p = max$	Set PFR to maximal allowed value
→	All other events		Event is ignored

Fig. 5. Time and level change events

The level change events b, c, d and e trigger an action that sets $\dot{V}_p$ (pump flow rate) equal or close to $\dot{V}_{in}$ (flow rate coming in). In order for this to happen, $\dot{V}_{in}$ must be determinable. If the level rises from "Level 3" to "Level 4", it must before that have risen from "Level 2" to "Level 3" in order for $\dot{V}_{in}$ to be known. As an example of the opposite, if the level were to oscillate between "Level 3" and "Level 4", the level change event c will not be triggered at the change from "Level 3" to "Level 4".

4. PUMP CALIBRATION

The solenoid valve introduced into the line enables the isolation of the level glass and the pump from the supply. Since the volumes in the level glass are known, one can use the fluid in the level glass for the calibration of the pump. The procedure is controlled by an calibration automaton. It implements a sequence of operations:

a: open valve	g: compute flow rates
b: close valve	h: change pump speed parameter
c: start pump	i: increase counter by 1
d: stop pump	j: start time
e: start control mode	k: note time
f: start x seconds countdown	

Table 1. Calibration actions

Calibration procedures (i), (ii) and (iii) are done in succession, one time after assembly of the system is finished:

(i) Find correlation between one pump input and flow rate. This is done by specifying a (preferably low) input signal to the pump and collecting the output flow into a container with known volume. The flow rate, $\dot{V}$, for the given input, u, is found by measuring the time, Δt required to fill the volume-calibrated container to volume V:

$$\dot{V}(u) = \frac{V}{\Delta t(u)} \tag{1}$$

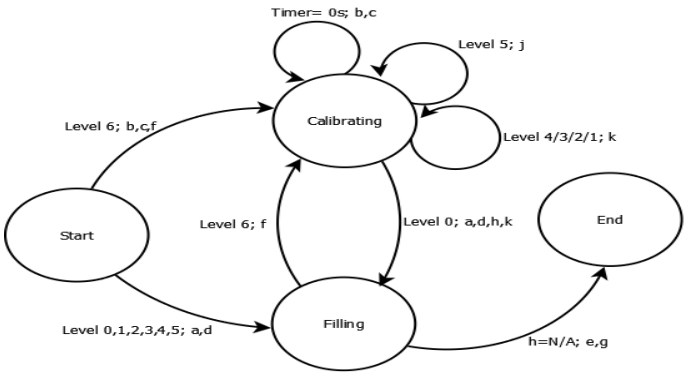


Fig. 6. (ii): Automaton state-flow diagram for level volume identification

(ii) Find volume between sensors using flow rate from the previous volume calibration (i). As one flow rate is known, this is used to calculate the relevant volumes. Figure 6 shows the state-flow diagram for the calibration automaton: The level glass is filled to and beyond the top sensor by opening the valve and stopping the pump. When the liquid level is sufficiently above the top sensor (which is determined by a timer of a few seconds), the valve closes and the pump is started at a given rpm corresponding to input used in (i), thus starting the emptying of the level glass. When the liquid level passes through the top sensor, i.e. enters "Level 5", a timer starts. For each new level entered, the time is noted. When the liquid level reaches "Level 0", the pump is stopped and the valve opens, allowing the liquid to fill the level glass again. This experiment is repeated a given number of times in order to also get statistical information. Once the last experiment is completed, the automaton algorithm switches back to control mode.

The volume of "Level n" is calculated as:

$$V(n) = \dot{V}(u) * (t_n(u) - t_{n-1}(u)) \tag{2}$$

The results are used by the control algorithm. The last calibration procedure uses the volume from top sensor to bottom sensor, $\sum v(i)$.

(iii) Find correlation between pump inputs and flow rates for entire input span, using the equation:

$$\dot{V}(u) = \frac{V}{t_5(u) - t_0(u)} \tag{3}$$

$t_5(u) - t_0(u)$ is the time it takes to the level to drop from "Level 5" to "Level 0", corresponding to the volume between the top and bottom level sensor. This is the same equation as the first calibration procedure, but since the volume of the level glass is now known, is can be used as fixed volume container. The procedure is shown in Figure 7. The level glass is filled and emptied, similar to the procedure explained in (ii), but the algorithm cycles through an array of specified pump speeds instead of

using the same speed for all the calibration cycles. It also does not note the time when the liquid level enters the intermediate levels, but the time is taken only when entering "Level 5" and "Level 0".

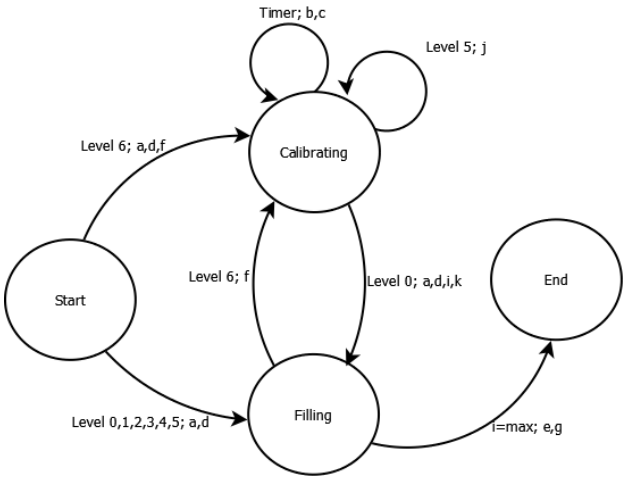


Fig. 7. (iii): Automaton state-flow diagram for flow calibration

5. PRACTICALITIES

5.1 Static volume

Operating at steady flow, the controller keeps the liquid within "Level 3", between the two middle sensors. This effectively traps the liquid in the level glass for an extended periods of time. If the properties of the fluid in the line change with time, then the liquid in the glass will be different. As the flow decreases in the line, fluid from the level glass will come back into the line thus changing temporarily the properties of the pumped fluid. In case of our distillation column, can potentially affect the composition in the column. Whilst the volume in the level glass is small this behaviour should be kept in mind and if it is potentially disturbing the operation of the plant the control automaton is to be modified so as to exchange the fluid in the level glass periodically by taking the necessary control actions. In our current application, the level glass holds ≈50 ml maximum, which under normal operation will not cause any problems.

5.2 Handling overflow

The level glass my overflow under certain operating conditions, for example if the line flow exceeds the maximal pump ration. But may also be caused by sensor or pump malfunction, or air bubbles. The likelihood of overflow can be reduced by having a large volume reservoir above the top sensor and having a time event for "Level 6" that stops or reduces the flow into the level glass. This has the potential downside of larger static volume, which was discussed in the previous subsection. There are multiple ways of handling overflow situations. There is no single optimal solution, and one should choose the one most fitting to the specific application. The following are some suggestions:

- Overflow can be directed by a hose into the waste reservoir. This has the added benefit of preventing overpressure.
- Overflow can be redirected into another stream.
- A one-way output valve could be used for letting air pass freely through without letting liquid pass through the top. The pipe will in this case lose the ability to overflow, creating the possibility of a pressure build-up.

5.3 Maintenance

All parts of the flow control set-up robust and have relatively high standing times, and are easily replaceable if so needed. The liquid level sensors are low price and are simply screwed into the side of the level glass. The level glass is made from one solid piece of see-through plastic material, and is cheap to both in material and manufacturing.

5.4 Regular calibration

Peristaltic pumps have a tendency to deform the hose over time, thereby altering the input/flowrate correlation. It is recommended to regularly check the hose for damage and wear. Since the calibration procedure is completely automated, it can be scheduled to regularly be repeated. Deviations may be taken as a sign for the ageing of the hose triggering a replacement. This also provides the control algorithm regularly with updated flow rate values. The calibration can be run manually or automatically, without stopping the overall process, given that the system allows for some accumulation of mass upstream to the valve. For a distillation set-up, accumulation is allowed in the tubing following the condenser or even in the condenser itself. Accumulation depends on how slowly the level glass is emptied, i.e. the minimum pump input that is being tested.

6. EXPERIMENTAL RESULTS

A distillation column is fitted with the discussed flow control scheme, as shown in Figure 8.

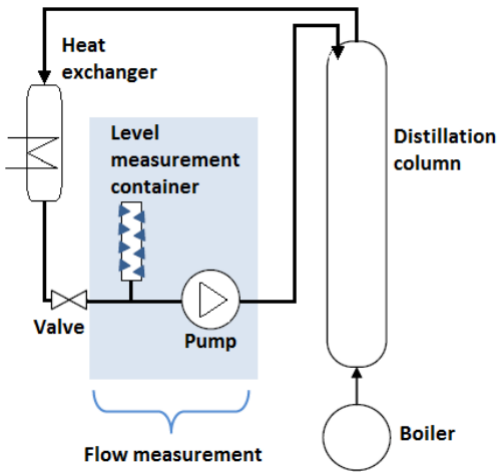


Fig. 8. Distillation column with event-driven flow control

The condensate accumulating at the low end of the condenser is entering the line to the pump in front of which the level glass with optical liquid level sensors is connected. The pump pushes the condensate back to the top of the column providing the required reflux. The control algorithm keeps track of the volumetric flow rate using the calibration relation. The peristaltic pump head is driven by a stepper motor that is equipped with a complete process interface that connects directly to the RS 485 instrumentation bus. The level glass is *approx* 15 cm high and holds ≈ 50 ml between the top and bottom sensor.

A calibration program is realised as a separate program making it possible to run the calibration independent of te control program. The program is also used to determine volumes between the sensors in the level glass.

6.1 Calibration results

Figure 9 shows the correlation between pump rpm, being the input to the pump, and the volumetric flow rate. The graph is linear up to ≈ 90 rpm at which point a "hump" appears. We found that the "hump" is reproducible. The detailed cause is not known, but can likely be attributed to some interaction between deformation of the hose and the workings of the peristaltic pump. The pump rarely exceeds 90 rpm when in normal operation, but the maximum speed setting that starts when the liquid enters "Level 6", is 150 rpm.

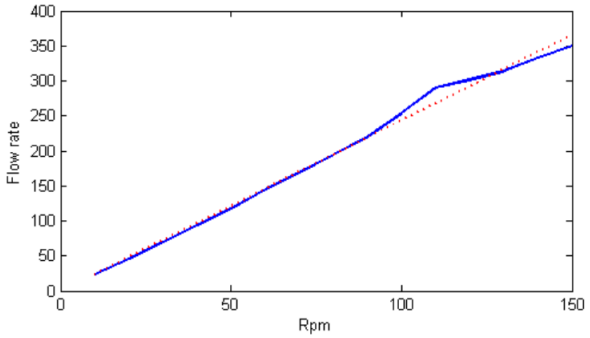


Fig. 9. line: Results from calibration procedure 3. Red line: linear regression.

The use of a linear correlation is easy and convenient, but can be inaccurate when getting into the non-linear region. If so necessary one may consider fitting a higher-order polynomial that is capable of capturing the "hump" feature reasonably accurate. Alternatively one may consider fitting piecewise either with linear or non-linear functions. Figure 10 shows a piecewise linear implementation of flow rate estimation from the distillation column set-up.

CONCLUSION

The event-based, discretized-state flow control system consists of a volume transporter, a cylindrical vertical glass with six discrete liquid level sensors, and a computer interface. The difference between inflow and outflow is estimated from the time taken for a change to occur in

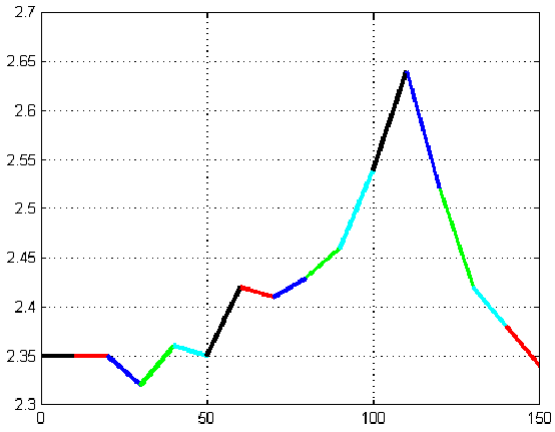


Fig. 10. Piecewise linear regression. x-axis:rpm, y-axis:ratio between flow rate and rpm.

the liquid height in the level glass detected by the discrete level sensors. The control algorithm adjusts the pump speed and thus the pumping ratio accordingly. The fully automated calibration procedure does minimally interfere with the operation of the assembly and can thus be periodically repeated thereby maintaining a high accuracy of the estimated flow rate through regular re-calibration.

The self-calibrating flow rate control structure is low-cost, robust and easily implemented and maintained. The realisation of the automaton algorithm is extremely simple and consists of a recursive two-table look-up procedure. The first table provides the next state given the current state and current event whilst the second gives the set of actions given the current state and the current event. It thus requires only memory for the automata tables and the few lines of code. The assembly can be used to manually or automatically calibrate both on and off-line. Its use has been demonstrated on a set of lab-scale distillation columns, where it acts as reflux measurement and pump system.

REFERENCES

P P H H Philips. *Modelling, Control and Fault Detection of Discretely-Observed Systems*. PhD thesis, TU Eindhoven, Eindhoven, The Netherlands, 2001.

P P H H Philips, W P M H Heemels, H A Preisig, and P P J Bosch van den. Control of quantised systems based on discrete-event models. *International Journal of Control*, 76(3):277–294, 2003.

H A Preisig. A mathematical approach to discrete- event dynamic modelling of hybrid systems. *Computers & Chemical Engineering*, 20:S1301– S1306, 1996.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Process Monitoring of Three Tank System

H. Aaltonen*

*Systems Engineering Laboratory, Department of Process and Environmental Engineering, P.O. Box 4300, 90014 University of Oulu, FINLAND (Tel: +358 8 5532442; e-mail: harri.aaltonen@oulu.fi).

Abstract: The purpose of this study is to develop the monitoring system for the level and temperature of a mixing tank system in closed loop. Principal component analysis (PCA) has been used to process monitoring where idea is to convert on-line data collected from the process into a few meaningful measures, and thereby assist the operators in determining the status of the operations. Once the PCA model of a data set has been obtained, confidence limits can be established by plotting the time series of the Q and T^2 statistics. The confidence limits were implemented in the DeltaV automation system by applying basic calculation blocks. Several process experiments were conducted in the pilot-scale when the performance of step responses and bypass output flow was compared. The results give the operator as a tool for process condition monitoring.

Keywords: Monitoring, principal component analysis, statistical analysis, automation system.

1. INTRODUCTION

This paper is a study of process monitoring system for the level and temperature of three tank system. A three tank system is popular experimental systems in control laboratories and we have used this system initially for the laboratory exercise, which is called “The measurement and control of conductivity”. Salt is dissolved in small tank from which the solution is pumped to the three tank system. The conductivity of water is measured in the second tank, and the conductivity is controlled by changing the amount of the solution.

The aim of the research is to monitoring the level and temperature of a mixing tank system, which are controlled in the first tank at the same time. The original process data is projected onto smaller number of principal component (or latent variables), thus reducing the dimension of the variables. Once the PCA model of a data set has been obtained, confidence limits can be established by plotting the time series of the Q and T^2 statistics (Chiang et al., 2001). The confidence limits are implemented in the DeltaV automation system.

Structure of the paper is as follows. Automation system is presented in the section 2. The process and PCA method are presented in the sections 3 and 4. Experiments and discussion is shown in the section 5. Conclusions is presented in the section 6.

2. AUTOMATION SYSTEM

The DeltaV automation system is developed by Emerson Process Management. Different processes can be planned and managed by using the system which is based on PC technique. The system uses several standard techniques such as Ethernet network connection, digital communication bus such as Foundation fieldbus, and data integration options such as OPC and XML. (Emerson Process Management, 2009); (DeltaV Books Online, 2010)

The DeltaV automation system contains software and hardware for process control. The system consists of a work station, a control unit, I/O units, a power supply and different system busses. A safety instrumented system is also possible to link to the automation system. A small DeltaV system can contain one ProfessionalPLUS workstation and one controller. An Ethernet switch is used to connect a single ProfessionalPLUS and controller. The control unit manages communication between I/O units. Analog I/O units, digital I/O units and the Fieldbus Foundation fieldbus enable effective data transfer. Furthermore, data can be collected from historian database by using MS Excel program. (Emerson Process Management, 2009); (DeltaV Books Online, 2010)

The system contains many software tools for the design and maintenance of the automations system. DeltaV software enables to design graphical user interfaces, to construct control modules, and to configure them, and to collect data. The most important design tools are DeltaV Explorer, Control Studio, and DeltaV Operate Configure. The control program can be made by applying the DeltaV Explorer and Control Studio programs. By using DeltaV Explorer, the structure of the system can be investigated. The components of system can define (areas, modules, equipment and alarms) by using the DeltaV Explorer. Control modules are designed and modified by using the Control Studio. The modules are graphically composed by using ready-made components, which are available in the programming libraries. Graphical user interfaces are designed by using the DeltaV Operate

Configure program which is used also during operation. The state of process can be observed by using the Process History View program. (Emerson Process Management, 2009); (DeltaV Books Online, 2010)

DeltaV InSight is a suite of tools that support a systematic approach to improving control by monitoring control performance, identifying and diagnosing problem loops, recommending tuning and maintenance improvements, and continuously adapting to changing conditions to optimize plant performance. The InSight applications enable you to examine abnormal operating conditions that have been detected by I/O and Control blocks and access the recommended tuning and associated process models that are developed on demand or automatically during normal plant operation. The time-critical process identification and adjustment of tuning functions are performed through the use of block modifiers that execute with the PID function block and operate independently of the InSight application. InSight tuning allows you to quickly tune a loop with minimal input. You can tune with InSight if you have Tuning and Control keys assigned to your user account. If this is your first time tuning with InSight, you might want to use simulator modules for PID and FLC blocks before tuning a live process. The principle of tuning is based on the relay feedback. The program sends pulses, which produce oscillation with constant amplitude and frequency in the process and the parameters of the process can be defined. (DeltaV Books Online, 2009)

3. PROCESS DESCRIPTION

The three tank system for control studies was built to allow student and researchers to test control strategies on a system that is as close as possible to an actual industrial plant. The pilot plant is operated with the DeltaV automation system. The field devices are connected mainly through Foundation Fieldbus (FF). In addition to the Operate view, Process History view is used to follow the dynamic behaviour of the process. The pilot plant used in the lab work consists of three tanks in series, a NaCl-solution tank, a membrane pump that pumps salt water into water tank 2, a conductivity sensor in the tank 2 that measures also temperature, flow measurements and several control valves that control the tank levels and water temperature. The temperature is measured in the first tank with Rosemount Transmitter Model 3244MV (FF) in measuring range of 0-100 °C, and the level with Rosemount Transmitter Model 3051 (FF) in measuring range 0-950 mmH_2O. The plant has five control loops: one warm water control loop, one cold water control loop, two level control loops and one conductivity control loop. The laboratory system is shown in Fig. 1.

Fig. 1. The three tank system.

4. PCA method

The basic idea of principal component analysis (PCA) is to reduce the dimensionality of a set considering a large number of interrelated variables, while retaining as much as possible of a variation present in the data set. This is achieved by transforming the measured data to a new set of variables, the principal components, which are uncorrelated. These principal components are ordered so that the first few retain most of the variation present in all of the original variables. PCA relies on an eigenvector decomposition of the covariance or correlation matrix of the process variables. (Wise and Callagher, 1996); (Chiang et al., 2001)

In this work I have collect data matrix $\mathbf{X}$ with two input and two output variables. For a given data matrix $\mathbf{X}$ with m rows and n columns the covariance matrix of $\mathbf{X}$ is defined as

$$\mathrm{cov}(\mathbf{X}) = \mathbf{X}^{\mathrm{T}}\mathbf{X}/(m-1) \tag{1}$$

The columns of $\mathbf{X}$ are autoscaled, i.e. adjusted to zero mean and unit variance by dividing each columns by its standard deviation (Wise and Callagher, 1996). Principal component analysis is a method of expressing a matrix $\mathbf{X}$ of input variables as outer products of two vectors, a score matrix $\mathbf{T}$ and a loading matrix $\mathbf{P}$ plus a residual matrix $\mathbf{E}$.

$$\mathbf{X} = \mathbf{TP} + \mathbf{E} \tag{2}$$

In PCA the eigenvalues λ_i associated with the eigenvectors p_i as (Wise and Callagher, 1996):

$$\mathrm{cov}(\mathbf{X})\,p_i = \lambda_i\,p_i \tag{3}$$

where p_i vectors are eigenvectors of the covariance matrix.

In PCA method the first eigenvector or principal component aligns with the greatest variation in the data, and the second principal component aligns with the greatest amount of variation that is orthogonal to the first principal component (Wise and Callagher, 1996).

Usually less than the entire principal components are used. It can calculate by cumulative sum of variances and select the principal component as over 72 percent of the total variance (Chiang et al., 2001).

The PCA model can be used in monitoring by plotting of Q and Hotelling's T^2 statistics.

The Q statistic indicates how well each sample conforms to the PCA model. Q statistic is a sum of squares of each row (sample) of $\mathbf{E}$ (Wise and Callagher, 1996):

$$Q = e_i e_i^T = x_i\left(\mathbf{I} - \mathbf{P}_k\mathbf{P}_k^T\right)x_i^T \tag{4}$$

where e_i is the ith row of $\mathbf{E}$, $\mathbf{P}_k$ is the matrix of the first k loadings vectors retained in the PCA model and $\mathbf{I}$ is the identify matrix.

Hotelling's T^2 statistic measures the variation in each sample within the PCA model. T^2 is the sum of normalized squared scores as (Wise and Callagher, 1996):

$$T^2 = t_i\boldsymbol{\lambda}^{-1}t_i^T = x_i\mathbf{P}\boldsymbol{\lambda}^{-1}\mathbf{P}^T x_i^T \tag{5}$$

where t_i is ith row of $\mathbf{T}_k$, the matrix of k scores vectors from the PCA model and $\boldsymbol{\lambda}^{-1}$ is the diagonal matrix containing inversed eigenvalues with the k eigenvectors retained in the model.

Once a PCA model of a data set has been obtained, confidence limits can be established for the overall residual Q and T^2 index. Given the eigenvalues λ_i of the covariance matrix of $\mathbf{X}$ confidence limits can be calculated for the Q as (Chiang et al., 2001):

$$Q_\alpha = \theta_1\left[\left(h_0 c_\alpha\sqrt{2\theta_2}\right)/\theta_1 + 1 + \left(\theta_2 h_0(h_0-1)\right)/\theta_1^2\right]^{1/h_0} \tag{6}$$

where $\theta_i = \sum_{j=k+1}^{n}\lambda_j{}^i$, and $h_0 = 1-\left(2\theta_1\theta_3\right)/3\theta_2{}^2$, and c_α is the standard normal deviate corresponding to the $(1-\alpha)$ percentile. Given the level of significance, α, the threshold for the Q statistic can be computed using (6) and be used to detect faults.

Statistical confidence limits for T^2 can be calculated by means of the F-distribution as follows (Wise and Gallagher, 1996):

$$T_\alpha{}^2 = \frac{k(m-1)}{m-k}F(k,m-1,\alpha) \tag{7}$$

where $F(k,m-1,\alpha)$ is the upper $100\,\alpha$ % critical point of the F-distribution with k and $m-1$ degrees of freedom.

4. EXPERIMENTS AND DISCUSSION

Monitoring system for the level and temperature of a mixing tank system were investigated, which are controlled in the first tank by separately PI controllers. Step response and disturbance with output flow of the first tank are conducted to define the behaviour of the process in the entire area of operation.

The process experiment, when level was changed, is presented in Fig. 2. The steady state level temperature was 22 degree °C, and the level was 270 mmH_2O. The set point of the level controller was changed from 270 to 280 mmH_2O, and then to 250 mmH_2O. Controllers gain and integration time were 4.60 and 1.25, and 50 and 150 seconds for the level and temperature loops.

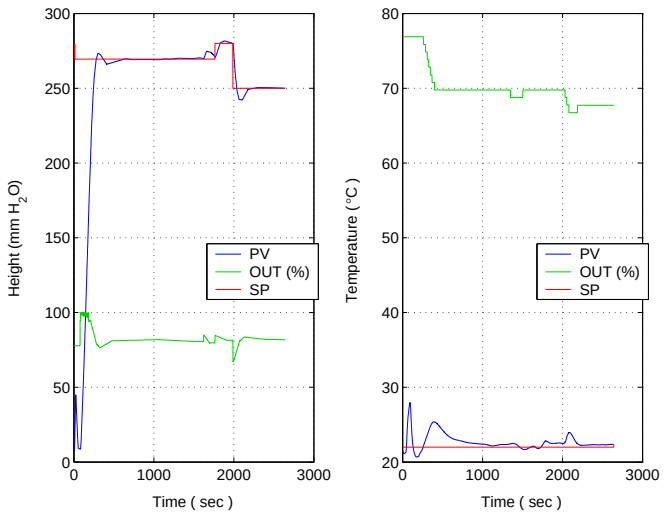


Fig. 2. Level and temperature control in the first tank using PI controllers.

PCA model was done using normal operation data, and it is selected a set of 300 data sample rows measured a 1 second sampling time in Fig. 2. Cumulative sum of the variances by each principal component is illustrated in Table 1. It shows that 90% of the total variance is captured by two principal components. Thus the four variables can be replaced with two new variables, which are linear combinations of the original variables, with little loss of information.

Table 1. Variance captured by PCA model.

Principal component number	Variance	Total variance captured
1	0.5684	0.5684
2	0.3338	0.9023
3	0.3429	0.9880
4	0.0481	1

Figure 3 plots the test data for the level step responses and bypass flow at the first tank; Fig. 4 shows the same for the process monitoring with Q and T^2 index. Confidence limit is 95% for the Q and T^2 indexes. The new steady state temperature was 22.1 degree °C, and the level was 270 mmH_2O for the DeltaV automation system.

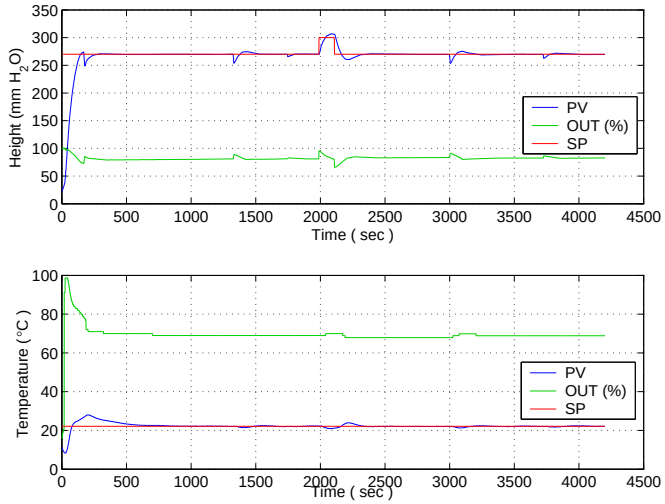


Fig. 3. Test data for the process monitoring in the first tank.

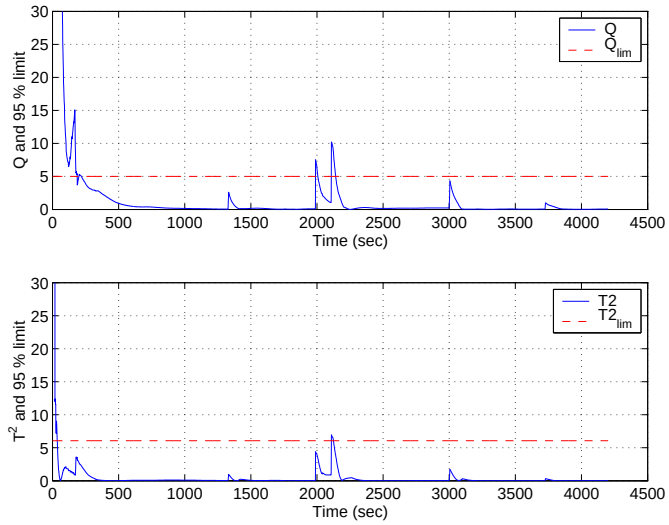


Fig. 4. Graphical illustration for fault detection using the Q and T^2 statistics.

The results shows peak upward (see Fig. 4 at 1320, 1740, 3000 and 3740 seconds) when the bypass flow was opened five, one, five and two seconds. Peak is also upward (see Fig. 4 at 1990 to 2110 seconds) when the set point of the level controller was changed from 270 to 300 mmH_2O, and then to 270 mmH_2O. Q and T^2 indexes were over 50 and over 20 times the 95 % limit when the step point of level controller was changed. There are no peak upward (see Fig. 4 at 3950 to 4050) when the set point of level was changed 940 to 930 mmH_2O, and then to 940 mmH_2O in the second tank. The diagnosis system is implemented in the DeltaV automation system.

5. CONCLUSIONS

In this study, it was used the three tank system that was built to allow students and researchers to study control strategies. Process monitoring system for the level and temperature of a mixing tank system were investigated, which are controlled in the first tank at the same time. Principal component analysis (PCA) has been used to process monitoring where idea is to convert on-line data collected from the process into a few meaningful measures, and thereby assist the operators in determining the status of the operations. Once the PCA model of a data set has been obtained, confidence limits can be established by plotting the time series of the Q and T^2 statistics. The diagnosis system is implemented in the DeltaV automation system.

REFERENCES

Chiang, L.H., Russell, E.L., and Braatz, R.D. (2001). *Fault detection and diagnosis in industrial systems.* Springer, London.

DeltaV Books Online (2010). *Emerson process management.*

Emerson Process Management (2009). *DeltaV digital automation system, system overview.* read 9 December 2011. URL: http://www2.emersonprocess.com/siteadmincenter/PM DeltaV Documents/Brochures/DeltaV-System-Overview-v11-Brochure.pdf

Wise, B.M. and Callagher, N.B. (1996).The process chemometrics approach to process monitoring and fault detection. *Journal of Process Control*, 6 (6), 329-448.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Model-based analysis of control performance in sewer systems

A. H. Mollerup*, M. Mauricio-Iglesias, N. B. Johansen***, D. Thornberg***, P. S. Mikkelsen**** and G. Sin****

** CAPEC, DTU Chemical Engineering, Technical University of Denmark, DK-2800 Lyngby, Denmark and Copenhagen Wastewater Innovation, DK-1432, Copenhagen K, Denmark (e-mail: molle@ kt.dtu.dk or ahm@ke.dk).*
*** CAPEC, DTU Chemical Engineering, Technical University of Denmark, DK-2800 Lyngby, Denmark, (e-mail: mmmi@ kt.dtu.dk; gsi@kt.dtu.dk).*
****Copenhagen Wastewater Innovation, DK-1432, Copenhagen K, Denmark (e-mail: nbj@ ke.dk; dt@spvand.dk).*
*****DTU Environment, Technical University of Denmark, DK-2800 Lyngby, Denmark, (e-mail: psm@ env.dtu.dk).*

Abstract: Design and assessment of control in wastewater systems has to be tackled at all levels, including supervisory and regulatory level. We present here an integrated approach to assessment of control in sewer systems based on modelling and the use of process control tools to assess the controllability of the process. A case study of a subcatchment area in Copenhagen (Denmark) is used to illustrate the combined approach in modelling of the system and control assessment.

Keywords: Sewer system, control, plantwide control, system understanding, modelling

1. INTRODUCTION

Since the EU Water Framework directive came into force in 2000, wastewater systems (sewer system and wastewater treatment plants) in Europe have been put under pressure to reduce the number of combined sewer overflows[1] (CSOs) from the system to protect the aquatic environment. And as the future climate changes are predicted to induce an increase in precipitation in the northern part of Europe (Watson et al., 1997), the strain on the performance of the wastewater systems will only become larger in the future.

To cope with the increasing pressure the wastewater system can be expanded by building larger pipes and new storage tanks. But expanding the wastewater system requires a significant capital investment and intensive civil building works and therefore the need for this should be limited as much as possible. Instead of expanding the system, research has shown that the implementation of Real Time Control (RTC) in the sewer system can increase the utilization of the existing storage volume and thereby reduce the need to build new storage capacity (Marinaki, Papageorgiou 1997). The use of RTC allows for the control of pumping stations and diversion or retention gates according to online measurements at critical points, thereby making it possible for the system operation to respond to rain events.

To further increase the utilization of the system, research is now primarily focused on how to do a system wide optimization. In particular, research on how to use real time optimisation (RTO) and model predictive control (MPC) in the control structure.

The problem of finding the optimum control structure that best serves both the sewer system and wastewater treatment plant(s) is a challenging and formidable problem. The main objective of the control system is to minimize the risk of flooding and the volume of overflow and bypass from the system.

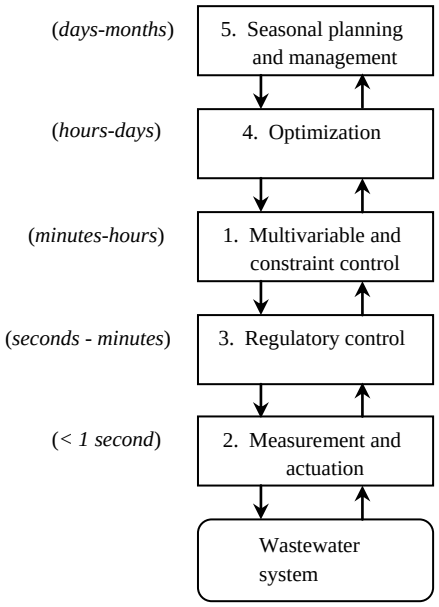


Fig. 1: The system decomposition with respect to time scale (based on Seborg et al. (2011)).

One way to address the control problem is to decomposition it according to the time scales of the actions. From this it becomes possible to identify a number of layers, linked by

[1] Combined Sewer Overflow: The discharge from a combined sewer system that handles both wastewater and stormwater. During rain the system can discharge from overflow structures into a recipient such as a stream, river, lake or sea, if the capacity of the system is exceeded.

master-slave relations. For the sewer system the decomposition can look as depicted in figure 1.

Based on decomposition one can look at the different levels of control separately, starting from the lowest level and moving upward.

Using the system decomposition it becomes obvious that the research today on control of the wastewater system is focused on higher layers in the system decomposition. However these methods are supervisory layer controllers that require a control structure as given. Therefore it becomes imperative that the regulatory level control functions optimally.

However, in practice the design and implementation of RTCs in wastewater systems have been done incrementally as the utility management companies have identified the potentials. Over time as more and more controls have been implemented in the wastewater system and the number of controlled actuators in the system has increased, important interactions among the different control loops may have appeared, that was not accounted for in the design of the individual controls. Depending on the interactions among the control loops this may appear impairing on the performance of the control system. Therefore attention should be paid to aspects related to the regulatory level such as the analysis of controllability properties of the system according to the design of control loops, pairing of variables, among others. This analysis of control problem and designing of control structure needs to be performed in a formal and structured way as required by process control engineering good practice (e.g. see (Skogestad 2002).

The aim of this work is therefore to formulate the problem of design and analysis of the regulatory level from a process control perspective and to develop a methodological approach to find the optimal solution.

2. METHODOLOGY

The methodology is developed assuming a hierarchical vertical problem decomposition based on time scale as it has been done with similar methodologies aimed at chemical processes. Other approaches in control design for whole plants are i) horizontal decomposition based on process units, ii) vertical decomposition based on process structure; and iii) vertical decomposition based on control objectives (Larsson, Skogestad 2000). These approaches are considered unsuitable for the following reasons. The horizontal decomposition approaches are unable to tackle material recycle (e.g. pumping water back to a previous tank) and assumes large buffer tanks between units (Douglas 1988). The process structure decomposition consists on splitting the process at different levels of representation (e.g. batch vs continuous, input-output) but it has been formulated for chemical plants and it is difficult to adapt to other processes. Finally, the control objective approach establishes different tasks for control (inventory, product specification, equipment constraints, economic performance) but fails to consider the cases where such tasks can lead to contradictory objectives (Price, Lyman & Georgakis 1994).

2.1 Generic approach to the control assessment problem

A suitable solution to the problem of control assessment should not only point out the deficiencies in the system but also address potential solutions to solve them. However, as stated previously, the approach cannot be that of pure design given that the capital cost of modifications in the control system of sewer system is very high and involves works in the public domain. The assessment should lead to incremental modifications that, albeit suboptimal, are feasible.

We propose here to carry out a decomposition of the system with respect to time in different layers linked by master-slave relations (Fig. 1). Such decomposition has been applied in chemical processes as a way to manage the complexity of control design. The methodology consists on the review of each of the control layers, their assessment and proposal of solutions and an eventual evaluation step.

To this respect, the classic indexes will be tested (relative gain array, closed-loop disturbance gain) as indicators of process controllability and ability to reject disturbances. The adaptation of controllability indexes to describe sewer systems will be investigated and new indexes will be proposed if appropriate.

3. CASE STUDY

The approach proposed for the control system assessment is illustrated here through an actual case study described below. The analysis was done considering three scenarios, namely:

1.Dry weather

2.Low intensity rain

3.Moderate intensity rain

Since the primary objective of the control system is to minimise the overflows to the environment, the scenario 3 (moderate intensity rain) is arguably the most relevant. No overflow is expected during dry weather or low intensity rain. Nevertheless, the energy consumption of the pumps is mainly related to the operation during scenarios 1 and 2 since they represent the normal operation. Therefore, they should also be investigated in order to achieve an energy efficient operation as well.

The evaluation of the system is performed with simulation of a long time period (in this work 30 years), using as an input historical rainfall data. The key performance indicators (KPI) are the number of CSO, the volume of the overflows, their frequency and the flooding episodes (overflows from virtual tanks).

3.1 Description of case study area.

The subcatchment area analysed in this work is a part of Copenhagen (Denmark) sewer system (Fig. 2), owned and maintained by Copenhagen Energy. It has a size of Y hectare, X km pipes and is additionally composed of 3 pumping stations, 2 storage tanks and 1 pipe basin. The inflows to the system are the wastewater dry weather flow and rainfall and the outflows are five flows to the environment (caused by

overflow) and a pipe that directs the flow downstream for treatment.

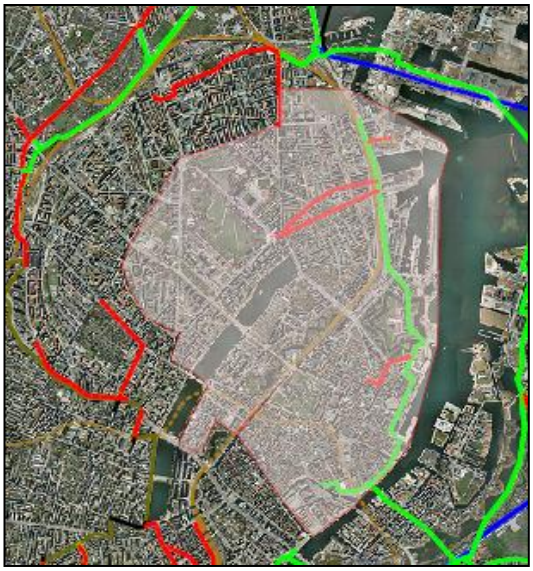


Fig. 2. Map of case study area.

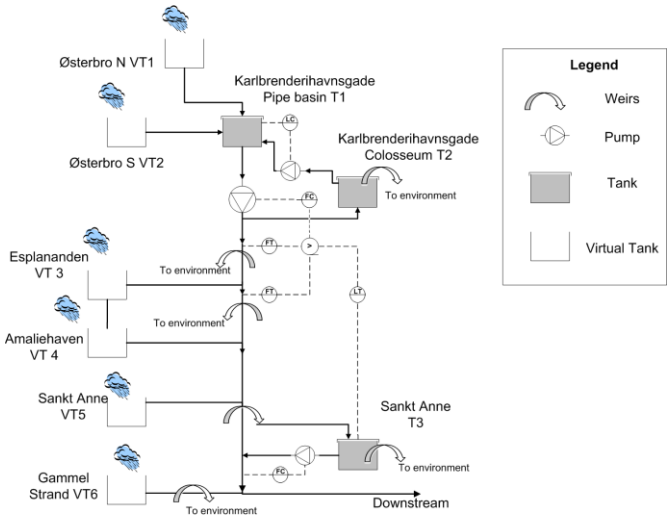


Fig. 3. Modelling the case study using virtual tank approach

The implemented control system of the area is composed of three decentralized loops as follows:

-Loop 1.Controls the level in the tank (T1) manipulating the outflow from the tank (T2), with a larger storage capacity.

-Loop 2. Selective control which manipulates the outflow from tank (T1) to the variable which is further from the setpoint among i) the flow measured downstream the pumping station, ii) the flow in the pipeline after the virtual tank (VT 3) and iii) the level of tank (T3)

-Loop 3. Manipulates the outflow of tank (T3) according to the flow downstream the pumping station.

3.2 Sewer system modelling

The catchment area and sewer system were modelled through the virtual tank approach (Ocampo-Martinez 2010). Virtual tanks are regions of the subcatchment area where the precipitation and flow are considered to be homogeneous. The global mass balance for a virtual tank is expressed as:

$$\frac{dV_i}{dt} = q_{in} + P_{eff} - q_{out} \quad (1)$$

where q_{in} is the inflow coming from other tanks, virtual tanks and the dry weather flow (household wastewater), I_{eff} is the effective precipitation and q_{out} is the outflow from the virtual tank, empirically modelled as:

$$q_{out} = \beta_i V_i \quad (2)$$

The parameter β_i (in s^{-1}) is a the volume/flow conversion coefficient (Ocampo-Martinez 2010). It can be determined from regression of historical data of flow and level. If the regression was not satisfactory for all the ranges of level/flow, β_i can be determined piecewise for two or more ranges at the expense of introducing a nonlinearity in (2). Indicators of wastewater composition have also been included and are modelled considered the virtual tanks as completely mixed:

$$\frac{dV_i C_j^i}{dt} = q_{in} C_j^{in} - q_{out} C_j^i \quad (3)$$

where the index i represents the virtual tank and j the compound considered. The compounds considered are the chemical oxygen demand (COD), total nitrogen compounds (TN), total phosphorous compounds (TP), and total suspended solids (TSS). It is assumed that precipitation has negligible concentration of these compounds and, consequently, does not appear in (3).

The tanks in the system have been modelled as completely mixed compartments and their outflow is assumed to be perfectly controlled by the corresponding control loops (level versus pumps).

Weirs are elements that are used to moderate the flow downstream (e.g. before a pipe node) and to prevent the backwater effects. Since weirs do not have storage capacity, they are modelled as:

$$q_{out} = \begin{cases} q_{in} & if\ q_{in} \le q_{max} \\ q_{max} & otherwise \end{cases} \quad (4)$$

Finally, pumps are modelled as perfect actuators, immediately setting the flow to the desired value.

3.2.1 Overflow modelling

The modelling of overflow is a key step, given that the objective of sewers system control is precisely to minimise the overflow to the environment (river, lakes, sea…). However, it should be noted that not all the overflows streams are spilled onto the environment; a number of them

are directed to other tanks or virtual tanks according to the links in Fig. 3. Overflow has been modelled as follows:

-From virtual tanks and tanks. The subcatchment areas are divided into smaller zones for which a maximum level (and therefore a maximum volume) is assumed before flooding. For tanks, a maximum storage capacity is known. Overflow is formally calculated the same way for the two structures as:

$$q_{overflow} = \begin{cases} 0 & if\ V \leq V_{max} \\ q_{in} - q_{out} & otherwise \end{cases} \quad (5)$$

-From weirs.

$$q_{overflow} = \begin{cases} 0 & if\ q_{in} \leq q_{max} \\ q_{in} - q_{max} & otherwise \end{cases} \quad (6$$

3.2.2 Dry weather flow modelling

Dry weather flow is the flow in the sewers system not influenced by a rain event. For an urban sewer system, it corresponds to the wastewater produced by households in the area covered by a virtual tank. Its contribution to the total flow during any rain event is in general negligible. However, it is important to take it into account since it is the source of polluting agents that can be spilled in case of overflow.

Dry weather flow has been modelled through as a four term harmonic function of periods 24h/12h (Breinholt et al. In Press) or by regression to a 3rd degree polynomial, repeated as an input with period 24h (Sin et al. 2008). In this work, the dry weather flow was simulated by a 24 piecewise function (each piece consisting of a straight line interpolating a 1h period) built from actual operation data (Fig. 4)

The COD, TN, TP and TS concentration of the dry weather flow was determined taking into account the amount of equivalent-persons per area covered by each virtual tank (Tchobanoglous, Burton 1991).

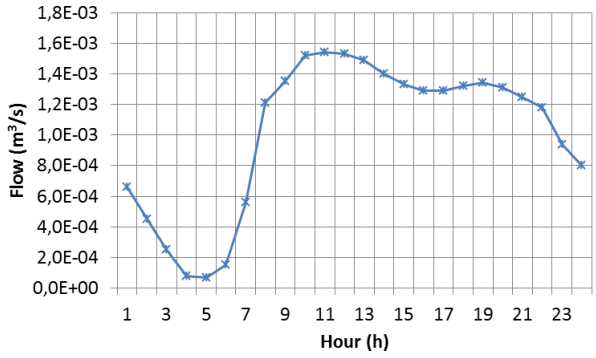


Fig. 4. Dry weather flow hourly variation

3.2.3 Precipitation modelling

To carry out the three-scenario analysis, box rain was used in order to reach a suitable steady state. As for the evaluation of the control base case and proposed improvements, historical rain data were used corresponding to the last 30 years in the subcatchment area. Other approaches in evaluation reported in literature are based on the use of designed rain events with a certain return period, in particular the so-called Chicago designed storm (Huff, Vogel & Changnon Jr 1981). However, it has been reported that, since designed rain events can overestimate or underestimate the actual rainfall of different region, sewer systems evaluations should be done with historical data or with rain events designed in purpose for a region (Raso, Malgrat & Castillo 1995).

4. CONCLUSIONS

An assessment of sewer system control was carried out by a model-based approach. It was illustrated with a case-study corresponding to a subcatchment area in Copenhagen (Denmark). Future work on this project will consist on the evaluation of the control system with historical precipitation data.

REFERENCES

Breinholt, A., Thordarson, F.O., Møller, J.K., Grum, M., Mikkelsen, P.S. & Madsen, H. In Press, "Grey-box modelling of flow in sewer systems with state-dependent diffusion", *Environmetrics*, .

Douglas, J.M. 1988, *Conceptual design of chemical processes*, McGraw-Hil.

Huff, F.A., Vogel, J.L. & Changnon Jr, S.A. 1981, "Real-time rainfall monitoring prediction system (Chicago).", *Journal of the Water Resources Planning and Management, Division, ASCE*, vol. 107, no. WR2 Proc Paper, 16557, pp. 435.

Larsson, T. & Skogestad, S. 2000, "Plantwide control - a review and a new design procedure", *Modeling, Identification and Control*, vol. 21, no. 4, pp. 209-240.

Marinaki, M. & Papageorgiou, M. 1997, "Central flow control in sewer networks", *Journal of Water Resources Planning and Management*, vol. 123, no. 5, pp. 274-283.

Ocampo-Martinez, C. 2010, *Model Predictive Control of Wastewater Systems* 1st edn, Springer.

Price, R.M., Lyman, P.R. & Georgakis, C. 1994, "Throughtput manipulation in plantwide control structures", *Industrial Engineering Chemistry Research*, , no. 5, pp. 1197.

Raso, P., Malgrat, P. & Castillo, F. 1995, "Improvement in the selection of design storms for the new master drainage plan of Barcelona", *Water Science and Technology*, vol. 32, no. 1, pp. 217-224.

Seborg D., Edgar T., Mellichamp d. & Doyle F. 2011, third edition. *Process Dynamics and Control*, Wiley.

Sin, G., De Pauw, D.J.W., Weijers, S. & Vanrolleghem, P.A. 2008, "An efficient approach to automate the manual trial and error calibration of activated sludge models", *Biotechnology and bioengineering*, vol. 100, no. 3, pp. 516-528.

Skogestad, S. 2002, "Plantwide control: Towards a systematic procedure", *Computer Aided Chemical Engineering*, vol. 10, no. C, pp. 57-69.

Tchobanoglous, G. & Burton, F. 1991, *Metcalf and Eddy Wastewater Engineering* McGraw-Hill.

Watson R., Zinyowera M. and Moss R. 1997. *IPCC Special Report on The Regional Impacts of Climate Change. An Assessment of Vulnerabilit*y. Cambridge University Press.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Dynamic Simulation of Oxy Combustion in a pilot scale CFB boiler

Matias Hultgren*. Laura Lohiniva*
Jenö Kovács. Jouni Ritvanen*****
Antti Tourunen****

** University of Oulu, Systems Engineering Laboratory, POB 4300, 90014 Oulu, Finland. firstname.lastname@oulu.fi.*
*** Foster Wheeler Energia Oy, Varkaus, Finland. jeno.kovacs@fwfin.fwc.com*
**** Lappeenranta University of Technology, LUT Energy, Lappeenranta, Finland. jouni.ritvanen@lut.fi.*
***** VTT Technical Research Centre of Finland, Jyväskylä, Finland. antti.tourunen@vtt.fi.*

Abstract: In the present work, the results of the oxy combustion circulating fluidized bed (CFB) test week of VTT Jyväskylä were analyzed. During the test week, dynamic combustion measurements with a 50-100 kW anthracite/petcoke-fired CFB pilot unit were carried out, including step tests in both air and oxy mode, oxy combustion load ramps with different strategies, as well as switches between air and oxy combustion. The principal aim of the testing was to study the air- and oxy-firing combustion process dynamics and to define the main differences between these combustion modes. These findings can later be used to develop and validate combustion controls for the Foster Wheeler Flexi-Burn^{TM} CFB technology. From the experimental tests, the furnace temperatures, heat transfer probe measurements, flue gas and oxidant oxygen contents, as well as flue gas CO_2 and H_2O percentages were examined and the phenomena behind the responses were discussed.

The outcomes of the pilot CFB test week were also modelled with a dynamic 1-D CFB hotloop model, which contains the furnace, the gas-solid separator (cyclone) and the solids return system. An oxidant gas mixing block for mixing pure O_2, recirculated flue gas (RFG) and air during oxy combustion and air-oxy-air switches is also included. The transient simulation model can be used for both dynamic and steady-state simulations with one or multiple fuel flows. The model is based on ideally mixed 1-D elements, for which mass and energy balance equations are solved against time. A combined energy equation for the gas and solid phases is used to solve the element temperatures. Hydrodynamics, combustion characteristics and heat transfer inside the modules are calculated using semi-empirical correlations. The goals of the simulation testing were to validate the model for oxy combustion and to compare air and oxy combustion. The simulations were conducted with the measured pure O_2, RFG and air flows as direct model inputs. A single fuel flow was used to represent the experimental anthracite and petcoke mass flows, which were calculated from the respective silo weights with least squares fits. 24-hour anthracite and petcoke ratio averages were used for the input fuel flow property calculations. Slight modifications to the fuel mass flow were also made to better match the flue gas O_2 of the experimental results. The differences in the simulated and measured flue gas O_2 were possibly caused by inaccuracies in the experimental fuel mass flows or by air leakage. With the fuel corrections, the simulation model showed satisfactory results for the dynamic tests and the model was thus successfully validated for oxy combustion. The simulation model is an important tool for designing and assessing control solutions for the oxy-CFB hotloop. Process simulation is also important for oxy-firing, because the separate control of the pure O_2 and RFG flows increases the number of possible control solutions in the solid fuel boiler.

This work is a part of the Foster Wheeler circulating fluidized bed and oxy combustion research. This investigation was co-financed under the EEPR09-CCS-COMPOSTILLA Project (European Union's "European Energy Programme for Recovery" programme). Oxy combustion is a natural continuation to the solid fuel combustion development to meet the challenges of climate change. Therefore, oxy-firing related control issues have to be investigated before the application of this emerging technology to power plants. The measurement experiments that are processed and simulated in this work are described by Mikkonen et al. in the study "Experimental Testing of Oxy Combustion in a pilot scale CFB boiler", which has been submitted to the 17th Nordic Process Control Workshop.

Keywords: oxy combustion, CFB, dynamic models, dynamic behaviour, furnace simulation.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Experimental Testing of Oxy Combustion in a Pilot Scale CFB boiler

Hannu Mikkonen*. Laura Lohiniva. Antti Tourunen*.**
Mikko Jegoroff*. Matias Hultgren. Jenö Kovács***.**

** VTT Technical Research Centre of Finland, Jyväskylä, Finland. firstname.lastname@vtt.fi*
*** Systems Engineering Laboratory, University of Oulu, Oulu, Finland. firstname.lastname@oulu.fi*
**** Foster Wheeler Energia Oy, Varkaus, Finland. firstname.lastname@fwfin.fwc.com*

Abstract: In the present work, measurement experience and results of oxy combustion in a pilot scale (20-100 kW) coal-fired circulating fluidized bed (CFB) are summarized. The pilot system at VTT Jyväskylä enables operation in both air and oxy combustion modes. The aim of the one-week test runs was to study combustion dynamics under both conditions with the aim of developing and validating combustion controls for Foster Wheeler's FlexiBurn^{TM} CFB combustion technology. The technology is currently being applied in a 30 MWth test unit close to Compostilla power plant in Spain, where also a 300 MW^{e} complete CCS (carbon capture and storage) demonstration plant is planned to be built in 2013-2015. Dynamic combustion tests in this study included various load step and ramp change tests in both air and oxy modes, operation strategy tests in oxy mode and tests for mode switching methods (from air to oxy and vice versa). The measurement results were analyzed, dynamic behavior in air and oxy mode were compared, and the observed and expected phenomena were discussed. The measurement period was then followed by dynamic simulations to evaluate how the results agreed with modeling work and expectations.

The pilot-scale CFB system operates with fuel power of 50-100 kW in oxy mode and 20-50 kW in air mode. The solids separation and recycling system includes a cyclone and loop seal. For operating in oxy combustion mode, a flue gas recirculation system and mixing with O_2 are installed. A coal mix of anthracite and petcoke was used as fuel. Additives (bed sand, limestone for emission control) can be fed into the furnace together with fuel or into the circulating material. The oxidant gas consists of air, bottled O_2, recirculated flue gas (RFG), or other bottled gases such as CO_2, or it can be a mixture of these. All input fuel, air, oxygen and RFG flows can be measured and controlled, with the exception of secondary RFG flow which is determined by the difference between total and primary RFG flows. The air distribution system is used to control total gas flow and feeding at primary and secondary levels. The furnace is also equipped with heating (electrical) and cooling (air and water) systems. The furnace consists of several separately controlled zones, which allows flexible control of process and combustion conditions (e.g. O_2 level, temperature). Pressure and temperature measurements are well located all along the furnace height, and there are several gas and solid material sampling ports. Flue gas composition is monitored with traditional online gas analyzers and an FTIR spectrometer. The pilot system is controlled by a PCS7 automation system, which also collects all measurement data. The pilot system can be operated in both air and oxy firing modes, and different mode switching methods can be used.

During the test week, combustion dynamics were studied under both air and oxy firing conditions. The aim was to develop and validate combustion controls for oxy CFB combustion technology. Dynamic combustion tests included a reference load step series for air combustion, load step series for oxy combustion, load ramp series for oxy combustion and switching strategies between air and oxy modes. Step tests were performed between 70-100% load. All input flows (fuel, oxidant gas) were changed stepwise. Settling time of one hour was allowed between two consecutive steps. In ramp tests, different ramp speeds and ranges were compared. Finally, measurement results from the experimental tests were analyzed. For example furnace temperatures, heat transfer probe measurements, flue gas composition, and O_2 content in both flue gas and oxidant were examined.

The assessment of measurement results versus dynamic CFB model simulations performed within this test study is presented in the paper "Dynamic Simulation of Oxy Combustion in a pilot scale CFB boiler" by Matias Hultgren et al. submitted to the 17th Nordic Process Control Workshop.

This investigation was co-financed under the EEPR09-CCS-COMPOSTILLA Project (European Union's European Energy Programme for Recovery programme).

Keywords: oxy combustion, CFB, dynamic behavior

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Greenhouse LED lighting control

Torsten Wik * **Anna-Maria Carstensen** * **Tessa Pocock** **

* *Department of Signals and Systems, Chalmers University of Technology, Göteborg, Sweden*
** *Heliospectra AB, Göteborg, Sweden*

Abstract: Industrial scale greenhouses have, during the last decade, reached a high level of automation. However, lighting control is in general still controlled manually because of the type of lamps (High Pressure Sodium) that are used. With High Brightness LEDs about to reach the market today sufficiently high power for greenhouse grown crops can be achieved, and this opens up for advanced lighting control. Optimized control will, however, be a difficult task because the needs of the plants differ between individual plants, crops, time of the day, time of the growth cycle, temperature, and of course the natural ambient light. In this approach to this problem we distinguish four different control loops: growth control, ambient light compensation, light stress detection and recovery, and spectrum optimization, where the focus of this work is on the latter two. In particular it is shown here that light induced photoinhibition, decreasing photosynthetic yield and potentially damaging the plants, can be remotely detected in a light environment.

1. INTRODUCTION

Contrary to most people's belief, green house lighting is a major energy consumer in Europe. The current electricity consumption is estimated to be around 150 TWh per year, which is about the same as the total electricity consumption in Sweden. Modern Dutch green houses are built in two storeys, 10ha in size, and with an electricity consumption of 10 MW powered by gas turbines. A lowered electricity consumption would clearly have a significant environmental impact and also allow for crops grown closer to the consumer.

Today, almost all full scale green houses use High Pressure Sodium (HPS) lamps for their lighting. These lamps are more or less of the same type as those used for highways. They are highly efficient in the sense that they give a lot of light for a given power. However, what has not been commonly known is that the spectra they produce do not fit well to the absorption spectrum of the photosynthesis. In fact, the mismatch, with a lot of power in the far red, implies that approximately one third of the emitted light energy can never be used by the plants, and often the wasted light is even higher (see Figure 1). Another problem with the HPS lamps is that they are not adjustable and slow to start, and therefore they are in general not controlled even though most other processes in a commercial green house are.

Today there are high power LEDs available on the market and a company, Heliospectra AB in Göteborg, is developing LED-based lamps for green houses. These lamps contain several different groups of LEDs having different colors to better fit the absorption spectrum of the photosynthesis (Figure 1). In Figure 2 the first commercial installation for one of Santa Marias (Swedeponics) green house for basil is shown.

Since LEDs are easily adjustable in power this opens up for feed forward as well as feedback control. This control

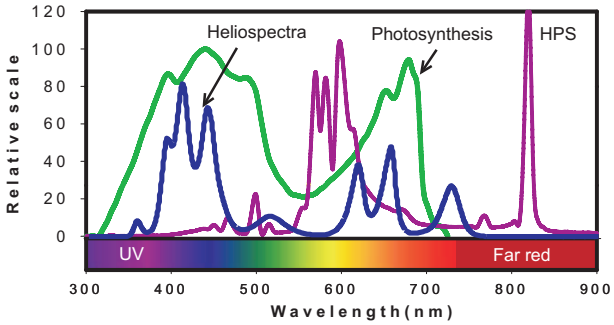


Fig. 1. The action spectrum of the photosynthesis, HPS lamps and the prototype LED-lamp.

Fig. 2. The first commercial installation of the Heliospectra lamp in a basil house at Swedeponic, Påarp, Sweden. (Photo: T. Pocock for Heliospectra).

possibility may cause potentially large energy savings through at least the following mechanisms:

- *Growth control.* Today the growers have poor possibilities to adjust the plant growth, which have the effect that a significant part of the harvest has to be thrown away because the demand does not match the produce. For this reason Swedeponic, for example, throw away 15% of their basil produced. Using variable light intensities we can control the growth rate within certain limits, hence minimizing this waste.

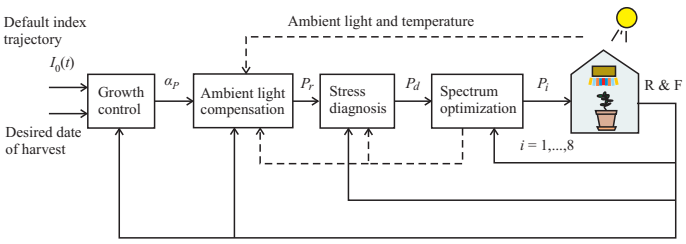


Fig. 3. The basic components of the LED lamp control system being developed.

- *Photoinhibition.* Excess light causes plant stress and the induction of photo protective mechanisms that lower the yield. The plants may even become damaged with a permanent decreased photosynthetic yield. (Note that the human eye cannot see when the light stress begins but only when the plants are actually being damaged!)
- *Light spectrum.* The intensities of the different LED-groups (colors) should be adjusted to the needs of the plant.

2. METHOD AND PRELIMINARY RESULTS

In this project we are aiming at a control system for a LED-lamp adjusting the LED intensities to the needs of the plants and the grower, based on sensors installed in the lamp. The system is characterized according to the above energy saving mechanisms (see Figure 3).

To determine the needs of the plant is a difficult task, but their status affects the light emitted (fluorescence) and the light reflected from the plants. A key research task is therefore how to use the measured emitted and reflected light to diagnose the plants.

2.1 Stress diagnosis

Analysis of fluorescence, using fluorescence indices such as F_v/F_m, from plants is a well established method for detecting plant stress. However, such standard methods require on leaf measures and a completely controlled environment. To be used in practice for automatic control, the stress has to be sensed remotely. Takayama et al. [2011] have remotely detected plant stress (draught) in tomato plants in a greenhouse using fluorescence. However, their method requires dark adaption (complete darkness for at least 20 minutes), and can thus only be used at night and not in daylight, which is the normal situation for our application.

In a series of experiments we have focused on diagnosing the signaling response to different light excitations with the purpose of finding when the plants become stressed by excess light. We have then found that the fluorescent light at 685 nm responded well to an excitation signal at 420 nm. As a first step we have investigated the responses to steps (Figure 4) and to slowly varying sinusoids (Figure 5). From the step responses it can be concluded that the plants exhibit different dynamics for increases in light intensity than for decreases. In particular it was found that the dynamics of the up-steps were significantly changed by excess light already at moderate (reversible) photoinhibition. A

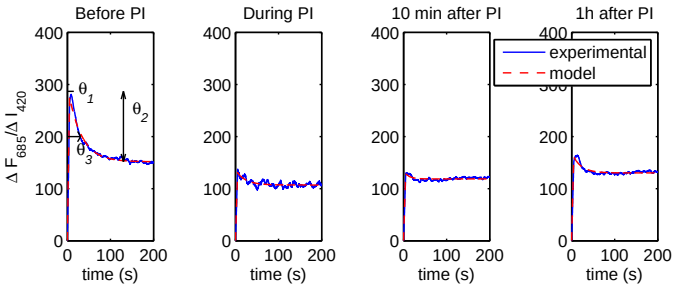


Fig. 4. First order transfer functions with direct term fit closely to the up-steps with clear parameter changes caused by the photo inhibition.

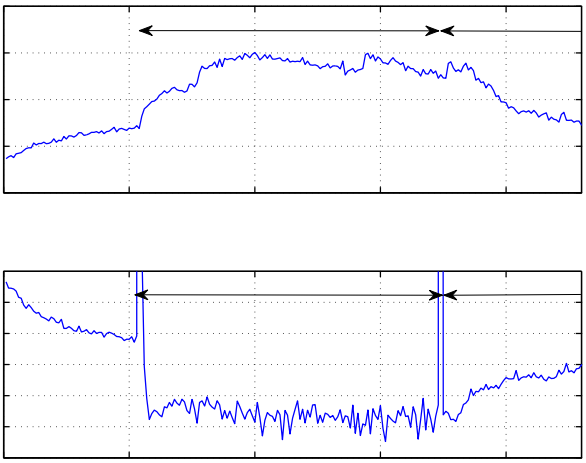


Fig. 5. The ratio between the amplitude of the variation at 60-1 Hz in the reflected light at 420 nm and the amplitude of the variation in the emitted light at 420 nm (top). The corresponding ratio between the variations in fluorescence at 685 nm and of the applied light at 420 nm (bottom).

relatively close fit of the responses to first order transfer functions on the form

$$G(s) = \theta_1 - \frac{\theta_2}{1 + s\theta_3}$$

to the measured responses could be achieved (Figure 4), with clear changes in model parameters (Figure 6), implying that photo inhibition can in fact be remotely detected in a light environment. In the next step, recursive identification of models having direction dependent dynamics (Rosenqvist [2004]) using a superimposed suitable binary excitation signal at 420 nm, for example, will be investigated as an approach to have a continuous online diagnosis.

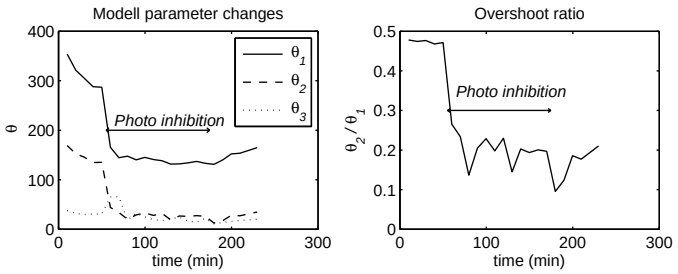


Fig. 6. Changes in model parameters (see Figure 4) during and after photo inhibition.

REFERENCES

F. Rosenqvist. *Direction-dependent processes - Theory and application.* PhD thesis, Chalmers University of Technology, SE-412 96 Göteborg, Sweden, 2004.

K. Takayama, H. Nishina, S. Iyoki, S. Arima, K. Hatou, Y. Ueka, and Y. Miyoshi. Early detection of drought stress in tomato plants with chlorophyll fluorescence imaging-practical application of the speaking plant approach in a greenhouse. In *18th IFAC World Congress*, pages 1785–1790, Milano, Italy, Aug. 28 - Sep. 2 2011.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Control of Blood Glucose for People with Type 1 Diabetes: an in Vivo Study

Dimitri Boiroux * **Signe Schmidt** **
Anne Katrine Duun-Henriksen * **Laurits Frøssing** **
Kirsten Nørgaard ** **Sten Madsbad** ** **Ole Skyggebjerg** *
Niels Kjølstad Poulsen * **Henrik Madsen** *
John Bagterp Jørgensen *

* *Department of Informatics and Mathematical Modeling, Technical University of Denmark, 2800 Kgs Lyngby, Denmark.*
** *Department of Endocrinology, Hvidovre Hospital, 2650 Hvidovre, Denmark*

Abstract:
Since continuous glucose monitoring (CGM) technology and insulin pumps have improved recent years, a strong interest in a closed-loop artificial pancreas for people with type 1 diabetes has arisen. Presently, a fully automated controller of blood glucose must face many challenges, such as daily variations of patient's physiology and lack of accuracy of glucose sensors. In this paper we design and discuss an algorithm for overnight closed-loop control of blood glucose in people with type 1 diabetes. The algorithm is based on Model Predictive Control (MPC). We use an offset-free autoregressive model with exogenous input and moving average (ARMAX) to model the patient. Observer design and a time-varying glucose reference signal improve robustness of the algorithm. We test the algorithm in two clinical studies conducted at Hvidovre Hospital. The first study took place overnight, and the second one took place during daytime. These trials demonstrate the importance of observer design in ARMAX models and show the possibility of stabilizing blood glucose during the night.

1. INTRODUCTION

Type 1 diabetes is a disease caused by destruction of the insulin producing beta-cells in the pancreas. Therefore, patients with type 1 diabetes must rely on exogenous insulin administration in order to tightly regulate their blood glucose. Blood glucose should preferably be kept in the range 4.0-8.0 mmol/l. Long periods of high blood glucose (hyperglycemia) can lead to long-term complications like nerve diseases, kidney diseases, or blindness. However, the dosing of insulin must be done carefully, because a too high dosage of insulin may lead to a too low blood glucose (hypoglycemia). Low blood glucose has immediate effects, such as coma or even death.

The conventional insulin therapy for people with type 1 diabetes consists of the injection of slow acting insulin once a day and rapid acting insulin several times per day. The slow acting insulin is used to counteract the continuous glucose production from the liver. The fast acting insulin compensates the intake of carbohydrates (CHO) during the meals. The decision on the dosage of short and fast acting insulin is based on several blood glucose measurements per day.

However, an increasing number of patients with type 1 diabetes use an intensive insulin therapy based on continuous glucose monitors (CGMs) and insulin pumps instead of the conventional therapy described above. This regime can reduce the risk of complications. CGMs can provide more frequent blood glucose measurements. In addition, insulin pumps can adjust to daily variations in insulin needs.

Nevertheless, the patients still need to be constantly involved in their decisions on the insulin treatment based on their CGMs and/or fingersticks measurements. A system consisting of a CGM, an insulin pump and a control algorithm that computes the insulin dose based on glucose measurements is called an artificial pancreas. The artificial pancreas provides closed-loop control of the blood glucose by manipulation of the insulin injection. The artificial pancreas has the potential to ease the life and reduce complications for people with type 1 diabetes. Its principle is illustrated in Fig. 1. Several review papers about closed-loop control of blood glucose for people with type 1 diabetes have been published (Hovorka et al. (2006), Cobelli et al. (2011), Bequette (2011)).

Previous publications have proven that model predictive control (MPC) has great potential for design of an artificial pancreas. Magni et al. (2009) established that MPC could reduce oscillatory behaviors compared to proportional integral derivative (PID) controllers. Boiroux et al. (2010) applied open-loop constrained nonlinear optimal control. Hovorka et al. (2010) tested an MPC-based controller on children and adolescents with type 1 diabetes.

In this paper we focus on overnight blood glucose control for people with type 1 diabetes using a CGM, an insulin

⋆ Funding for this research as part of the DIACON project from the Danish Council for Strategic Research (NABIIT project 2106-07-0034) is gratefully acknowledged.

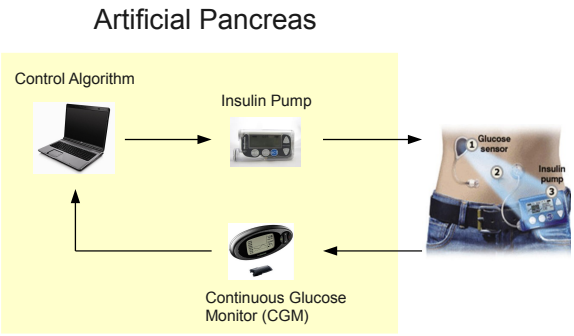


Fig. 1. Closed-loop glucose control. Glucose is measured subcutaneously using a continuous glucose monitor (CGM). Insulin is dosed by an insulin pump.

Fig. 2. Picture of the pilot trial.

pump, and a controller based on MPC. Many factors, such as meal intake, physical exercise, stress, illness, alcohol consumption etc. affect insulin needs. Also, hormone release during the night may cause elevated blood glucose in the early morning. This particular phenomenon is called the dawn phenomenon. The main goal of a closed-loop controller is to compensate these effects by adjusting the amount of injected insulin based on frequent glucose measurements coming from a CGM.

The paper is structured as following. Section 2 describes the material and methods used for the studies. We discuss the design of the controller in Section 3. Section 4 shows the results for the two clinical studies conducted at Hvidovre Hospital. Conclusions are provided in Section 5.

2. METHODS AND MATERIAL

This section describes the clinical protocol and the internally developed graphical user interface for the clinical studies.

2.1 Clinical protocol

The clinical trial consists of a randomized cross-over study including 12 patients with type 1 diabetes. The goal is to compare overnight glucose control during open-loop and closed-loop insulin administration. We investigate the cases where

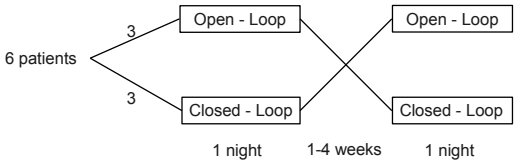


Fig. 3. Study design.

- The insulin bolus matches the evening meal (6 patients in total)
- The insulin bolus is underdosed (6 patients in total)

The study design for the 2 cases is illustrated in Fig. 3.

The scenario during the clinical studies is the following:

- The patient arrives at 16:00.
- A meal is consumed at 18:00 and an insulin bolus is administrated. The meal size is determined by the weight of the patient. The bolus size depends on the patient and the scenario (meal with the correct bolus or underbolused meal).
- The loop is closed at 22:00 (for closed-loop studies only).
- The closed-loop ends at 07:00 the following day (for closed-loop studies only).

The purpose of the first part of the study (when the insulin bolus matches the evening meal) is to validate the ability of the controller to compensate for overnight physiological changes in patients. The second part of the study (when meals are underbolused) must ensure that the controller can bring and keep blood glucose in the range 4.0-8.0 mmol/L.

The patient is equipped with 2 Dexcom Seven Plus CGMs and a Medtronic Paradigm insulin pump. The CGMs provide glucose measurements every 5 minutes. The clinician decides on the sensor used by the controller, based on the accuracy of the sensor during the days before the study. The other CGM can be used as a backup device. Insulin is administrated to the patient through small discrete insulin injections (also called microboluses) every 15 minutes.

It must be pointed out that the pump used for the trials has discrete increments of 0.025U for the microboluses, and a minimum continuous insulin injection (or basal rate) of 0.025 U/hr. The controller handles these restrictions by using hard constraints on the minimal insulin infusion rate and by rounding the suggested microbolus to the nearest 0.025U (see Section 3.6).

In addition, blood samples are taken every 30 minutes in order to measure more accurately the blood glucose (in case of prolonged period of low blood glucose, the sampling time is set to 15 minutes). The blood glucose was measured by Hemocue and after the trial by YSI. These values are not provided to the controller.

The clinician has the authority to prevent severe hypoglycemia by injection of intravenous glucose. Such a decision is based on the glucose history.

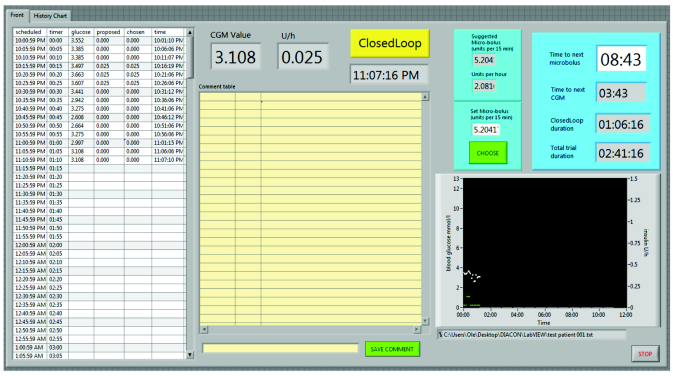


Fig. 4. Graphical User Interface screenshot

2.2 Graphical User Interface

Fig 4 provides an overview of the graphical user interface developed for the artificial pancreas. The glucose sensor provides a glucose measurement every 5 minutes. The glucose measurements are transmitted from the sensor to the software via a wireless receiver.

The graphical user interface returns a new insulin microbolus suggestion every 15 minutes. At these times, it also returns the glucose prediction and insulin prediction profiles. The decision on the insulin microbolus can be overruled if there is a safety risk for the patient. The exact time before the next microbolus suggestion is provided by the graphical user interface.

It is also possible to add comments if necessary. These comments have no influence on the microboluses computation, but are stored.

3. CONTROLLER DESIGN

This section presents the detailed description of the controller. The controller computes a discrete-time offset-free ARMAX model. This model is then used to optimize the future injections of insulin. The controller must be designed in a robust and safe way for the patient, especially regarding low blood glucose. We use here a time-varying glucose setpoint to avoid insulin overdose.

3.1 Model computation

Several research groups investigated low-order models to describe glucose-insulin dynamics. Kirchsteiger et al. (2011) used a third order transfer function, Finan et al. (2009) identified ARX models and Percival et al. (2010) applied a first order transfer function with a delay. In this paper we use a Single Input-Single Output (SISO) second order continuous-time transfer function

$$Y(s) = G(s)U(s), \quad G(s) = \frac{K}{(\tau s + 1)^2} \tag{1}$$

The input $U(s)$ is the insulin intake and the output $Y(s)$ is the blood glucose, both expressed in terms of deviation variables from a steady state, K is the static gain and τ is the time constant. The gain and the time constant are computed from known patient-specific parameters. These parameters are the insulin action time and the insulin sensitivity factor (ISF). They can be estimated for each individual patient by looking at the impulse response for a small insulin bolus. The insulin action time τ corresponds to the time that blood glucose takes to reach its minimum. The insulin sensitivity factor (ISF) corresponds to the maximum decrease in blood glucose per unit of insulin bolus. These parameters are empirically estimated by the patient and his/her physician. However, these parameters may dramatically vary from day to day for a given patient.

The impulse response in the temporal domain of the transfer function (1) is

$$y(t) = K\frac{t}{\tau^2}\exp(-t/\tau) \tag{2}$$

We shall now relate the insulin sensitivity factor and the insulin action time to the gain K and the time constant τ in (2). The insulin action time corresponds to the time to reach the minimum blood glucose, it is therefore equal to τ. We find K by computing the output of the impulse response (2) at its minimum, i.e. at time $t = \tau$. It gives

$$y(\tau) = -I_{SF} = \frac{K}{\tau}\exp(-1) \tag{3}$$

Isolating K in the above equation yields to

$$K = -\tau\exp(1)I_{SF} \tag{4}$$

The transfer function (1) can be reformulated as a discrete-time transfer function model in the form

$$y(t) = G(q^{-1})u(t), \quad G(q^{-1}) = \frac{\bar{B}(q^{-1})}{\bar{A}(q^{-1})} \tag{5}$$

which is equivalent to

$$\bar{A}(q^{-1})y(t) = q^{-n_k}\bar{B}(q^{-1})u(t) \tag{6}$$

$\bar{A}(q^{-1})$ and $\bar{B}(q^{-1})$ are

$$\bar{A}(q^{-1}) = 1 + \bar{a}_1 q^{-1} + \bar{a}_2 q^{-2} \tag{7a}$$
$$\bar{B}(q^{-1}) = \bar{b}_1 q^{-1} + \bar{b}_2 q^{-2} \tag{7b}$$

Fig 5 depicts the exact impulse response and its second order approximation for a virtual patient. This patient is simulated using the model developed by Hovorka et al. (2004). The figure demonstrates that a second order model can provide a fairly good approximation of a patient with type 1 diabetes. Current insulin, such as the Novorapid insulin documented in Nov (2002) has a similar impulse response shape, but can provide even faster action (the minimum in glucose is reached in 60-90 minutes).

3.2 Observer design for the first study

Odelson et al. (2006), Jørgensen and Jørgensen (2007) and Åkesson et al. (2008) proposed several methods for Kalman filter tuning. In our controller we use the following discrete-time, linear ARMAX model

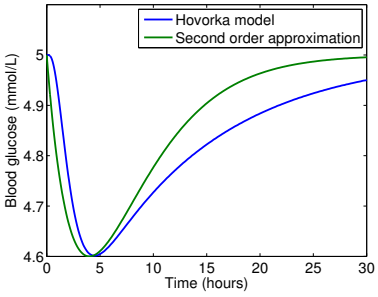


Fig. 5. Example of second order approximation compared to the exact impulse response. The bolus size is here 0.1U. The parameters are: τ=4 hours and $I_{SF} = 4$ mmol/L/U.

$$A(q^{-1})y(t) = q^{-n_k}B(q^{-1})u(t) + C(q^{-1})\varepsilon(t) \tag{8}$$

A, B and C are polynomials, and q^{-1} is the backward shift operator. We assume that $\varepsilon(t) \sim N_{iid}(0, \sigma)$. In the first pilot study we used the following ARMAX model description

$$A(q^{-1})y(t) = B(q^{-1})u(t) + (1 - \alpha q^{-1})e(t) \tag{9}$$

in which

$$A(q^{-1}) = (1 - q^{-1})\bar{A}(q^{-1}) \tag{10}$$
$$B(q^{-1}) = (1 - q^{-1})\bar{B}(q^{-1}) \tag{11}$$

The model (9) is able to provide offset-free tracking due to the integrator. The parameter $\alpha \in [0; 1]$ is a tuning parameter. $\alpha = 0$ corresponds to an integrated ARX model, while $\alpha = 1$ corresponds to an ARX model without integrator. For further details about the choice of α, see e.g. Huusom et al. (2010).

The ARX model (9) may be realized as a stationary state space model in innovation form

$$x_{k+1} = Ax_k + Bu_k + K\varepsilon_k \tag{12}$$
$$y_k = Cx_k + \varepsilon_k \tag{13}$$

The matrices A, B, C and K are written in the canonical form

$$A = \begin{bmatrix} -a_1 & 1 & 0 \\ -a_2 & 0 & 1 \\ -a_3 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$
$$K = \begin{bmatrix} -\alpha - a_1 \\ -a_2 \\ -a_3 \end{bmatrix} \quad C = [1\ 0\ 0] \tag{14}$$

Fig. 6 shows the glucose and insulin predictions for the first study. It can be noticed that the prediction is mostly based on the two previous observations (which show an increasing blood glucose) rather than on the global trend (which shows a decreasing blood glucose).

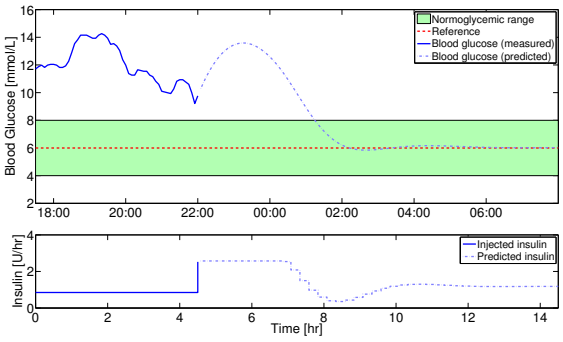


Fig. 6. Example of blood glucose prediction for the first study. It can be seen that the controller relies more on the local trend than on the global trend.

3.3 Observer design for the second study

In this section we consider the general ARMAX model (8) in which we assume

$$C(q^{-1}) = 1 + c_1q^{-1} + c_2q^{-2} + c_3q^{-3} \tag{15}$$

and

$$A(q^{-1}) = (1 - q^{-1})\bar{A}(q^{-1}) \tag{16}$$
$$B(q^{-1}) = (1 - q^{-1})\bar{B}(q^{-1}) \tag{17}$$

in order to preserve the offset-free control property. Therefore, the Kalman gain K in equation (14) becomes

$$K = \begin{bmatrix} c_1 - a_1 \\ c_2 - a_2 \\ c_3 - a_3 \end{bmatrix} \tag{18}$$

(the matrices A, B and C remain unchanged). The design of observer consists of setting the eigenvalues of $A - KC$. Having the eigenvalues close to 0 makes the state estimation error rapidly vanish, but on the other hand the observer will be more sensitive to noise. Having the eigenvalues close to 1 makes the observer less sensitive to noise (and therefore more relying on the global trend) but introduces a delay in the predictions. It can be shown that these eigenvalues are the roots of the polynomial

$$\chi(z) = z^3 + c_1z^2 + c_2z + c_3 \tag{19}$$

$\chi(z)$ is the characteristic polynomial of $A - KC$, and the coefficients c_i, $i = 1, 2, 3$ are the same as the ones in equations (15) and (18). Let α, β_1 and β_2 be the roots of (19). We assume that $\alpha \in \mathbf{R}$, and that β_1 and β_2 are either real or complex conjugate. Furthermore, these roots must all lie inside the unit circle.

As for the first study, we fixed $\alpha = 0.99$. The choice of β_1 and β_2 has been made using data from the first pilot study. For modeling purpose, we considered the stochastic continuous-time model and measurements at discrete times, i.e.

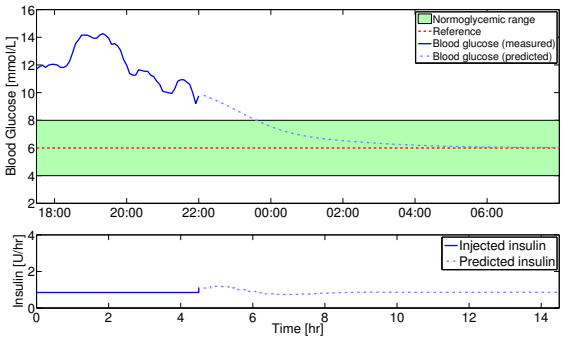


Fig. 7. Example of blood glucose prediction for the first study with the new observer. The controller is able to predict more accurately the blood glucose trend.

$$dx(t) = f(t, x(t), u(t))dt + \sigma d\omega(t) \tag{20a}$$
$$y_k = h(t_k, x(t_k)) + v_k \tag{20b}$$

$x(t)$ are the system states, $u(t)$ are the known inputs (insulin injections, meals and intravenous glucose injections) and y_k are discrete outputs (CGM measurements). The function f is a continuous-time state-space description of the transfer function (1).

We used the internally developed software "Continuous Time Stochastic Modelling" (CTSM) to estimate the variances (variance of process noise and measurement noise) with the maximum likelihood method. We took these variances to compute the predictive Kalman gain K, and hence β_1 and β_2. The computation of β_1 and β_2 yielded

$$\beta_{1,2} = 0.8078 \pm 0.1581i \tag{21}$$

These roots give

$$c_1 = -2.6056 \quad c_2 = 2.2770 \quad c_3 = -0.6708 \tag{22}$$

Fig. 7 illustrates an other example of blood glucose and insulin prediction. We have generated these prediction plots by taking the same data sequence in which we designed the observer. Unlike the previous case in Fig. 6, the controller is able to predict more accurately the blood glucose trend.

3.4 Computing the j-steps ahead predictions

If the k-th glucose measurement y_k is available, the one-step ahead prediction of the states and outputs is

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k} + Bu_{k|k} + K\varepsilon_k \tag{23a}$$
$$\hat{y}_{k+1|k} = C\hat{x}_{k+1|k} \tag{23b}$$

ε_k is the innovation term

$$\varepsilon_k = y_k - C\hat{x}_{k|k-1} \tag{24}$$

In the case where the k-th glucose measurement y_k is not available, the one-step ahead prediction of the states and outputs is

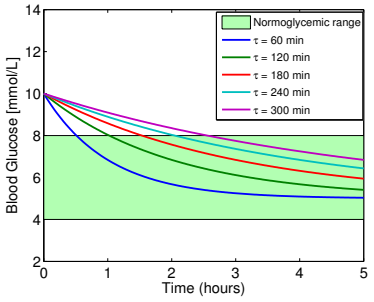


Fig. 8. Example of time-varying reference signal for different values of the time constant τ_r.

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k} + Bu_{k|k} \tag{25a}$$
$$\hat{y}_{k+1|k} = C\hat{x}_{k+1|k} \tag{25b}$$

Similarly, the $j+1$ steps ahead predictions of the states and the outputs for $j = 1, 2, \ldots$ are

$$\hat{x}_{k+j+1|k} = A\hat{x}_{k+j|k} + Bu_{k+j|k} \tag{26a}$$
$$\hat{y}_{k+j+1|k} = C\hat{x}_{k+j+1|k} \tag{26b}$$

3.5 Time-varying glucose setpoint

The glucose trajectory is exponentially decreasing when the blood glucose is above the target, which robustifies the controller with respect to plant-model mismatches. Consequently, the reference blood glucose is

$$\hat{r}_{k+j|k}(t) = \hat{y}_{k|k} \exp\left(-\frac{t_j}{\tau_r}\right) \tag{27}$$

The choice of the tuning parameters τ_r has an influence on the rapidness and the robustness of the controller. Small values of τ_r provide a faster return to the euglycemic range, while larger values of τ ensure a more robust control. The glucose setpoint profiles for different values of the time constant τ_r are shown in Fig. 8.

3.6 Model Predictive Control with Soft Constraints

At the time t_k, the open loop convex quadratic program solved online is

$$\min_{\{u_{k+j}, v_j\}_{j=0}^{N-1}} \phi = \frac{1}{2}\sum_{j=0}^{N-1} \|\hat{y}_{k+j+1|k} - \hat{r}_{k+j+1|k}\|_2^2 + \lambda\|\Delta u_{k+j}\|_2^2 + \kappa\|v_{k+j}\|_2^2 \tag{28a}$$

$$s.t. \quad \hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + Bu_k + Ke_k \tag{28b}$$
$$\hat{y}_{k+1|k} = C\hat{x}_{k+1|k} \tag{28c}$$
$$\hat{x}_{k+j+1|k} = A\hat{x}_{k+j|k} + Bu_k \tag{28d}$$
$$\hat{y}_{k+j+1|k} = C\hat{x}_{k+j+1|k} \tag{28e}$$
$$u_{\min} \leq u_{k+j} \leq u_{\max} \tag{28f}$$
$$G_{\min} - y_{k+1} \leq v_{k+j} \tag{28g}$$
$$v_j \geq 0 \tag{28h}$$

in which $\hat{x}_{k|k-1}$ and $e_k = y_k - C\hat{x}_{k|k-1}$ are given. $u_{\min}$ and $u_{\max}$ are the minimum and the maximum insulin infusion rates allowed by the pump. $\Delta u_{k+j} = u_{k+j} - u_{k+j-1}$ is the variation in the insulin infusion rate. $G_{\min}$ depicts the lower bound on blood glucose. The reference signal $\hat{r}_{k+j+1|k}$ is time-varying and its computation is given in section 3.5.

The slack variables v_j are introduced to penalize hypoglycemia. The hard input constraints (28f) limit the insulin infusion rate. The penalty term $\kappa\|v_{k+j}\|_2^2$ is used to avoid hypoglycemia and the penalty term $\lambda\|\Delta u_{k+j}\|_2^2$ prevents the insulin infusion rate from varying too aggressively.

For the study we choose $N = 120$, i.e. a 10 hour prediction horizon, and

$$\begin{aligned} &u_{min} = -u_{ss} + 0.025, \quad u_{max} = u_{ss}, \\ &\lambda = \frac{10}{u_{ss}^2}, \quad \kappa = 1000 \end{aligned} \tag{29}$$

We remind here that the input variables are deviation variables from the steady state u_{ss}. Consequently, the choice of $u_{min} = -u_{ss} + 0.025$ allows the controller to deliver the minimum basal rate (0.025U/hr), and $u_{max} = u_{ss}$ prevents the pump from overdosing the insulin. The high value of κ makes hypoglycemia undesirable.

4. STUDIES RESULTS

In this section we discuss the two studies conducted at Hvidovre Hospital on the same patient. The patient has an insulin sensitivity factor equal to 5 mmol/L/U and an insulin action time equal to 5 hours. Her basal insulin is $u_{ss} = 0.85$ U/hr.

4.1 Pilot studies results

Fig. 9 depicts the blood glucose and insulin profiles for the first pilot study. The study started at 17:30. A meal has been consumed at 18:00. An insulin overdosing led to severe hypoglycemia and an intravenous glucose injection at approximately 00:00. A microbolus decision has been overruled at 01:30.

Fig. 10 depicts the blood glucose and insulin profiles for the second pilot study. Intravenous glucose has been administrated at 10:00 and 12:00 to compensate for a too high insulin sensitivity. The sensor has to be calibrated at 12:15 and 14:45. In despite of these disturbances, the controller was able to keep the blood glucose within the range 4.0-8.0 mmol/L after the second glucose administration. In addition, the intravenous glucose is not included in the model, and therefore can be considered as an unknown disturbance. However, it can be noticed that insulin is still slightly overdosed.

5. CONCLUSION

This contribution presents a closed-loop controller for people with type 1 diabetes. We described a practical way of computing the glucose-insulin dynamics model. The controller has been tested two times on the same patient. The most noticeable difference between the two studies was the observer design. The trial results illustrated the importance of observer design in state space models in innovation form, and how modelling based on prior data can be used to design the observer. Improvements are being implemented on the controller in order to ensure a more robust control of blood glucose and avoid the observed insulin overdosing during the second pilot study.

REFERENCES

Novorapid Product Monograph. Can be downloaded at: http://www.novonordisk.com/images/diabetes/pdf /Novorapid%20Product%20Monograph.pdf, 2002.

B. M. Åkesson, J. B. Jørgensen, N. K. Poulsen, and S. B. Jørgensen. A generalized autocovariance least-squares method for kalman filter tuning. *Journal of Process Control*, 18:769 – 779, 2008.

B. W. Bequette. Challenges and progress in the development of a closed-loop artificial pancreas. In *American Control Conference 2012 (ACC 2012)*, 2011. Submitted.

D. Boiroux, D. A. Finan, N. K. Poulsen, H. Madsen, and J. B. Jørgensen. Implications and limitations of ideal insulin administration for people with type 1 diabetes. In *UKACC International Conference on Control 2010*, pages 156 – 161, 2010.

C. Cobelli, E. Renard, and B. Kovatchev. Artificial pancreas: past, present, future. *Diabetes*, 60:2672 – 2682, 2011.

D. A. Finan, F. J. Doyle, C. C. Palerm, W. C. Bevier, H. C. Zisser, L. Jovanovič, and D. E. Seborg. Experimantal evaluation of a recursive model identification technique for type 1 diabetes. *J Diabetes Sci Technol*, 3:1192 – 1202, 2009.

R. Hovorka, V. Canonico, L. J. Chassin, U. Haueter, M. Massi-Benedetti, M. O. Federici, T. R. Pieber, H. C. Schaller, L. Schaupp, T. Vering, and M. E. Wilinska. Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. *Physiological Measurement*, 25:905–920, 2004.

R. Hovorka, M. E. Wilinska, L. J. Chassin, and D. B. Dunger. Roadmap to the artificial pancreas. *Diabetes Research and Clinical Practice*, 74:178 – 182, 2006.

R. Hovorka, J. M. Allen, D. Elleri, L. J. Chassin, J. Harris, D. Xing, C. Kollman, T. Hovorka, A. M. F. Larsen, M. Nodale, A. De Palma, M. E. Wilinska, C. L. Acerini, and D. B. Dunger. Manual closed-loop insulin delivery in children and adolescents with type 1 diabetes: a phase 2 randomised crossover trial. *Lancet*, 375:743 – 751, 2010.

Jakob Kjøbsted Huusom, Niels Kjølstad Poulsen, Sten Bay Jørgensen, and John Bagterp Jørgensen. Tuning of methods for offset free MPC based on ARX model representations. In *2010 American Control Conference (ACC)*, pages 2355–2360, Baltimore, MD, USA, 2010.

J. B. Jørgensen and S. B. Jørgensen. Comparison of prediction-error modelling criteria. In *Proceedings of the 2007 American Control Conference (ACC 2007)*, 2007.

Harald Kirchsteiger, Giovanna Castillo Estrada, Stephan Pölzer, Eric Renard, and Luigi del Re. Estimating interval process models for type 1 diabetes for robust control design. In *Preprints of the 18th IFAC World Congress*, pages 11761 – 11766, 2011.

L. Magni, D. M. Raimondo, C. Dalla Man, G. De Nicolao, B. P. Kovatchev, and C. Cobelli. Model predictive con-

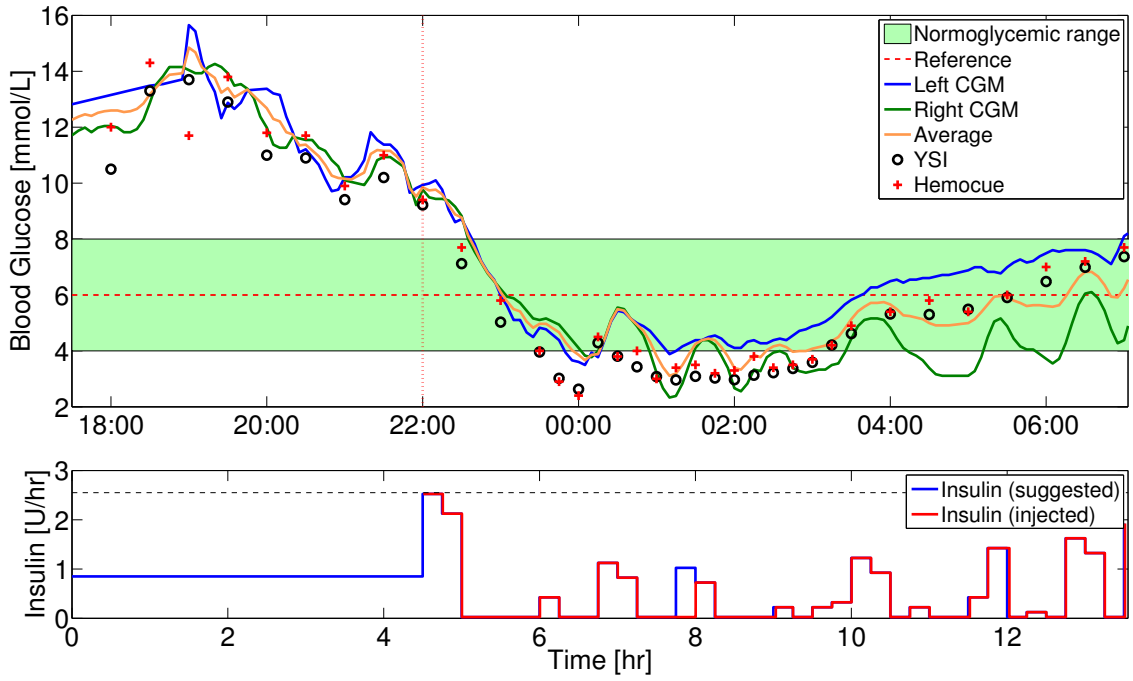


Fig. 9. Blood glucose and insulin profiles for the first pilot study. The insulin infusion rates are computed based on the right CGM (green curve).

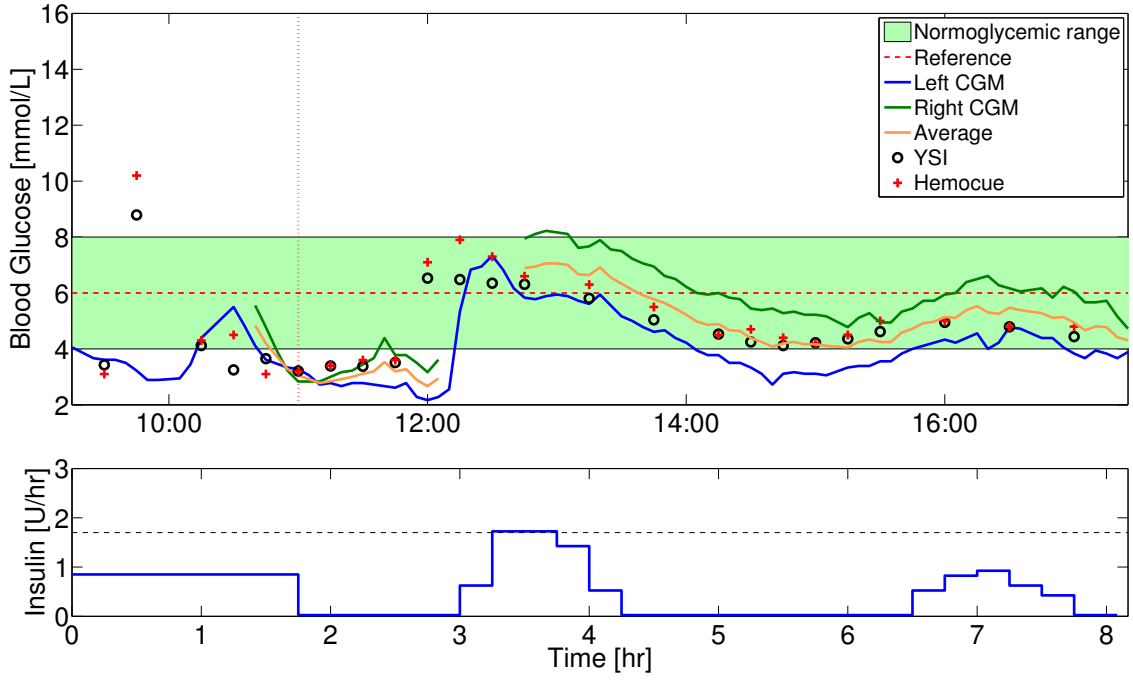


Fig. 10. Blood glucose and insulin profiles for the second pilot study. The insulin infusion rates are computed based on the left CGM (blue curve).

trol of glucose concentration in type I diabetic patients: An in silico trial. *Biomedical Signal Processing and Control*, 4(4):338–346, 2009.

B. J. Odelson, A. Lutz, and J. B. Rawlings. The autocovariance least-squares method for estimating covariances: application to model-based control of chemical reactors. *IEEE transactions on control systems technology*, 14(3):532 – 540, 2006.

M. W. Percival, W. C. Bevier, Y. Wang, E. Dassau, H Zisser, L. Jovanovič, and F. J. Doyle III. Modeling the effects of subcutaneous insulin administration and carbohydrate consumption on blood glucose. *Diabet Sci Technol*, 4(5):1214–1228, 2010.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

A Dantzig-Wolfe Decomposition Algorithm for Linear Economic MPC of a Power Plant Portfolio

Laura Standardi [*], **Kristian Edlund** [**] ,
Niels Kjølstad Poulsen [*], **John Bagterp Jørgensen** [*]

[*] *DTU Informatics, Technical University of Denmark*
({laus,nkp,jbj}@imm.dtu.dk)
[**] *DONG Energy (kried@dongenergy.dk)*

Abstract: Recently, the interest in renewable energy sources is increasing. In the short future, their penetration in the power systems will be significantly higher than today. Denmark is working on achieving its goal by 2020 of having 30% of the energy production provided by renewable sources. 50% of the total power consumption is expected to stem from wind turbines. Due to the inherent stochasticity in renewable energy systems (RES), their energy production is usually complicated to forecast and control. The aim of the smart grid in which consumers as well as producers are controlled is to allow for larger variation in the power production due to the significant amount of renewable energy. The multiple power generators and consumers must be coordinated to balance the supply and demand for power at all times.
The aim of this study is to examine a control technique for large scale distributed energy systems (DES), where a significant amount of renewable energy sources are present. Economic Model Predictive Control (MPC) is applied to control the power generators, minimizing the cost and producing the amount of energy required. We examine the large scale scenario, where multiple power generators and consumers such as e.g. electrical vehicles, heat pumps for domestic heating, and refrigeration and cooling systems must be controlled to balance the supply and demand for power. The system is very large scale. To address the large scale of the system and be able to compute the control decisions within a sample period, Dantzig-Wolfe decomposition is used for solution of the resulting linear program describing the Economic MPC of such systems. The controller obtained has been tested by simulations of a power portfolio system.

Keywords: Decoupled subsystems, Model based control, Predictive control, Optimization, Power system control, Decomposition

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

A model-based FDD method of interacting control loops and its application to a mixing tank process

Karhan Özdenkçi*, Alexey Zakharov*, Sirkka-Liisa Jämsä-Jounela*

*Aalto University, School of Chemical Technology, P.O.Box 16100, 00076, Aalto, Finland
(Tel: +358-9-470 22668; E-mail: karhan.ozdenkci@aalto.fi).

Abstract: The paper addresses fault detection and diagnosis (FDD) of interacting control loops and proposes a model-based method to diagnose all types of faults: sensor faults regarding output variables (output sensor faults) and measured disturbance variables (disturbance sensor faults), actuator faults and process faults. The method consists of an output observer and a bank of special observers. The output observer detects a fault and indicates whether it is an output sensor fault. If only one residual of the output observer deviates, the fault is concluded to be an output sensor fault; otherwise, the special observers are investigated to determine the faulty actuator or disturbance sensor.

One special observer is assigned for each manipulated input (i.e. actuator) and disturbance input variable (i.e. disturbance sensor). Each of these observers involves Kalman filter as state estimation and an auxiliary PI controller. The auxiliary PI manipulates its specific input variable in the estimation of the particular observer to compensate the fault impact on the residual of the pairing output. If and only if the fault is the actuator or disturbance sensor fault for which the observer is assigned, other output residuals of the observer approach zero as well; in other words, the auxiliary PI effectively compensates the fault impact on the estimations of the particular observer. Consequently, all output residuals of one special observer approach zero whereas output residuals deviate in other special observers. To sum up, the method involves the structured residual approach to diagnose actuator and disturbance sensor faults; the output residuals of each special observer are insensitive to a single actuator or disturbance sensor fault and sensitive to all other faults. If there is no specific observer with all zero output residuals, the fault is concluded to be a process fault. The method is beneficial for its ease of implementation and the ability to correctly diagnose all types of faults.

Finally, the method is applied to the fault detection and diagnosis of a basic mixing process. The simulation results demonstrate that the method can effectively diagnose all types of faults in strongly interacting control loops; on the other hand, the reliable diagnosis can hardly be achieved in the case of weak interaction. This indicates the key role of interactions in fault diagnosis. Thus, it should be aimed to maximize the interactions within each subsystem when decomposing a large-scale process. In a large-scale process, the proposed method can be used for a subsystem including MIMO control loops of which the dynamic model is available. Furthermore, the ability to diagnose disturbance sensor faults is beneficial since evaluating the diagnostic results of two sequential subsystems can improve the resolution.

Keywords: Fault detection and diagnosis, interacting control loops, MIMO control loops, structured residuals, actuator faults, sensor faults

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

A New Static Estimator for Estimation of Primary Variables from Combination of Secondary Measurements

Maryam Ghadrdan[1] , Sigurd Skogestad[1] , Ivar J. Halvorsen[2]

1 Department of Chemical Engineering, Norwegian University of Science and Technology, N-7491 Trondheim, Norway, Email: ghadrdan@nt.ntnu.no, skoge@nt.ntnu.no,
2 SINTEF ICT, Applied Cybernetics, N-7465 Trondheim, Norway, Email: ivar.j.halvorsen@sintef.no

Reliable and accurate measurement of product compositions is one of the important issues in distillation column control. However, direct composition measurements are expensive, not so reliable and there is also a significant delay. Temperature measurements are robust, fast and cheap and represents proven technology in the process industry.

Our estimation method (we call it Loss method) is a reformulation of the well-developed Self-optimizing method proposed by Skogestad (2000) to make it adequate for the purpose of estimation. We have formulated the Loss method for open-loop and closed-loop estimators. With the term “open-loop” estimator, it is implied that the predicted variables are not used for control purposes. In "closed-loop" the estimates are used as feedback in control loops. The estimators are optimal in the sense that they give the smallest prediction error which is defined as the difference between the true value and the estimated value. We have compared the performance of the new static estimator with partial least squares (PLS) estimators (Di Ruscio 2000) and Steady State Kalman Filter on a distillation model. Since the focus of our work is on chemical processes, the time scales at which the sensor noise characteristics change are much larger than the time scale at which we study the system. Thus we assume the system and noise covariances are time-invariant. As a result, steady state Kalman Filter is used in this study. PLS and Kalman Filter are two of the estimators which have shown good performance for vast number of applications. However, the idea of the proposed estimator is to give optimal solution for the scenario in which the estimator is supposed to be used, which is proved mathematically.

References

Skogestad, S. (2000). “Plantwide control: the search for the self-optimizing control structure.” J. Proc. Control 10:487-507

Di Ruscio, D. (2000). "A weighted view on the partial least-squares algorithm." Automatica 36: 831-850.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Abstract

Design of Optimal Low-Order Feedforward Controllers for Disturbance Rejection

Martin Hast, Tore Hägglund
Dept. of Automatic Control, Lund University,

Feedforward is an efficient method to reduce control errors both for reference tracking and disturbance rejection, given that the disturbances acting on the system are measurable. This paper treats the subject of disturbance rejection and presents an analytic solution to the problem of designing a feedforward lead-lag filter which minimizes the integrated square error when the system is subject to a measurable step disturbance, d. The resulting feedforward controller is optimal in an open-loop setting, see Figure 1, for first-order plants with time delays, P_i.

In general, due to for instance modeling errors or non-measurable disturbances, feedforward alone is not sufficient in order to reject disturbances in a satisfying fashion. The structure depicted in Figure 2 with $H = 0$ is a common control structure that utilizes both a feedforward and a feedback controller. However, using this structure, the feedback controller C will interact with, or even counteract the feedforward controller. Hence, the feedforward and feedback controllers should be designed jointly. An alternative feedforward scheme was presented in [1]. By choosing $H = P_2P_3 - P_2P_1G_{ff}$, the controller interaction can be eliminated. Furthermore, the design of the feedforward controller reduces to the equivalent problem in the open-loop case.

Along with the presented optimal design rule, examples that illustrate the behavior of the resulting feedforward controller is presented. The performance of the obtained optimal feedforward controllers have also been compared with other available design methods e.g., [2].

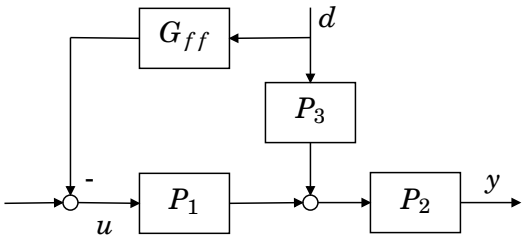


Figure 1 Open-loop structure.

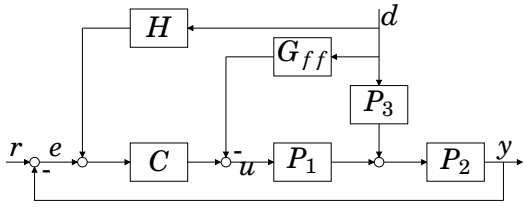


Figure 2 Closed-loop feedforward structure.

References

[1] C. Brosilow and B. Joseph. *Techniques of Model-Based Control*. Prentice Hall PTR, 2002.

[2] José Luis Guzmán and Tore Hägglund. Simple tuning rules for feedforward compensators. *Journal of Process Control*, 21(1):92–102, January 2011.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Educational Toolkit for Teaching FDD Methods

**Venla Kuuluvainen*, Vesa-Matti Tikkala*, Esa Kiiski*,
Sirkka-Liisa Jämsä-Jounela***

** Aalto University School of Chemical Technology, P.O. Box 16100
FI-00076 AALTO (email: venla.kuuluvainen@aalto.fi)*

Abstract: Tightening global competition within the process industries sets higher standards for product quality, production efficiency and safety. These standards can be met with effective fault detection and diagnosis (FDD) systems that can manage complex processes. FDD methods have been widely studied in the literature and applications have been presented to all control hierarchy levels; supervisory, stabilizing and basic control levels. Although the process monitoring task has partly shifted from operators to computers, the knowledge requirements of process operators have not decreased. In addition, the design and implementation of monitoring systems has generated new requirements for engineers entering the field.

This paper presents a toolkit developed for use in a master's level course on process monitoring methods. The course is essential in the education of process automation master students at the Aalto University School of Chemical Technology. The toolkit contains FDD methods for all control hierarchy levels of the plant automation system and an application to a three layer board machine as a case study. The toolkit has been tested with industrial data from the board machine. Additionally, utilizing a user friendly interface, three guided exercises around the use cases were developed for educational purposes: a supervisory level FDD method for a thickness sensor fouling; a stabilizing level FDD method for detecting faults in the drying groups; and a basic level FDD for detecting valve stiction.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Experimental study of anti-slug control on small-scale test rig

Esmaeil Jahanshahi, Sigurd Skogestad
Department of Chemical Engineering, Norwegian University of Science and technology, Trondheim, NO-7491 (e-mail: skoge@ntnu.no).

Abstract: Slug flow condition at offshore oilfields leads to many disadvantages and an effective method is needed to prevent it. Reducing the opening of the top-side choke valve is one solution which reduces the production rate. On the other hand design changes like building slug catcher or full separation are costly. Therefore, active control of the production choke valve is the recommended anti-slug solution. The aim is to have a non-oscillatory flow regime together with maximum of average valve opening which gives the maximum production rate.

Focus of this research is to find structure of a simple yet robust structure for anti-slug control system. In order to find good control variables for stabilization, experiments on a small-scale test rig are performed. The control variables or combination of control variables (in cascade) which could be able to stabilize the system with larger average of the choke valve opening are preferable control variables.

Similar research has been done previously, but it is repeated in this work using a new closed loop tuning method by Skogestad and Grimholt (2011) Using this systematic approach ensures that the results from the different experiments are comparable.

Also, bifurcation diagrams are generated for the test rig and a simplified model by Jahanshahi and Skogestad (2011) is fitted to the experimental data. Finally, experimental results are compared to the results of the controllability analysis using the simplified model.

Keywords: Oil production, two-phase flow, riser slugging, anti-slug control, controllability analysis.

References

Jahanshahi, E. and Skogestad, S. (2011). Simplified dynamical models for control of severe slugging in multiphase risers. In 18th IFAC World Congress, 1634–1639. IFAC.

Skogestad, S. and Grimholt, C. (2011). The SIMC method for smooth PID controller tuning, Chapter for planned PID book, Springer, 29

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Fault Analysis as a Tool for Fault Detection and Diagnosis Development in Large-Scale Systems

Vesa-Matti Tikkala*, Helena Laavi*, Sirkka-Liisa Jämsä-Jounela*

** Aalto University School of Chemical Technology, P.O. Box 16100, 00076 Aalto, Finland (Tel: 358-9-470 23858; e-mail: vesa-matti.tikkala@aalto.fi).*

Abstract: Global competition forces process industries to continuously optimize plant operation. One of the latest trends for efficiency and plant availability improvement is to set up fault diagnosis and maintenance systems for on-line industrial use. Because model or data-based diagnosis of all components cannot be realized on-line on a large-scale basis, a scalable procedure has first to be developed to narrow down the whole plant to the most likely faulty components responsible for abnormal process behavior. One of the key elements in the development of such systems is fault analysis. The aim of the fault analysis is to identify the main causes for production losses and plant shutdowns, to find out the typical faults and the faulty devices and thus to identify the main focus areas for FDD system development. This paper addresses fault analysis performed for a large-scale industrial board machine process. The analysis results are presented and discussed based on this industrial case study.

Keywords: Fault detection, fault diagnosis, large-scale systems, paper industry, statistical analysis

1. INTRODUCTION

Tightening global competition, increased product quality requirements as well as safety and environmental regulations have forced the process industries to continuously optimize the efficiency and profitability of its plants. Better profitability can generally be achieved through process optimization, by cutting costs and by reducing down-time caused by unplanned and planned shutdowns. The optimization can be further enhanced by focusing on preventing off-spec production caused by process disturbances and faults. To this end, there has been an increasing interest in applying process monitoring and fault diagnosis methods in the process industry. Reviews on the applications have been published e.g. by Isermann (2011) and Patton *et al.* (2000).

Process knowledge has always played a key role in the development of fault detection and diagnosis (FDD) systems for process industries. As a result, the FDD methods have been classified in three categories based on the type of knowledge they possess: quantitative model-based, qualitative model-based and process history based methods (Venkatasubramanian *et al.*, 2003a, b, c). The methods in each category have their strengths and weaknesses, and it has been stated that no single method is meeting the requirements for a good diagnostic system (Dash and Venkatasubramanian, 2000). To overcome the disadvantages, hybrid approaches that either combine the results of different methods or combine incomplete process knowledge of methods from different categories have been proposed (e.g. Chung *et al.*, 1994; Lee *et al.*, 2003; Vedam and Venkatasubramanian, 1999).

These methods are generally sufficient for unit processes and small-sized processes, but they become in most cases inefficient in large-scale processes. Therefore, process decomposition-based strategies have been developed to tackle the challenges of large-scale systems. A process can be decomposed in a structural or functional manner by utilizing either a top-down or a bottom-up strategy. For example, Prasad *et al.* (1998) have proposed a decomposition methodology based on the structure of a chemical plant. However, there are no well-defined criteria to evaluate the optimality of these decomposition schemes.

A methodology for developing a FDD system for a large-scale process consists of the main four phases: process decomposition, fault analysis, selection and construction of the diagnostic technique for each subsystem, and integration of the diagnostic information (Jämsä-Jounela, 2011). The final phases of the methodology consist of the validation and industrial implementation of the algorithms (Kettunen, 2010).

A major step in increasing the economic efficiency is to improve the plant availability. Because model or data-based diagnosis of all components cannot be realized on-line on a large-scale basis, a scalable procedure has first to be developed which narrows down the whole plant to the most likely components responsible for the abnormal behavior of the plant.

When developing the FDD system for industrial use, the most important phase is the fault analysis. The aim of the fault analysis is to find out the focus areas for the FDD system development and is accomplished through examining the main reasons for the production losses. In addition, the identification of the most common faults, their characteristics, locations, causes, and the faulty devices

decides eventually the selection of the suitable FDD methods to be used in the system under development.

This paper presents fault analysis performed at a large-scale industrial board machine process in order to facilitate the development of a FDD system. The paper is structured as follows. First, Chapter 2 presents the process description and its control strategy. Second, Chapter 3 describes the test case process decomposition and Chapter 4 presents the performed fault analysis. Finally, Chapter 5 concludes the findings of using the fault analysis as a tool in fault detection and diagnosis development of the board machine.

2. PROCESS DESCRIPTION AND ITS CONTROL STRATEGY

The board making process begins with the preparation of raw materials in the stock preparation section. Different types of pulp are refined and blended according to a specific recipe in order to achieve the desired composition and properties for the board grade to be produced. The consistency of the stock is controlled with dilution water.

The blended stock passes from the stock preparation to the short circulation. First, the stock is diluted in the wire pit to the correct consistency for web formation. Next, the diluted stock is cleaned and screened, after which it passes to the head box, from where it is sprayed onto the wire in order to form a solid board web.

The excess water is first drained through the wire and later by pressing the board web between rolls in the press section. The remaining water is evaporated off in the drying section using steam-heated drying cylinders. After the drying, the board is calendered in two phases in order to achieve desired surface properties. Details of the process can be found in Cheng *et al.* (2011) and in Tikkala *et al.* (2011).

The main control system of the board machine is the quality control system (QCS), which represents the highest level in the control hierarchy. By utilizing model-predictive control schemes, it controls the main quality variables, basis weight moisture, and thickness, in the machine direction and in the cross direction. The quality variables are measured after the calender section with a measurement scanner that traverses constantly across the web. The calculated control actions are delivered as setpoints to lower level controllers.

In the machine direction, the stock flow controller setpoints are adjusted according to the basis weight controller, while the steam pressure setpoints in the drying section are governed by the moisture controller. In the cross direction, the QCS system is controlling special actuators that adjust the profiles of the quality variables. The basis weight profile is controlled with the dilution water in the middle layer headbox, while the moisture profile is controlled with a steam box located before the press section and with a moisturizing device in the drying section. The thickness profile is controlled at the second calender.

These controls are supported by a large number of basic controls, adjusting pressures, flows, level, etc. around the board machine.

3. PROCESS DECOMPOSITION

When developing the FDD system application for large-scale industrial plants, a decomposition methodology based on the structure of the factory is recommended (Prasad *et al.*, 1998). Based on the plant main objectives, the plant is decomposed first to unit processes, second equipment, and finally to field instruments. Plant topology, PI-diagrams, and expert knowledge are used for specifications.

In this case study, the board machine has been first decomposed in nine sections as follows: stock preparation, short circulation, broke processing, wire section, press section, drying section, calender section, reeling, and QCS. Next, the sections are decomposed in equipment and field instruments. As an example, the decomposition of the board machine focusing on the drying section is shown in Fig. 1.

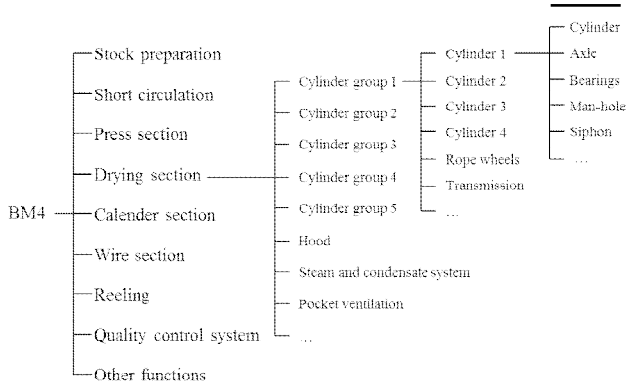


Fig. 1. Decomposition of the board machine focusing on the drying section

4. FAULT ANALYSIS

In the year 2009, the automation system of the board machine at Imatra Mills was updated and the 1st calendar was renewed. Due to these major updates, the board machine was selected as a good candidate for the FDD project.

The fault analysis aims at finding the main focus areas for FDD system development. For this purpose, the long-term production and maintenance data from the year 2010 were collected for the purpose of this study.

In the fault analysis, production losses were first studied by analyzing the web breaks and shutdowns. Next, fault statistics are used to identify the typical fault types in each unit process and the corresponding faulty devices. Finally, the results of the interviews of the plant personnel are used in confirming the findings of the main focus areas for the FDD system development.

4.1 Analysis of the Production Losses

Web breaks and shutdowns are studied as a first phase of the fault analysis. The distribution of the production time, the web breaks, and the shutdowns of the test case are presented in Table 1. These events caused interruptions in the board production for one third of the analyzed time interval during the year 2010.

Both unplanned and planned shutdowns involved in total a three-month interference of the production which was remarkably longer than the additional two-week interference created by the web breaks. Additionally, the statistics show that the web breaks are nearly always due to operational reasons whereas unplanned shutdowns can often have equally a maintenance source. The operational causes consist mainly of process disturbances whereas maintenance faults are, for example, mechanical failures.

The studied year was exceptional in terms of normal production efficiency as it was the first complete production year after implementing the new equipment. The plant experts stated that the reported data depict typical numbers at this stage of implementation of the new device. Typically, start-up related problems last three years.

Table 1. Distribution of production time, web breaks, and shutdowns, and the cause distribution of the web breaks and unplanned shutdowns

Event	Duration		Cause		
	h	*%*		*h*	*%*
Web break	13.2	5	Maintenance	0.6	4
			Operational	12.5	95
			Unspecified	0.2	1
Unplanned shutdowns	42.7	15	Maintenance	21.3	50
			Operational	20.4	48
			Unspecified	1.0	2
Planned shutdowns	49.9	16			
Normal production	186.1	64			
Total	**288.9**				

When producing special products, the sensitivity of each product to web breaks and shutdowns has to be carefully checked. The effect of the produced board grade on the frequency of the web breaks and shutdowns was studied on two types of board, A and B. The types are categorized into three or four board grade blocks according to the board basis weight: low, mid, and high.

The statistics show that board grade blocks with the lowest basis weight have an increased risk for web breaks (Table 2). Also, the board grade block with the highest basis weight is susceptible to web breaks as seen in the case of the board type B. This type of dependency on the board grade block basis weight was not shown for shutdowns, as can be seen in Table 3. The analysis of the shutdowns reveals that the production of the grades in the block B (high) suffers also from repetitive shutdowns.

Table 2. The statistics of the web breaks by grade blocks of board types A and B of various basis weights. The cases with the highest probabilities of breaks are highlighted

Grade block	Web Breaks		
	Production time (h)	Number of breaks	Percentage of production time loss (%)
A (low)	1073	87	6
A (mid-low)	2365	365	4
A (mid-high)	269	36	5
A (high)	199	10	3
B (low)	360	30	6
B (mid)	813	41	4
B (high)	91	8	10

Table 3. The statistics of the shutdowns by grade blocks of board types A and B of various basis weights. The cases with the highest probabilities of shutdowns are highlighted

Grade block	Shutdowns		
	Production time (h)	Number of shutdowns	Percentage of production time loss (%)
A (low)	1073	42	17
A (mid-low)	2365	172	21
A (mid-high)	269	9	15
A (high)	199	3	17
B (low)	360	11	14
B (mid)	813	22	13
B (high)	91	6	31

The results of the analysis of the production losses suggest that the different operation conditions should be considered when developing FDD systems for a board machine e.g. focusing the development only on the specific board grades in question.

4.2 Distribution of the Faults by Fault Types, Process Sections, and Devices

The aim of the fault statistics is also to identify the most typical fault types, the faultiest unit processes, and the devices connected to the faults.

In the study of the typical fault types, a malfunction was reported as the most common fault type. This includes the problems incurred by a device that is functioning but in an incorrect way. As can be seen in Fig. 2, other remarkable fault types were leakages and other damages that produced over 10 % of all faults each. Clogging and jamming or loosening and disengagement presented every tenth fault. Vibration alone gave almost 5 % of the faults.

To study the fault distribution by unit processes, the results of the decomposition of the board machine were used. Among the first eight sections, the faults were distributed quite evenly, but the QCS had twice as many faults as the other sections as seen in Fig. 3.

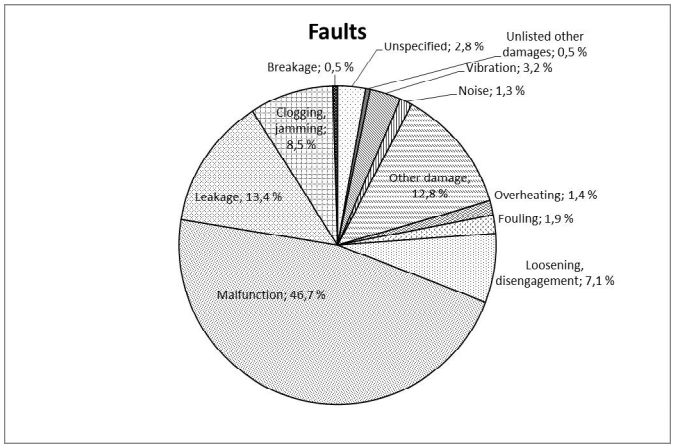


Fig. 2. Distribution of faults by the fault type

The faults that could not be restricted into one of the unit processes were put into the category named Other functions. These include, among others, faults located in the ventilation of the machine hall and other faults from the supporting facilities of the plant.

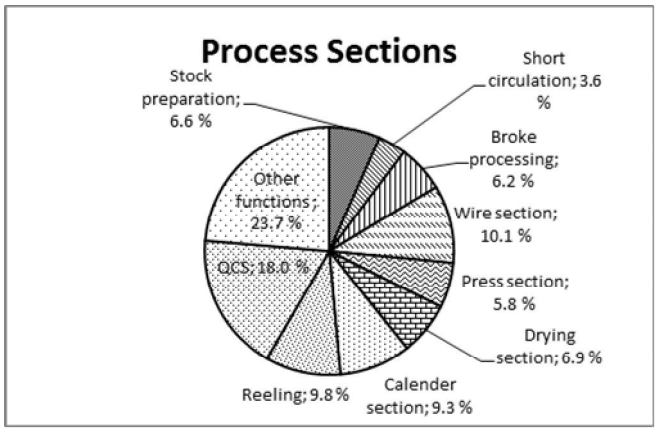


Fig. 3 Distribution of the faults by the process sections

Within the specific process sections, the main fault types and devices were identified. Next, the QCS faults were studied separately as will be explained in Chapter 4.4.

The devices of the process units and their main faults are listed in Table 4. As can be seen in the table, process control devices caused the majority of the faults. In addition, valves were typical sources of malfunctions and leakages.

Well-controlled drying is vital in board making, but the drying section suffers most from various leakages as can be seen in Table 5. Furthermore, the highlighted results indicate that problems in operation of valves cause every fifth fault. Also leakages of pumps and pipes are highlighted as recommended targets of the FDD system development.

Table 4. Classification of the main process devices and their main fault types. The devices chosen for further FDD analysis are highlighted

Fault type	Percentage of all faults	Device
Malfunction	39.3 %	Actuators*
		Automations*
		Control systems*
		Positioners*
		Sensors*
		Transmitters*
		Valves
		Pumps
		Drives
		Drying cylinders
		Hydraulic devices
Leakage	15.5 %	Valves
		Pumps
		Pipes
		Hydraulic devices
		Sensors*
		Rolls
		Heat exchangers
		Tanks
Vibration	4.1 %	Roll

* Process control devices

Table 5. The fault types by device in the drying section of the board machine. The focus areas of FDD development in this section are highlighted, other remarkable fault sources are bordered with dashed line

	DRYING SECTION	Fault type						Total
		Leakage	Loosening, disengagement	Malfunction	Noise	Other damage	Overheating	
Device	Drive	-	-	-	-	-	2.3 %	**2.3 %**
	Drying cylinder	-	6.8 %	4.5 %	-	2.3 %	-	**13.6 %**
	Gear and transmission	4.5 %	-	-	-	2.3 %	-	**6.8 %**
	Heat exchanger	2.3 %	-	-	-	-	-	**2.3 %**
	Mechanical	-	-	2.3 %	-	4.5 %	-	**6.8 %**
	Other mechanical device	-	-	4.5 %	2.3 %	-	-	**6.8 %**
	Pipe	9.1 %	-	-	-	-	-	**9.1 %**
	Positioner	-	-	9.1 %	-	-	-	**9.1 %**
	Pressure device	2.3 %	-	-	-	-	-	**2.3 %**
	Pump	11.4 %	-	-	-	6.8 %	2.3 %	**20.5 %**
	Roll	6.8 %	2.3 %	-	-	2.3 %	-	**11.4 %**
	Valve	2.3 %	-	6.8 %	-	-	-	**9.1 %**
	Total	**38.6 %**	**9.1 %**	**27.3 %**	**2.3 %**	**18.2 %**	**4.5 %**	**100.0 %**

4.3 SISO Level Fault Study

The SISO level study provides supporting data for choosing suitable locations for further FDD algorithm development. The scope of the study was limited to the most common fault types that were identified as malfunctions and leakages.

The malfunctions were mostly related to valves and located in the stock preparation and in the drying section. In addition, also the faults related to the consistency sensor were abundant in the stock preparation and involved further problems in other process sections.

4.4 QCS Fault Analysis

The QCS faults were analysed separately due to its substantial importance to the board making process.

Table 6 lists the origins of all faults occurred in the QCS. Furthermore, the table compares the shares of each cause between the typical QCS fault type categories that are named as malfunctions, sensor malfunctions, and actuator malfunctions.

Table 6. Causes of all QCS faults, sorted by malfunctions, sensor malfunctions, and actuator malfunctions. The main cause is highlighted.

All causes		Mal-functions	Sensor mal-functions	Actuator mal-functions
Component failure	1.8 %	1.1 %	1.6 %	0.0 %
Corrosion/ oxidation	0.9 %	1.1 %	0.0 %	9.1 %
Exceptional conditions	1.8 %	2.1 %	1.6 %	9.1 %
Impurities, moisture	38.6 %	46.8 %	69.4 %	9.1 %
Misoperation	6.1 %	7.4 %	0.0 %	9.1 %
Normal wear	7.9 %	8.5 %	8.1 %	18.2 %
Other failure	3.5 %	3.2 %	1.6 %	0.0 %
Program fault	5.3 %	6.4 %	0.0 %	0.0 %
Safety switch	2.6 %	3.2 %	0.0 %	9.1 %
Unknown/ unspecified	31.6 %	20.2 %	17.7 %	36.4 %
Total	**100.0 %**	**100.0 %**	**100.0 %**	**100.0 %**

As can be seen from the table, impurities and moisture gave almost two faults out of five. In addition, normal wear was a remarkable issue invoking approximately 8 % of all faults.

Considering only malfunctions in general, almost half of the faults in the QCS were due to impurities and moisture, as can be seen in Table 6. Furthermore, the table shows that among sensor malfunctions, this cause was even more common unlike among actuator malfunctions where normal wear was the most common source for a malfunction. Among the sensor malfunctions and the malfunctions in general, normal wear was the second most frequent cause. Other more regular reasons for malfunctions in general were misoperation and program faults even though sensors and actuators hardly suffered from these causes.

4.5 Recommendations for the Main Focus Areas of the FDD Development

As a result of the fault analysis, the following areas were identified as the main focus areas for the FDD development: QCS (board thickness measurements), the drying section (clogging, jamming, and leakages of valves; condensate problems), valves (malfunctions and leakages), and consistency sensor (malfunctions).

At the highest process control and monitoring level, the FDD development should be set on the QCS due to its high share of the faults and its substantial importance to the board making process. Especially, the faults in the measurements of board thickness are recommended to be studied.

At the unit process level, the FDD development is recommended to be directed to the drying section that plays a key role due to its importance and strongly affects to the function of the other sections of the process. Especially, the clogging, jamming, and leakages of valves, and the condensate problems are selected as good candidates for the FDD development.

At the SISO level, malfunctions and leakages are selected for the FDD development In addition, the faults of the consistency sensor, of which the proper function is crucial for obtaining the right board quality, is one candidate for FDD development.

The board making experts confirmed the presented key areas for the FDD system development at the board machine at the Imatra Mills.

5. CONCLUSIONS

In this paper, the fault analysis of a board machine was presented. The performed analysis provided a practical tool and substantial benefits in focusing the FDD development of the large-scale system.

The case study of an industrial board machine confirmed that the fault analysis suited well in screening the target areas of FDD improvements. The FDD development is necessary in this study at all process hierarchy levels, but the needs vary by the level and by the process section.

ACKNOWLEDGEMENT

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2011) under grant agreement n° 257580.

REFERENCES

Cheng, H.; Tikkala, V-M.; Zakharov, A.; Myller, T.; Jämsä-Jounela, S-L. (2011). An Application of the Enhanced Dynamic Causal Digraph Fault Diagnosis Method on a

Three-layer Board Machine. *IEEE Transactions on Control Systems Technology*. 19(3), 644-655.

Chung, H., Bien, Z., Park, J., and Seong, P. (1994). Incipient multiple fault diagnosis in realtime with applications to large-scale systems. *IEEE Transactions on Nuclear Science*. 41(4), pp. 1692-1703.

Dash, S. and Venkatasubramanian, V. (2000). Challenges in the industrial applications of fault diagnosis systems. *Computers and Chemical Engineering*, 24, pp. 785-791.

Isermann, R. (2011). *Fault-Diagnosis Applications,* p. 354. Springer-Verlag, Berlin-Heidelberg.

Jämsä-Jounela, S.-L. (2011). Fault diagnosis methods and their applications in the process industry, invited paper to *COSY 2011, Special International Conference on Complex Systems: Synergy of Control, Computing & Communication*s, Ohrid, Macedonia.

Kettunen, M. (2010). *Data-based fault-tolerant model predictive control: an application to a complex dearomatization process.* Dissertation, Aalto University, Espoo.

Lee, G., Song, S.-O., and Yoon, E. S. (2003). Multiple-fault diagnosis based on system decomposition and dynamic PLS. *Industrial and Engineering Chemistry Research*. 42(24), pp. 6145-6154.

Patton, R.J., Uppal, F. J., and Lopez-Toribio, C.J. (2000). Soft computing approaches to fault diagnosis for dynamic systems: a survey. In: Proceedings of *IFAC Symposium on fault detection, supervision and safety for technical proceses,* Budapest, Hungary (pp. 298-311).

Prasad, P., Davis, J. Jiranpinyo, Y., Josephson, J., and Bhalodia, M. (1998). Structuring diagnostic knowledge for large-scale process systems. *Computers and Chemical Engineering*, 22(12), pp. 1897-1905.

Tikkala, V.-M., Myller T., Kulomaa, T., Hämäläinen, V., Jämsä-Jounela, S.-L. (2011) Data Analysis and Monitoring of Thickness Sensor Fouling Using Self-Organized Maps, *PAPYRUS Workshop,* Corsica, France.

Vedam, H. and Venkatasubramanian, V. (1999). PCA-SDG based process monitoring and fault diagnosis. *Control Engineering Practice*, 7, pp. 903-917.

Venkatasubramanian, V., Rengaswamy, R., and Kavuri, S.N. (2003a), A review of process fault detection and diagnosis part II: Qualitative models and search strategies. *Computers and Chemical Engineering,* 27(3), pp. 313-326.

Venkatasubramanian, V., Rengaswamy, R., Kavuri, S.N., and Yin, K. (2003b), A review of process fault detection and diagnosis part III: Process history based methods. *Computers and Chemical Engineering,* 27(3), pp. 327-346.

Venkatasubramanian, V., Rengaswamy, R., Yin, K., and Kavuri, S.N. (2003c), A review of process fault detection and diagnosis part I: Quantitative model-based methods. *Computers and Chemical Engineering,* 27(3), pp. 293-311.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Application of shape-based stiction detection methods to the critical valves of a board machine

Octavio Pozo Garcia* Sirkka-Liisa Jämsä-Jounela*

**Aalto University School of Chemical Technology,*
Department of Biotechnology and Chemical Technology
P.O. Box 16100, 00076 Aalto
Finland (e-mail: octavioado@gmail.com)

Abstract: Operation of a modern industrial plant can be severely affected by the operation of its key devices, such as valves. Namely, the presence of stiction in a valve can cause poor performance in the control loop and consequently, in the worst cases, degrade the efficiency of the whole process and the quality of the product. Thus, the performance of the control loops must be maintained which requires development and implementation of FDD methods for valve stiction. This paper presents an application of the histogram and curve fitting stiction detection methods for the valves of a board machine, and both methods are based on the shape analysis of oscillating signals in the control loop. The FDD application is tested and validated on case studies selected from the fault analysis and the operational data of an industrial board machine. The results show that both methods can successfully diagnose the presence of valve stiction in an industrial environment.

Keywords: Fault detection and diagnosis, valve, oscillations, shape analysis, stiction, board machine

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Heuristic Evolutionary Random Optimizer for Non-linear Optimization Including Control and Identification Problems

Pasi Airikka*

**Metso Corporation, Tampere, P.O Box 237, FIN-33101*
FINLAND (Tel: +358-40-5872730; e-mail: firstname.lastname@metso.com).

Abstract: Optimal process control solutions such as model predictive controllers or simply PID controllers with optimized performance have a cost function to be minimized. Often, there are additional constraints which cannot be violated when searching the optimum solution in limited but typically multidimensional search space. This paper presents a novel optimizer for solving multidimensional non-linear optimization tasks including control and identification problems. The proposed method combines heuristic and gradient-based methods adding randomness which eventually results in a hybrid optimization solver allowing non-linearities and constraints of an optimized cost function.

Keywords: Optimization, non-linear, evolutionary, heuristic, iterative, control, identification.

1. INTRODUCTION

Optimization of a given cost function is a backbone in control design and, especially, system identification. The cost function is often a scalar valued function with one-dimensional or multi-dimensional search space. The function to be optimized (typically minimized) can be linear or non-linear and with or without constraints. An extensive and elaborate insight to system identification problems with various cost functions to be minimized are given in Söderström & Stoica (1989) and Ljung (1999).

A good example of a control design based optimization problem for the PI controllers is given in Åström et. al (1998) and more can be found in Åström & Hägglund (1995). The multidimensional optimization problems can be faced e.g with model predictive controls. A good review to those formulations are given in Maciejowski (2002).

Numerical non-linear multidimensional optimization methods are numerous and they can be classified e.g in terms of the applied theory. The gradient-based methods include a family of Newton based and quasi-Newton based methods such as Davidon-Fletcher-Powell method (Fletcher 1987, Davidon 1991). The Levenberg-Marquardt method is rather similar to the these methods but is considered a method of its own (Levenberg 1944, Marquardt 1963).

The downhill simplex method (alias amoeba method or Nelder-Mead method) is a good example of non-linear optimizaton method that does not require derivatives of the cost function (Nelder & Mead 1965). Another totally different classes of optimizers for non-linear problems are genetic algorithms (Holland 1975) and evolutionary programming based methods. A good review to both are given in Bäck (1996). What these methods share with the Nelder-Mead method is that they neither require computation of the derivatives of the cost function to be optimized.

This paper presents a novel optimization method combining features of search direction based (gradient-based) methods and evolutionary and genetic algorithms. The resulted iterative method Heuristic Evolutionary Random Optimizer (HERO) does not require computation of the cost function derivatives but involves a heuristic evaluation of search direction at each iteration. Furthermore, genetic algorithm – type mutation based on randomness is included in the solver.

HERO solver is applicable for non-linear optimization tasks with or without equality and inequality constraints. In this paper, the method is described and finally tested with some well-known benchmark functions and, also, with a PI control design and a system identification task.

2. ESSENCE OF OPTIMIZATION

By definition, optimization is about finding the solution for either minimizing or maximizing the given cost function in the presence of given constraints (constrained optimization) or without them (unconstrainted optimization). The constraints can be of either equality or inequality type, or both. Typically, the cost function is a scalar-valued function with several variables for optimization in the multidimensional search space. The cost function itself can be linear or non-linear such as quadratic.

There are numerous different optimization problems and almost equally many optimization algorithms (solvers). Yet, for any optimization solver, the following pitfalls are lurking and challenging any solver:

- Large search space and/or several dimensions
- Shape of search space
- Equality and/or inequality constraints
- Existence of several local optima
- Numerical accuracy requirements

To meet the requirements above, Heuristic Evolutionary Random Optimizer (HERO) was designed with the idea of coming up with an optimization algorithm that would combine the essence of derivative-based and evolutionary algorithms for solving multidimensional non-linear optimization problems.

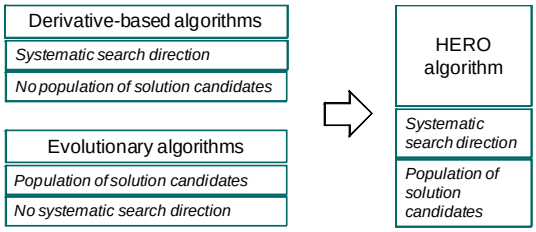


Figure 1. HERO optimizer combines features of both derivative-based and evolutionary optimization routines.

3. PROPOSED OPTIMIZER

Consider a scalar-valued cost function f with an n-dimensional search space subject to minimization:

$$\min_{\mathbf{x}} f(\mathbf{x}), \mathbf{x} \in \Re^n, \ \mathbf{x}_{\min} \le \mathbf{x} \le \mathbf{x}_{\max} \tag{1}$$

There may be equality and/or inequality constraints such as

$$g(\mathbf{x}) \le 0 \tag{2}$$

which can be treated by including a penalty in the cost function (1) as follows

$$\min_{\mathbf{x}} \left(f(\mathbf{x}) + \max\{0, \rho \cdot g(\mathbf{x})\} \right), \ \rho > 0 \tag{3}$$

3.1 Initialization

Given the task of (1) or (3), HERO solver needs to be initialized first. In initialization, a set of population members $\mathbf{x}_i(0)$, i =1…N are placed at equidistant intervals over the search space. Consequently, the population size is N = *grid* ^ n, where *grid* is the resolution of the grid. The grid size is a design parameter of HERO solver. Figure 2 illustrates what the initialized grid might look like for a two-dimensional search space (*grid* = 5). The initial population with N = 25 is evenly spread over the search space.

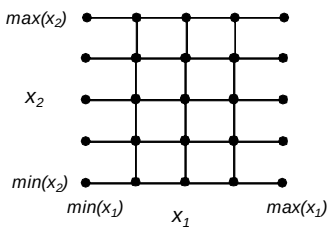


Figure 2. Initialized grid for solution population for a 2-dimensional optimization problem.

3.2 Iteration cycle

After initialization, all the initialized population members $\mathbf{x}_i$, (i = 1…N) are being iterated based on the previous population members. The iteration formula for updating the population members is:

$$\mathbf{x}_i(k+1) = \mathbf{x}_i(k) + \alpha\, \mathbf{d}_i(k) + \mu\, \mathbf{e}_i(k) \tag{4}$$

where k is iteration cycle, i = 1…N denotes the i^{th} population member, $\mathbf{d}_i$ search direction for the i^{th} population member, $\mathbf{e}_i$ random factor for the i^{th} population member, α search speed (0-1) and μ randomness factor (0-1). Given (4), a new population member $\mathbf{x}_i(k+1)$ is formed from the previous member $\mathbf{x}_i(k)$ corrected by the deterministic search direction $\mathbf{d}_i(k)$ and the random factor $\mathbf{e}_i(k)$.

3.3 Search direction and speed

The search direction $\mathbf{d}_i(k)$ is an essential part in a gradient/derivative-based optimization algorithm where the direction is typically computed using the first (Jacobian) or the second derivatives (Hessian) of the cost function or their estimates. Here, however, the search direction $\mathbf{d}_i(k)$ is computed without cost function derivatives as follows:

$$\mathbf{d}_i(k) = \mathbf{x}_{best}(k) - \mathbf{x}_i(k) \tag{5}$$

where $\mathbf{x}_{best}(k)$ is the best solution of all N population members at the iteration cycle k. The computation formula for the search direction is intuitively appealing as obviously all the population members are always being guided towards the best solution at that time. This heuristic selection is rather simple but, as proven later with test benches, it is also rather efficient.

The search speed α (0-1) is a design parameter determining the length of the step towards the best current solution at each iteration cycle. By setting $\alpha = 0$ there is no deterministic search direction but only the random factor affecting. By setting $\alpha = 1$, the population member is always taken exactly to the current best solution. It is advisable to avoid these extreme choices by allowing values $0 < \alpha < 1$. The search speed can be given an interpretation of a first-order low-pass filter with a time constant T indicating how many iterations would be ideally needed for reaching the best solution at that iteration. In the later simulations, the search speed with T = 3 is set to $\alpha = 1 - \exp(-1/T) = 0.72$.

3.4 Random factor

Random factor $\mathbf{e}_i(k)$ is white noise having its elements initialized to zero. The propability of having non-zero elements depends on progress of finding the best solution. If during consecutive iterations there is no progress, propability increases and, vice versa. The randomness factor μ (0-1)

can be made adaptive by adjusting it as a function of search direction norm during each iteration cycle. The random factor and its adaptation has a lot of common with mutation operation in genetic algorithms. Consequently, the random factor has an important role in preventing the algorithm to get stuck with a local optimum.

3.5 Termination of optimization

An optimization algorithm always needs a termination criterion. It is suggested to use a pre-set limit for either a minimum absolute change in the solution or a minimum absolute change in the cost function to be minimized. In addition, the maximum number of iterations should be limited prior to calling the optimization routine.

4. TEST BENCHES FOR HERO

There are several known test benches for testing optimization routines. In this paper, three of them are selected: functions of Rosenbrock, Rastrigin and six-hump camel back. In addition, a novel cost function for tuning a PI controller is given and used for testing performance of the proposed optimization routine. Also, a real system identification case of a power boiler is given using the introduced HERO solver. The boiler can be considered as a first-order dynamic multivariable process with two inputs and one output. The data used for model identification covers half a year's boiler operation.

4.1 Rosenbrock function

The Rosenbrock function is

$$f(\mathbf{x}) = (1-x_1)^2 + 100(x_2 - x_1^{\,2})^2 \qquad (6)$$

with the constraints -3 < x_1 < 3 and -3 < x_2 < 3. The function has a long, narrow, parabolic shaped flat valley with a single global optimum $f(1,1) = 0$. The valley is rather easy to find but the optimum is not due to the valley's flatness (figure 3).

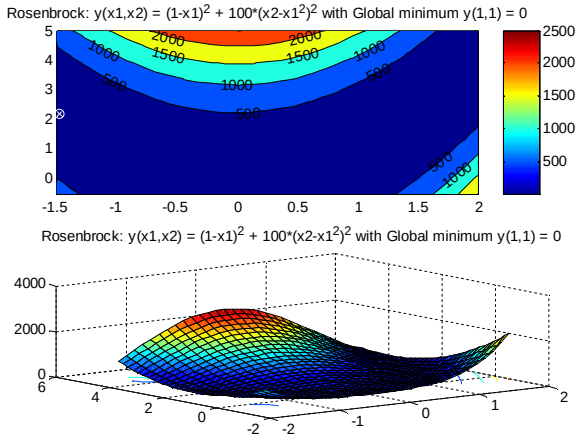


Figure 3. Rosenbrock function. Upper: bird-eye view with colormap. Lower: 3D view.

For comparison, HERO solver is compared to genetic algorithm (GA) and Nelder-Mead (NM) algorithm. Testing is based on 100 optimization runs that were executed with the results given in the table 4. Based on the results, HERO has the best performance (cost function average) with no single failure during 100 different optimization runs.

Table 4. Comparison of optimizaton runs between GA, NM and HERO solvers for Rosenbrock function.

Algorithm	Cost function Average	Cost function Std	Failed optimization runs (%)	Details
GA	0.0427	0.0521	0	Population = 36, Crossover rate = 0.2, Mutation rate = adaptive
NM	0.0710	0.2568	6	Random initialization
HERO	0.000277	0.000375	0	Grid = 6, a = 0.72, μ = adaptive

4.2 Rastrigin function

The Rastrigin function is

$$f(\mathbf{x}) = 20 + (x_1^{\,2} - 10\cos(2\pi x_1)) + (x_2^{\,2} - 10\cos(2\pi x_2)) \qquad (7)$$

with the constraints -3 < x_1 < 3 and -3 < x_2 < 3. The function has several local minima regularly distributed with a single global optimum $f(0,0) = 0$. It is difficult to find the global optimum (figure 5).

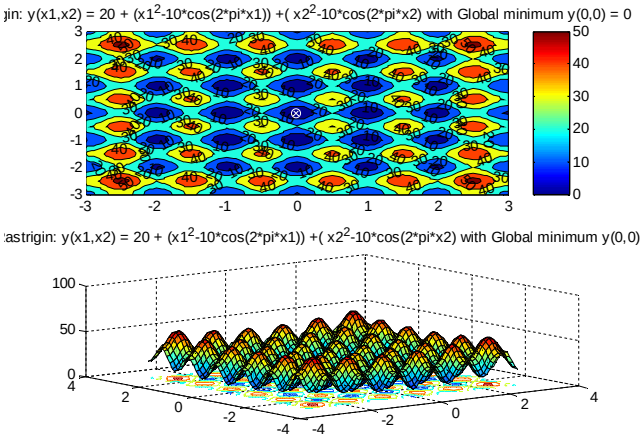


Figure 5. Rastrigin function. Upper: bird-eye view with colormap. Lower: 3D view.

For comparison, HERO solver is again compared to genetic algorithm (GA) and Nelder-Mead (NM) algorithm with 100 optimization runs with the results given in the table 6. Based on the results, HERO has no single failure during 100 different optimization runs and it gives the second best performance (cost function average) after the GA method.

Table 6. Comparison of optimizaton runs between GA, NM and HERO solvers for Rastrigin function.

Algorithm	Cost function Average	Cost function Std	Failed optimization runs (%)	Details
GA	0.000002	0.000008	0	Population = 36, Crossover rate = 0.2, Mutation rate = adaptive
NM	2.3183	1.9444	90	Random initialization
HERO	0.0299	0.0240	0	Grid = 6, a = 0.72, μ = adaptive

4.3 Six-hump camel back function

The six-hump camel back function is

$$f(\mathbf{x}) = (4 - 2.1x_1^2 + x_1^{4/3})x_1^2 + x_1x_2 + (-4 + 4x_2^2)x_2^2 \tag{8}$$

with the constraints $-3 < x_1 < 3$ and $-3 < x_2 < 3$. The six-hump camel back function has six local minima two of which are also global minima $f(-0.0898,0.7126) = f(0.0898,-0.7126) = -1.0316$. The global minimum is especially hard to find due to six local minima (figure 7).

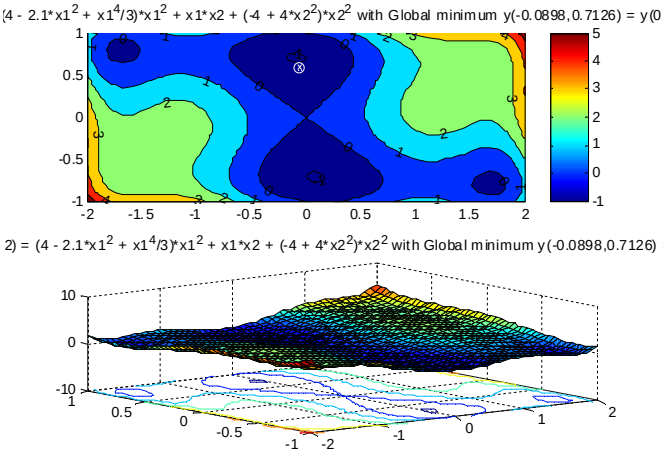


Figure 7. Six-hump camel back function. Upper: bird-eye view with colormap. Lower: 3D view.

Again, for comparison, HERO solver is compared to genetic algorithm (GA) and Nelder-Mead (NM) algorithm with 100 optimization runs with the results given in the table 8. Based on the results, HERO has no single failure during 100 different optimization runs and it also has the best performance (cost function average).

Table 8. Comparison of optimizaton runs between GA, NM and HERO solvers for six-hump camel back function.

Algorithm	Cost function Average	Cost function Std	Failed optimization runs (%)	Details
GA	-1.0277	0.0123	0	Population = 36, Crossover rate = 0.2, Mutation rate = adaptive
NM	-0.8478	0.5813	14	Random initialization
HERO	-1.0316	0	0	Grid = 6, a= 0.72, μ = adaptive

4.4 PI controller optimization

The PI controlled process model is a pure delay process

$$g(s) = e^{-Ls} \tag{9}$$

with a dead time of L = 1 sec. Given a cost function of an integrated error due to a step load disturbance disturbing the closed control loop

$$J_{IE} = \int_0^\infty e(t)dt = \frac{1}{k_i} \tag{10}$$

and a constraint on maximum sensitivity M_s = 1.4, the PI controller parameters proportional gain k_p and integral gain k_i can be obtained by minimizing the cost function J_{IE}. It has been shown by Åström et. al (1998) that these parameters are k_p = 0.158 and k_i = 0.472. The method given by Åström et. al is rather elegant with variable elimination for simplified iterable equations. However, for comparison with HERO solver, the cost function J_{IE} is now re-formulated as follows:

$$J = (M_s - \hat{M}_s)\int_0^\infty e(t)dt = \frac{M_s - \hat{M}_s}{k_i} \tag{11}$$

where M_s is designed maximum sensitivity and $\hat{M}_s$ real maximum sensitivity. The given formulation (11) simplifies optimization as the constraint of M_s is now included directly in the cost function J to be minimized leaving no constraints other than a stability region determined by proportional gain k_p and integral gain k_i as illustrated in figure 9.

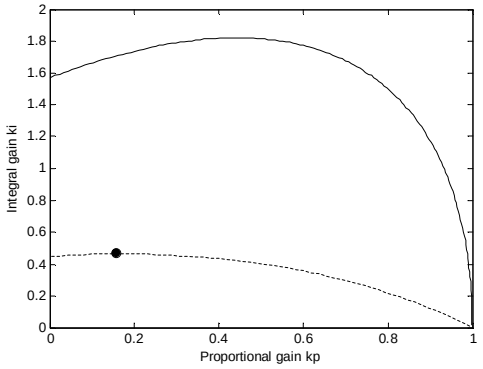


Figure 9. PI controller tuning space with a stability region (limited by solid line) and with a set of PI tunings for M_s = 1.4 (tunings on dotted curve). The optimal tuning for minimizing J_{IE} for M_s = 1.4 is marked with a black dot.

As previously, HERO solver is compared to genetic algorithm (GA) and Nelder-Mead (NM) algorithm with 100 optimization runs with 50 iterations each with the results given in the table 10. Based on the results, HERO has no single failure during 100 different optimization runs and it also has the best performance in terms of the optimal solution. In this comparison, all the methods yielded rather equally low cost function values which is due to the flat shape of the M_s curve (fig. 9) near the optimum.

Table 10. Comparison of optimizaton runs between GA, NM and HERO solvers for a PI controller design. The optimal parameters are k_p = 0.158 and k_i = 0.472.

Algorithm	Optimal Kp Average	Optimal Ki Average	Failed optimization runs (%)	Details
GA	0.1926	0.4153	0	Population = 36, Crossover rate = 0.2, Mutation rate = adaptive
NM	0.2949	0.4274	3	Random initialization
HERO	0.1619	0.4589	0	Grid = 6, a = 0.72, μ = adaptive

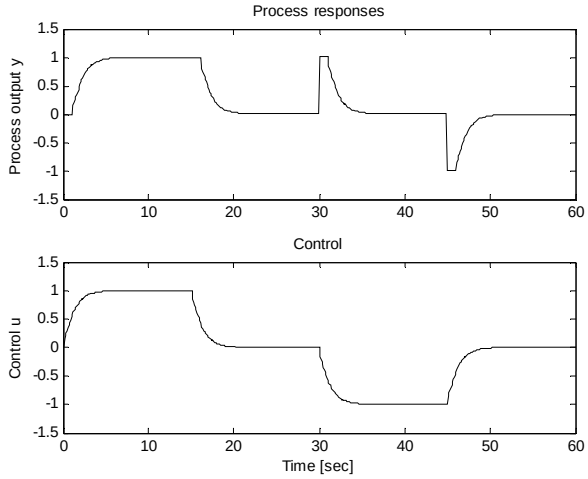


Figure 11. Process responses for the optimal PI controller tuning. Upper: Process output for setpoint up/down changes and up/down load disturbances. Lower: Control signal.

4.5 Process model identification

There is a co-generation power plant with a 280 MW bubbling fluidized bed boiler burning biomass, peat and sludge. The boiler inputs are fuel flow, total combustion air and feed water flow and the boiler output is its generated steam. The task is to identify a simple first-order multivariable model using the real collected process data on 6 months (180 days) with 1 hour sampling rate.

Basically, the data set with 1 hour sampling reflects steady states of the boiler. Therefore, the feed water flow can be ignored as a model input because, typically, in steady-states with a constant steam drum level, the feed water flow equals to the steam flow out of the boiler. Consequently, the simplified first-order discrete-time 2x1 process model subject to identification can be formulated as:

$$y(k+1) = \mathbf{A}y(k) + \mathbf{B}\begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix} + e(k) \qquad (12)$$

where y = generated steam flow (kg/sec), u_1 = fuel power (MW), u_2 = total combustion air (kg/sec) and e = white noise (kg/sec). Figure 12 shows the measurements of boiler live steam, fuel power and total combustion air for 6 months. There are faster hour and daily based changes due to a varying boiler load but also a slower decreasing trend due to a seasonal shift from winter to summer.

The cost function J to be minimized for identifying the process model is the 2-norm of the model residual:

$$J(\mathbf{A},\mathbf{B}) = \left\|\mathbf{y} - \hat{\mathbf{y}}\right\|_2 = \left(\sum_{k=1}^{N} \left| y(k) - \hat{y}(k) \right|^2 \right)^{1/2} \qquad (13)$$

where N = (6 x 30 x 24) / 2 = 2160.

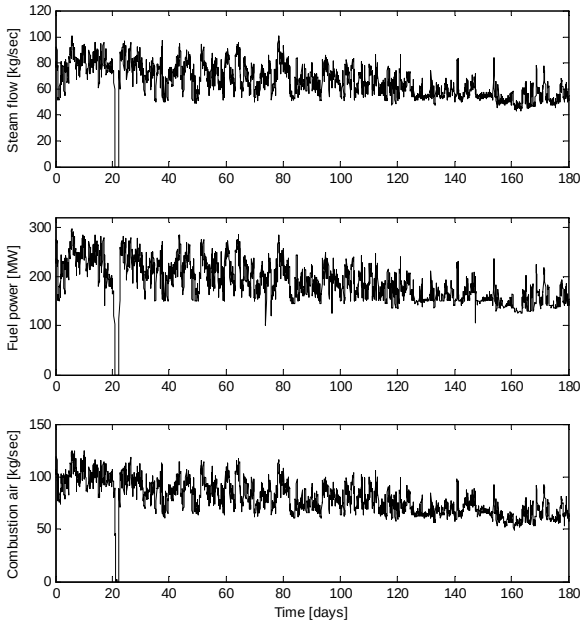


Figure 12. Boiler measurements for 6 months with 1 hour sampling. Top: generated steam (kg/sec). Middle: Fuel power (MW). Bottom: Total combustion air (kg/sec).

For model identification, the first 3 months (90 days) of the recorded process data is selected. The HERO solver is run with 200 iterations and with *grid* = 3 only. There are three parameters to be estimated in **A** and **B** in the range of -2…2. As a result, the minimized cost function receives a value of J = 234.1 after 200 iterations with the model parameters $\mathbf{A} = 0.8284$ and $\mathbf{B} = \begin{bmatrix} 0.0368 & 0.050 \end{bmatrix}$ having a static gain 0.51. The identification results are plotted in figure 13. The parameters seem to have almost converged before HERO solver was stopped after 200 iterations.

For comparison, model identification (12) using a Least Squares (LS) method yields the parameters $\mathbf{A} = 0.8438$ and $\mathbf{B} = \begin{bmatrix} 0.0294 & 0.0547 \end{bmatrix}$ having a static gain 0.54. The cost function J = 445.9 is clearly higher than with the HERO solver, however, it must be kept in mind that the LS method has a slightly different cost function which it minimizes and comparison of that cost function would not be beneficial for HERO solver.

For model validation, the last 3 months (90 days) of the recorded process data is used. Figure 14 shows both comparison of real and predicted process output zoomed for 20 days and the model residual for the whole validation data period of 90 days.

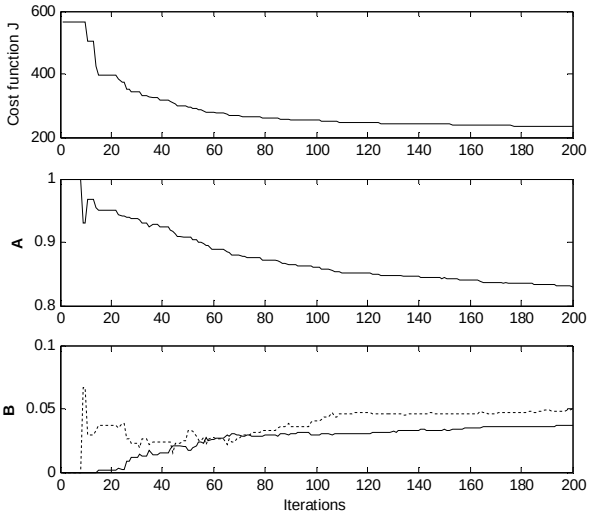


Figure 13. Model identification of the power boiler with HERO. Top: Cost function *J*. Middle: Model parameter **A**. Bottom: Model parameters **B**.

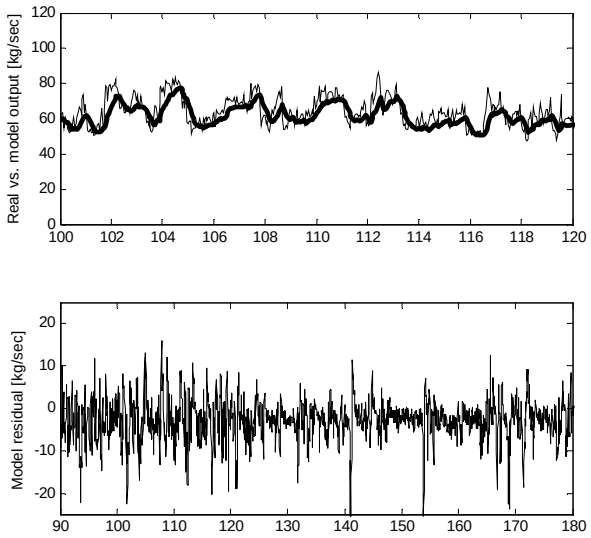


Figure 14. Model identification fit of the power boiler data. Upper: Zoomed set of real (thin) and predicted boiler output (thick) for 20 days. Lower: Model residual (kg/sec) for the whole validation data of 90 days.

5. CONLUSION

This paper presented a novel method called HERO (Heuristic Evolutionary Random Optimizer) for solving non-linear multidimensional optimizaton problems with constraints. The iterative method combines features of search direction based and evolutionary/genetic algorithm based algorithms. The proposed HERO solver for optimization task as given by (1) or (3) does not require computation of the derivatives/gradients of the cost function. It is population-based method such as genetic algorithms having also a mutation function in terms of the random term. Calculation of the search direction is based on heuristic approach of heading towards the currently best solution. And, in order to not get stuck with a local minimum, there is adaptation mechanism based on mutation to excite the solver over the local solution.

REFERENCES

Bäck, T. (1996). *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms*. Oxford University Press US, January 1996.

Davidon W.C (1991). Variable Metric Method for Minimization, *SIOPT Vol. 1,. Issue 1*, pp. 1–17.

Fletcher, Roger (1987). *Practical methods of optimization* (2nd ed.), New York: John Wiley & Sons.

Holland, J.H. (1975), *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor.

Levenberg, K. (1944). A Method for the Solution of Certain Non-Linear Problems in Least Squares. *The Quarterly of Applied Mathematics*, 2, pp. 164–168.

Ljung L. (1999). *System Identification — Theory For the User*, 2nd edition, PTR Prentice Hall, Upper Saddle River, N.J.

Maciejowski, J.M (2002). *Predictive Control with Constraints*, Prentice Hall, 2002.

Marquardt, D. (1963). An Algorithm for Least-Squares Estimation of Nonlinear Parameters, *SIAM Journal on Applied Mathematics*, 11 (2), pp. 431–441.

Nelder, J. A., R.A. Mead (1965). A simplex method for function minimization, *Computer Journal*, 7, pp. 308–313.

Press, W.H., S.A Teukolsky, W.T Vetterling, B.P Flannery, (2007). *Numerical Recipes: The Art of Scientific Computing*, 3rd edition, New York, Cambridge University Press.

Söderström, T., P. Stoica (1989). *System Identification*, Prentice Hall, Upper Saddle River, N.J.

Åström K.J, T. Hägglund (1995). *PID Controllers: Theory, Design and Tuning*, 2nd ed., Instrument Society of America, USA.

Åström, K.J, H. Panagopoulos, T. Hägglund (1998). Design of PI Controllers Based on Non-Convex Optimization, *Automatica*, Vol. 35, No. 5, pp. 585-601.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Industrial validation of the dynamic BioGrate boiler model

Alexandre Boriouchkine *, Sirkka-Liisa Jämsä-Jounela **

* Aalto University, School of Science and Technology, Department of Biotechnology and Chemical Technology, Process Automation Research Group, 00076 Aalto, Finland. e-mail: alexandre.boriouchkine@aalto.fi (corresponding author, phone: +358-9-470 23178).
** Aalto University, School of Science and Technology, Department of Biotechnology and Chemical Technology, Process Automation Research Group, 00076 Aalto, Finland. e-mail: sirkka-l@hut.fi

Abstract

Obligations for increasing the sustainability level of energy resources have presented industry with a challenge to replace fossil fuels with renewable ones. This poses additional challenges since, energy systems based on fossil sources have been used for decades, whereas, sustainable systems are a relatively young technology. Consequently, the complete replacement of non-renewable fuels with renewable ones requires considerable technological effort. Therefore, significant amount of research is required before sustainable energy systems could reach the required maturity.

One of the most mature renewable energy technologies and the most commercially appealing is biomass combustion. Biomass combustion is successfully utilized in BioGrate-boiler technology, which has been developed by MW Biopower, nevertheless, biomass combustion presents several difficulties. Specifically, the properties of biomass vary significantly depending on its origin, processing and handling techniques. Moreover, variable properties cause large fluctuations in power production and thus, set challenges for an existing control strategy to maintain the process within its constraints. As a consequence, improving the existing control strategy and, thus, boiler efficiency, requires an extensive knowledge and understanding of biomass combustion. This knowledge is not always possible to obtain through direct process measurements due to aggressive furnace environment. Thus, it must be obtained by an alternative way, namely, through a mechanistic model.

This paper considers the utilization of first principles model of a BioGrate boiler in a biomass combustion study. The dynamic model of a BioGrate boiler, upon which the study is based, is heterogeneous, including solid and gas phases. Furthermore, the model considers chemical reactions in both gas and solid phases. In addition, fuel movement on the grate is included into the model. The energy required by the process is employed through a radiation function validated by industrial data. The model is implemented in a MATLAB environment and tested with industrial data. The results of detailed industrial validation for the model are presented and discussed.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Iterative Methods for MPC on Graphical Processing Units

Nicolai Fog Gade-Nielsen *, **John Bagterp Jørgensen** *
Bernd Dammann *

* *DTU Informatics, Technical University of Denmark*
({nfga,jbj}@imm.dtu.dk)

Abstract:
The high floating point performance and memory bandwidth of Graphical Processing Units (GPUs) makes them ideal for a large number of computations which often arises in scientific computing, such as matrix operations. GPUs achieve this performance by utilizing massive parallelism, which requires reevaluating existing algorithms with respect to this new architecture. This is of particular interest to large-scale constrained optimization problems with real-time requirements.
The aim of this study is to investigate different methods for solving large-scale optimization problems with focus on their applicability for GPUs. We examine published techniques for iterative methods in interior points methods (IPMs) by applying them to simple test cases, such as a system of masses connected by springs. Iterative methods allows us deal with the ill-conditioning occurring in the later iterations of the IPM as well as to avoid the use of dense matrices, which may be too large for the limited memory capacity of current graphics cards.

Keywords: Model based control, Predictive control, Optimization, Iterative methods, Graphical Processing Unit

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Minimum Backed-off Operating Point Selection

Nabil M*[#], Sridharakumar Narasimhan*, Sigurd Skogestad[#]

*Indian Institute of Technology Madras, Chennai 600 036
India (e-mail: nabil123@gmail.com, sridharkrn@iitm.ac.in).
[#] Norwegian University of Science and Technology(NTNU), N7491 Trondheim
Norway, (e-mail: skoge@ntnu.no)

Abstract

Profitability is the major concern of a chemical plant and is achieved by operating the plant at the nominal point obtained from the optimizer. The optimizer minimizes a suitable cost function subject to equality and inequality constraints. Often, the solution of the optimizer is constrained i.e, there are several active constraints. Typically, it is assumed that these active constraints are controlled. However, presence of uncertainties in the form of measurement noise, modeling error, parametric uncertainties and disturbances cause violation of constraints. Thus, the suggested approach to ensure feasible operation is to keep away from constraints which in turn results in the lost profit [1]. Therefore the objective of our work is to propose a formulation that obtains the trade off solution between feasibility and profitability.

In this work, we propose a stochastic formulation that ensures feasible operation for the prescribed confidence limit. In this formulation, we have assumed full state feedback and disturbance as the only source of uncertainty and is characterized by Gaussian white noise. Following Peng et al[2], the dynamic operating region is defined for the given disturbances which follow from the closed loop covariance analysis of the state space model of the process. Controller selection also plays a crucial role in shaping the dynamic operating region while the size of the region is characterized by the prescribed confidence limit and variance of the disturbance considered. Thus consideration of the controller gain as a decision variable is important in determining the optimal operating point which minimizes the loss in profit. Therefore, the focus of our work is to propose an optimization formulation that determines the minimum backed off operating point by finding a suitable controller gain.

The formulated minimum back off operating point selection problem is a non convex non linear program with linear cost to be solved which does not guarantee an optimal solution. Thus the problem is relaxed by exploiting convexity in the constraints using Linear Matrix Inequality (LMI) theory. Since not all constraints could be convexified, we propose a simple two stage iterative procedure that reduce the variability of the economically important (i.e., active constrained) variables by progressively increasing the variability of the economically unimportant variables at each iteration. The optimal solution from our proposed solution methodology is compared with Peng et al[2] solution obtained using computationally expensive branch and bound technique. Work in progress to extend the current formulation for quadratic cost such that back-off for the nominally unconstrained variables could also be determined.

Acknowledgement: We acknowledge the financial support of the Research Council of Norway through Yggdrasil project no. 210897 / F11.

References: 1. Loeblein C, Perkins JD. Structural Design for On-Line Process Optimization: I . Dynamic Economics of MPC. *AIChE Journal*. 1999;45(5):1018-1029.

2. Peng JK, Manthanwar AM, Chmielewski DJ. On the tuning of predictive controllers: The minimum back-off operating point selection problem. *Industrial & engineering chemistry research*. 2005;44(20):7814–7822.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Model Predictive Control of a Nonlinear System with Known Scheduling Variable

Mahmood Mirzaei * **Niels Kjølstad Poulsen** *
Hans Henrik Niemann **

* *Department of Informatics and Mathematical Modeling, Technical University of Denmark, Denmark, (e-mail: mmir@imm.dtu.dk, nkp@imm.dtu.dk).*
** *Department of Electrical Engineering, Technical University of Denmark, Denmark, (e-mail: hhn@elektro.dtu.dk)*

Abstract: Model predictive control (MPC) of a class of nonlinear systems is considered in this paper. We will use Linear Parameter Varying (LPV) model of the nonlinear system. By taking the advantage of having future values of the scheduling variable, we will simplify state prediction. Consequently the control problem of the nonlinear system is simplified into a quadratic programming. Wind turbine is chosen as the case study and we choose wind speed as the scheduling variable. Wind speed is measurable ahead of the turbine, therefore the scheduling variable is known for the entire prediction horizon.

Keywords: Model predictive control, linear parameter varying, nonlinear systems, wind turbines, LIDAR measurements.

1. INTRODUCTION

Model predictive control (MPC) has been an active area of research and has been successfully applied on different applications in the last decades (Qin and Badgwell (1996)). The reason for its success is its straightforward ability to handle constraints. Moreover it can employ feedforward measurements in its formulation and can easily be extended to MIMO systems. However the main drawback of MPC was its on-line computational complexity which kept its application to systems with relatively slow dynamics for a while. Fortunately with the rapid progress of fast computations, better optimization algorithms, off-line computations using multi-parametric programming (Baotic (2005)) and dedicated algorithms and hardware, its applications have been extended to even very fast dynamical systems such as DC-DC converters (Geyer (2005)). Basically MPC uses a *model* of the plant to *predict* its future behavior in order to compute appropriate control signals to *control* outputs/states of the plant. To do so, at each sample time MPC uses the current measurement of outputs/states and solves an optimization problem. The result of the optimization problem is a sequence of control inputs of which only the first element is applied to the plant and the procedure is repeated at the next sample time with new measurements (Maciejowski (2002)). This approach is called receding horizon control. Therefore basic elements of MPC are: a model of the plant to predict its future, a cost function which reflects control objectives, constraints on inputs and states/outputs, an optimization algorithm and the receding horizon principle. Depending on the type of the model, the control problem is called linear MPC, hybrid MPC, nonlinear MPC etc. Nonlinear MPC is normally computationally very expensive and generally there is no guarantee that the solution of the optimization problem is a global optimum. In this work we extend the idea of linear MPC using linear parameter varying (LPV) systems to formulate a tractable predictive control of nonlinear systems. To do so, we use future values of a disturbance to the system that acts as a scheduling variable in the model. However there are some assumptions that restrict our solution to a specific class of problems. The scheduling variable is assumed to be known for the entire prediction horizon. And the operating point of the system mainly depends on the scheduling variable.

⋆ This work is supported by the CASED Project funded by grant DSF-09- 063197 of the Danish Council for Strategic Research.

2. PROPOSED METHOD

Generally the nonlinear dynamics of a plant could be modeled as the following difference equation:

$$x_{k+1} = f(x_k, u_k, d_k) \tag{1}$$

With x_k, u_k and d_k as states, inputs and disturbances respectively. Using the nonlinear model, the nonlinear MPC problem could be formulated as:

$$\min_{u} \quad \ell(x_N) + \sum_{i=0}^{N-1} \ell(x_{k+i|k}, u_{k+i|k}) \tag{2}$$

$$\text{Subject to} \quad x_{k+1} = f(x_k, u_k, d_k) \tag{3}$$

$$u_{k+i|k} \in \mathbb{U} \tag{4}$$

$$\hat{x}_{k+i|k} \in \mathbb{X} \tag{5}$$

Where ℓ denotes some arbitrary norm and $\mathbb{U}$ and $\mathbb{X}$ show the set of acceptable inputs and states. As it was mentioned because of the nonlinear model, this problem is computationally too expensive. One way to avoid this problem is to linearize around an equilibrium point of the system and use linearized model instead of the nonlinear model. However for some plants assumption of linear

model does not hold for long prediction horizons as the plant operating point changes, for example based on some disturbances that act as a scheduling variable. An example could be a wind turbine for which wind speed acts as a scheduling variable and changes the operating point of the system.

2.1 Linear MPC formulation

The problem of linear MPC could be formulated as:

$$\min_{u_0,u_1,\ldots,u_{N-1}} \|x_N\|_{Q_f} + \sum_{i=0}^{N-1} \|x_{k+i|k}\|_Q + \|u_{k+i|k}\|_R \quad (6)$$

$$\text{Subject to} \quad x_{k+1} = Ax_k + Bu_k \quad (7)$$

$$u_{k+i|k} \in \mathbb{U} \quad (8)$$

$$\hat{x}_{k+i|k} \in \mathbb{X} \quad (9)$$

Assuming that we use norms $1, 2$ and ∞ the optimization problem becomes convex providing that the sets $\mathbb{U}$ and $\mathbb{X}$ are convex. Convexity of the optimization problem makes it tractable and guarantees that the solution is the global optimum. The problem above is based on a single linear model of the plant around one operating point. However below we formulate our problem using linear parameter varying systems (LPV) in which the scheduling variable is known for the entire prediction horizon.

2.2 Linear Parameter Varying systems

Linear Parameter Varying (LPV) systems are a class of linear systems whose parameters change based on a scheduling variable. Study of LPV systems was motivated by their use in gain-scheduling control of nonlinear systems (Apkarian et al. (1995)). LPV systems are able to handle changes in the dynamics of the system by parameter varying matrices.

Definition (LPV systems) let $k \in Z$ denote discrete time. We define the following LPV systems:

$$x_{k+1} = A(\gamma_k)x_k + B(\gamma_k)u_k \quad (10)$$

$$A(\gamma_k) = \sum_{j=1}^{n_\gamma} A_j\gamma_{k,j} \quad B(\gamma_k) = \sum_{j=1}^{n_\gamma} B_j\gamma_{k,j} \quad (11)$$

Which $A(\gamma_k)$ and $B(\gamma_k)$ are functions of the scheduling variable γ_k. The variables $x_k \in \mathbb{R}^{n_x}, u_k \in \mathbb{R}^{n_u}$, and $\gamma_k \in \mathbb{R}^{n_\gamma}$ are the state, the control input and the scheduling variable respectively.

2.3 Problem formulation

Using the above definition, the linear parameter varying (LPV) model of the nonlinear system with disturbances is of the following form:

$$\tilde{x}_{k+1} = A(\gamma_k)\tilde{x}_k + B(\gamma_k)\tilde{u}_k + B_d(\gamma_k)\tilde{d}_k \quad (12)$$

This model is formulated based on deviations from the operating point. However we need the model to be formulated in absolute values of inputs, states and disturbances. Because in our problem the steady state point changes as a function of the scheduling variable and we need to introduce a variable to capture its bahavior. In order to rewrite the state space model in the absolute form we use:

$$\tilde{x}_k = x_k - x_k^* \quad (13)$$

$$\tilde{u}_k = u_k - u_k^* \quad (14)$$

$$\tilde{d}_k = d_k - d_k^* \quad (15)$$

x_k^*, u_k^* and d_k^* are values of states, inputs and disturbances at the operating point. Therefore the LPV model becomes:

$$\begin{aligned} x_{k+1} = A(\gamma_k)(x_k - x_k^*) + B(\gamma_k)(u_k - u_k^*) \\ + B_d(\gamma_k)(d_k - d_k^*) + x_{k+1}^* \end{aligned} \quad (16)$$

Which could be written as:

$$x_{k+1} = A(\gamma_k)x_k + B(\gamma_k)u_k + B_d(\gamma_k)d_k + \lambda_k \quad (17)$$

with

$$\lambda_k = x_{k+1}^* - A(\gamma_k)x_k^* - B(\gamma_k)u_k^* - B_d(\gamma_k)d_k^* \quad (18)$$

Now having the LPV model of the system we proceed to compute state predictions. In linear MPC predicted states at step n is:

$$\begin{aligned} x_{k+n} = A^n x_k + \sum_{i=0}^{n-1} A^i B u_{k+(n-1)-i} \\ \text{for} \quad n = 1, 2, \ldots, N \end{aligned} \quad (19)$$

However in our method the predicted state is also a function of scheduling variable $\Gamma_n = (\gamma_{k+1}, \gamma_{k+2}, \ldots \gamma_{k+n})^T$ for $n = 1, 2, \ldots, N-1$ and we assume that the scheduling variable is known for the entire prediction. Therefore the predicted state could be written as:

$$x_{k+1}(\gamma_k) = A(\gamma_k)x_k + B(\gamma_k)u_k + B_d(\gamma_k)d_k + \lambda_k \quad (20)$$

And for $n \in \mathbb{Z}, n \geq 1$:

$$\begin{aligned} x_{k+n+1}(\Gamma_n) = &\prod_{i=n}^{0} A(\gamma_{k+i})x_k \\ &+ \sum_{j=0}^{n-1} \left(\prod_{i=n-j}^{1} A(\gamma_{k+i}) \right) B(\gamma_{k+j})u_{k+j} \\ &+ \sum_{j=0}^{n-1} \left(\prod_{i=n-j}^{1} A(\gamma_{k+i}) \right) B_d(\gamma_{k+j})d_{k+j} \\ &+ \sum_{j=0}^{n-1} \left(\prod_{i=n-j}^{0} A(\gamma_{k+i}) \right) \lambda_{k+(n-1)-j} \\ &+ B(\gamma_{k+n})u_{k+n} + B_d(\gamma_{k+n})d_{k+n} + \lambda_{k+n} \end{aligned} \quad (21)$$

Using the above formulas we write down the stacked predicted states which becomes:

$$X = \Phi(\Gamma)x_k + \mathcal{H}_u(\Gamma)U + \mathcal{H}_d(\Gamma)D + \Phi_\lambda(\Gamma)\Lambda \quad (22)$$

with

$$X = (x_{k+1} \; x_{k+2} \; \ldots \; x_{k+N})^T \quad (23)$$

$$U = (u_k \; u_{k+1} \; \ldots \; u_{k+N-1})^T \quad (24)$$

$$D = (d_k \; d_{k+1} \; \ldots \; d_{k+N-1})^T \quad (25)$$

$$\Gamma = (\gamma_k \; \gamma_{k+1} \; \ldots \; \gamma_{k+N-1})^T \quad (26)$$

$$\Lambda = (\lambda_k \; \lambda_{k+1} \; \ldots \; \lambda_{k+N-1})^T \quad (27)$$

In order to summarize formulas for matrices $\Phi, \Phi_\lambda, \mathcal{H}_u$ and $\mathcal{H}_d$, we define a new function as:

$$\psi(m, n) = \prod_{i=m}^{n} A(\gamma_{k+i}) \quad (28)$$

Therefore the matrices become:

$$\Phi(\Gamma) = \begin{pmatrix} \psi(1,1) \\ \psi(2,1) \\ \psi(3,1) \\ \vdots \\ \psi(N,1) \end{pmatrix}$$

$$\Phi_\lambda(\Gamma) = \begin{pmatrix} I & 0 & 0 & \dots & 0 \\ \psi(1,1) & I & 0 & \dots & 0 \\ \psi(2,1) & \psi(2,2) & I & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \psi(N-1,1) & \psi(N-1,2) & \psi(N-1,3) & \dots & I \end{pmatrix}$$

$$\mathcal{H}_u(\Gamma) = \begin{pmatrix} B(\gamma_k) & 0 & \dots & 0 \\ \psi(1,1)B(\gamma_k) & B(\gamma_{k+1}) & \dots & 0 \\ \psi(2,1)B(\gamma_k) & \psi(2,2)B(\gamma_{k+1}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \psi(N-1,1)B(\gamma_k) & \psi(N-1,2)B(\gamma_{k+1}) & \dots & B(\gamma_{N-1}) \end{pmatrix}$$

$$\mathcal{H}_d(\Gamma) = \begin{pmatrix} B_d(\gamma_k) & 0 & \dots & 0 \\ \psi(1,1)B_d(\gamma_k) & B_d(\gamma_{k+1}) & \dots & 0 \\ \psi(2,1)B_d(\gamma_k) & \psi(2,2)B_d(\gamma_{k+1}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \psi(N-1,1)B_d(\gamma_k) & \psi(N-1,2)B_d(\gamma_{k+1}) & \dots & B_d(\gamma_{N-1}) \end{pmatrix}$$

After computing the state predictions as functions of control inputs (22), we can write down the optimization problem similar to a linear MPC problem as a quadratic program:

$$\begin{aligned} \min_{U} \quad & X^T Q X + U^T R U \\ \text{Subject to:} \quad & U \in \mathbb{U} \\ & X \in \mathbb{X} \end{aligned} \tag{29}$$

3. CASE STUDY

The case study here is a wind turbine. Wind turbine control is a challenging problem as the dynamics of the system changes based on wind speed which has a stochastic nature. The method that we propose here is to use wind speed as a scheduling variable. With the advances in LIDAR technology (Harris et al. (2006)) it is possible to measure wind speed ahead of the turbine and this enables us to have the scheduling variable of the plant for the entire prediction horizon.

3.1 Modeling

Nonlinear model For modeling purposes, the whole wind turbine can be divided into 4 subsystems: Aerodynamics subsystem, mechanical subsystem, electrical subsystem and actuator subsystem. The aerodynamic subsystem converts wind forces into mechanical torque and thrust on the rotor. The mechanical subsystem consists of drivetrain, tower and blades. Drivetrain transfers rotor torque to electrical generator. Tower holds the nacelle and withstands the thrust force. And blades transform wind forces into toque and thrust. The generator subsystem converts mechanical energy to electrical energy and finally the blade-pitch and generator-torque actuator subsystems are part of the control system. To model the whole wind turbine, models of these subsystems are obtained and at the end they are connected together. A wind model is obtained and augmented with the wind turbine model to be used for wind speed estimation. Figure 1 shows the basic subsystems and their interactions. The dominant dynamics of the wind turbine come from its flexible structure. Several degrees of freedom could be considered to model the flexible structure, but for control design mostly just a few important degrees of freedom are considered. In figure 2 basic degrees of freedom which are normally being considered in the design model are shown. However in this work we only consider two degrees of freedom, namely the rotational degree of freedom (DOF) and drivetrain torsion. Nonlinearity of the wind turbines mostly comes from its aerodynamics. Blade element momentum (BEM) theory (Hansen (2008)) is used to numerically calculate aerodynamic torque and thrust on the wind turbine. This theory explains how torque and thrust are related to wind speed, blade pitch angle and rotational speed of the rotor. In steady state, i.e. disregarding dynamic inflow, the following formulas can be used to calculate aerodynamic torque and thrust.

$$Q_r = \frac{1}{2}\frac{1}{\omega_r}\rho\pi R^2 v_e^3 C_p(\theta, \omega, v_e) \tag{30}$$

$$Q_t = \frac{1}{2}\rho\pi R^2 v_e^2 C_t(\theta, \omega, v_e) \tag{31}$$

In which Q_r and Q_t are aerodynamic torque and thrust, ρ is the air density, ω_r is the rotor rotational speed, v_e is the effective wind speed, C_p is the power coefficient and C_t is the thrust force coefficient. The absolute angular position of the rotor and generator are of no interest to us, therefore we use $\psi = \theta_r - \theta_g$ instead which is the drivetrain torsion. Having aerodynamic torque and modeling drivetrain with a simple mass-spring-damper, the whole system equation with 2 degrees of freedom becomes:

$$J_r\dot{\omega}_r = Q_r - c(\omega_r - \frac{\omega_g}{N_g}) - k\psi \tag{32}$$

$$(N_g J_g)\dot{\omega}_g = c(\omega_r - \frac{\omega_g}{N_g}) + k\psi - N_g Q_g \tag{33}$$

$$\dot{\psi} = \omega_r - \frac{\omega_g}{N_g} \tag{34}$$

$$P_e = Q_g\omega_g \tag{35}$$

In which J_r and J_g are rotor and generator moments of inertia, ψ is the drivetrain torsion, c and k are the drivetrain damping and stiffness factors respectively lumped in the low speed side of the shaft and P_e is the generated electrical power. For numerical values of these parameters and other parameters given in this paper, we refer to (Jonkman et al. (2009)).

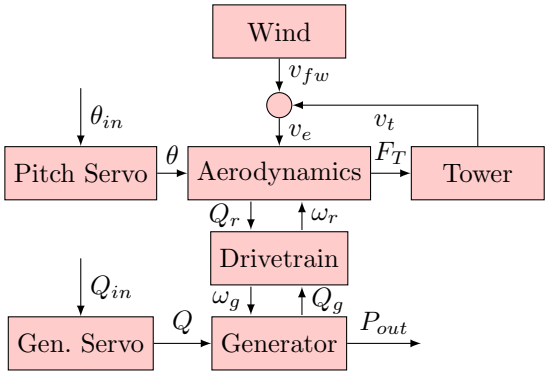


Fig. 1. Wind turbine subsystems

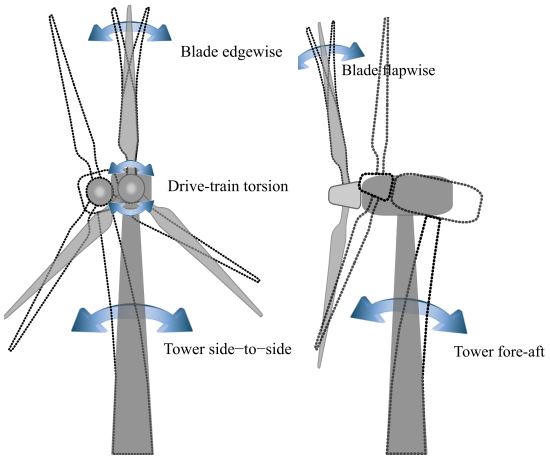


Fig. 2. Basic degrees of freedom

Linearized model As it was mentioned in the previous section, wind turbines are nonlinear systems. A basic approach to design controllers for nonlinear systems is to linearize them around some operating points. For a wind turbine, the operating points on the quasi-steady C_p and C_t curves are nonlinear functions of rotational speed ω_r, blade pitch θ and wind speed v. To get a linear model of the system we need to linearize around these operating points. Rotational speed and blade pitch are measurable with enough accuracy, however this is not the case for the effect of wind on the rotor. Wind speed changes along the blades and with azimuth angle (angular position) of the rotor. This is because of wind shear and tower shadow and stochastic spatial distribution of the wind field. Therefore a single wind speed does not exist to be used and measured for finding the operating point. We bypass this problem by defining a fictitious variable called effective wind speed (v_e) which shows the effect of wind in the rotor disc on the wind turbine. In our two DOFs model only the aerodynamic torque (Q_r) and electric power (P_e) are nonlinear. Taylor expansion is used to linearize them.

$$\Delta Q_r(\omega,\theta,v_e) = \underbrace{\frac{\partial Q_r}{\partial \omega}}_{a}\Delta\omega + \underbrace{\frac{\partial Q_r}{\partial \theta}}_{b_1}\Delta\theta + \underbrace{\frac{\partial Q_r}{\partial v_e}}_{b_2}\Delta v_e \quad (36)$$

$$\Delta P_e = \underbrace{\frac{\partial P_e}{\partial \omega_g}}_{Q_{g_0}}\Delta\omega_g + \underbrace{\frac{\partial P_e}{\partial Q_g}}_{\omega_{g_0}}\Delta Q_g \quad (37)$$

For the sake of simplicity in notations we use Q_r, P_e, θ, ω and v_e instead of ΔQ_r, ΔP_e, $\Delta\theta$, $\Delta\omega$ and Δv_e around the operating points from now on. Using the linearized aerodynamic torque, the 2 DOFs linearized model becomes:

$$\dot{\omega}_r = \frac{a-c}{J_r}\omega_r + \frac{c}{J_r}\omega_g - \frac{k}{J_r}\psi + b_1\theta + b_2 v_e \quad (38)$$

$$\dot{\omega}_g = \frac{c}{N_g J_g}\omega_r - \frac{c}{N_g^2 J_g}\omega_g + \frac{k}{N_g J_g}\psi - \frac{Q_g}{J_g} \quad (39)$$

$$\dot{\psi} = \omega_r - \frac{\omega_g}{N_g} \quad (40)$$

$$P_e = Q_{g_0}\omega_g + \omega_{g_0} Q_g \quad (41)$$

A more detailed description of the model and linearization is given in (Mirzaei et al. (2011)).

LPV model Collecting all the discussed models, matrices of the state space model become:

$$A(\gamma) = \begin{pmatrix} \frac{a(\gamma)-c}{J_r} & \frac{c}{J_r} & -\frac{k}{J_r} \\ \frac{c}{N_g J_g} & -\frac{c}{N_g^2 J_g} & \frac{k}{N_g J_g} \\ 1 & -1 & 0 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & Q_{g_0} & 0 \end{pmatrix} \quad (42)$$

$$B(\gamma) = \begin{pmatrix} b_1(\gamma) & 0 \\ 0 & -\frac{1}{J_g} \\ 0 & 0 \end{pmatrix} \quad D = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & \omega_{g_0} \end{pmatrix} \quad (43)$$

In which $x = (\omega_r\ \omega_g\ \psi)^T$, $u = (\theta\ Q_g)^T$ and $y = (\omega_r\ \omega_g\ P_e)^T$ are states, inputs and outputs respectively. In the matrix B, parameter b_1 is uncertain. Therefore the uncertain linear state space model becomes:

$$\dot{x} = A(\gamma)x + B(\gamma)u$$
$$y = Cx + Du$$

3.2 Control objectives

The most basic control objective of a wind turbine is to maximize captured power during the life time of the wind turbine. This means trying to maximize captured power when wind speed is below its rated value. This is also called maximum power point tracking (MPPT). However when wind speed is above rated, control objective becomes regulation of the outputs around their rated values while trying to minimize dynamic loads on the structure. These objectives should be achieved against fluctuations in wind speed which acts as a disturbance to the system. In this work we have considered operation of the wind turbine in above rated (full load region). Therefore we try to regulate rotational speed and generated power around their rated values and remove the effect of wind speed fluctuations.

3.3 Offset free control

Persistent disturbances and modeling error can cause an offset between measured outputs and desired outputs. To avoid this problem we have employed an offset free reference tracking approach (see Muske and Badgwell (2002) and Pannocchia and Rawlings (2003)). Our RMPC solves the regulation problem around the operating point. However we regulate around the operating point (x_k^* and u_k^*) which results in offset from desired outputs. To avoid this problem in our control algorithm we shift origin in our regulation problem to x_k^0 and u_k^0 instead. In order to find new origins, we have augmented linear model of the plant with a disturbance model that adds fictitious disturbances to the system. The fictitious disturbances compensate the difference between measured outputs and desired outputs. State space model of the augmented system is:

$$\tilde{x}_{k+1} = \tilde{A}\tilde{x}_k + \tilde{B}u_k \quad (44)$$

$$y_k = \tilde{C}\tilde{x}_k + Du_k \quad (45)$$

in which the augmented state and matrices are:

Table 1. Performance comparison between gain scheduling approach and linear MPC

Parameters	Proposed approach	Linear MPC
SD of ω_r (RPM)	0.111	0.212
SD of P_e (Watts)	4.686×10^4	8.048×10^4
Mean value of P_e (Watts)	4.998×10^6	4.998×10^6
SD of pitch (degrees)	2.67	2.95
SD of shaft moment (N.M.)	256	293

$$\tilde{x}_k = \begin{pmatrix} \hat{x}_{k+1} \\ \hat{d}_{k+1} \\ \hat{p}_{k+1} \end{pmatrix} \qquad \tilde{A} = \begin{pmatrix} A & B_d & 0 \\ 0 & A_d & 0 \\ 0 & 0 & A_p \end{pmatrix} \tag{46}$$

$$\tilde{B} = (B \ 0 \ 0)^T \qquad \tilde{C} = (C \ 0 \ C_p) \tag{47}$$

$\hat{x}_k, \hat{d}_k$ and $\hat{p}_k$ are system states, input/state and output disturbances respectively. (A, B, C, D) are matrices of the linearized model, B_d and C_p show effect of disturbances on states and outputs respectively. A_d and A_p show dynamics of input/state and output disturbances. For more information and how to choose these matrices we refer to (Muske and Badgwell (2002)) and (Pannocchia and Rawlings (2003)). Since the disturbances are not measurable, an extended Kalman filter is designed to estimate them. The estimated disturbances are used to remove any offset between desired outputs and measured outputs. Based on this model and estimated disturbances, x_k^0 and u_k^0 which are offset free steady state input and states can be calculated:

$$\begin{pmatrix} A - I & B \\ C & D \end{pmatrix} \begin{pmatrix} x_k^0 \\ u_k^0 \end{pmatrix} = \begin{pmatrix} -B_d \hat{d}_k \\ -C_p \hat{p}_k \end{pmatrix} \tag{48}$$

After calculating these values, we simply replace x_k^* and u_k^* in (18) with x_k^0 and u_k^0 which results in:

$$\lambda_k = x_{k+1}^0 - A(\gamma_k)x_k^0 - B(\gamma_k)u_k^0 - B_d(\gamma_k)d_k^* \tag{49}$$

4. SIMULATIONS

In this section simulation results for the obtained controller are presented. The controller is implemented in MATLAB and is tested on a full complexity FAST (Jonkman and Jr. (2005)) model of the reference wind turbine (Jonkman et al. (2009)). Simulations are done with realistic turbulent wind speed, with Kaimal model (iec (2005)) as the turbulence model and TurbSim (Jonkman (2009)) is used to generate wind profile. In order to stay in the full load region, a realization of turbulent wind speed is used from category C of the turbulence categories of the IEC 61400-1 (iec (2005)) with $18m/s$ as the mean wind speed.

4.1 Stochastic simulations

In this section simulation results for a stochastic wind speed is presented. Control inputs which are pitch reference θ_{in} and generator reaction torque reference Q_{in} along with system outputs which are rotor rotational speed ω_r and electrical power P_e are plotted in figures 3-6 (red-dashed lines are results of linear MPC and solid blue lines show the results of the proposed approach.) Simulation results show good regulations of generated power and rotational speed. Table 1 shows a comparison of the results between the proposed approach and MPC approach based on linearization at each sample point (Henriksen (2007)).

As it could be seen from the table and figures, the proposed approach gives better regulation on rotational speed and generated power (smaller standard deviations) while maintaining a smaller shaft moment and pitch activity.

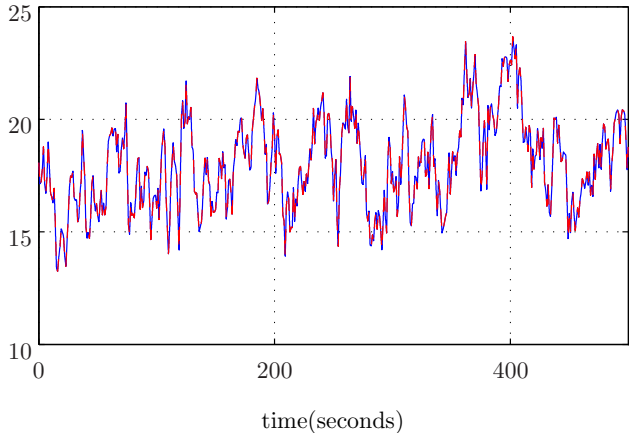


Fig. 3. Blade-pitch reference (degrees, red-dashed line is linear MPC and solid blue line is the proposed approach)

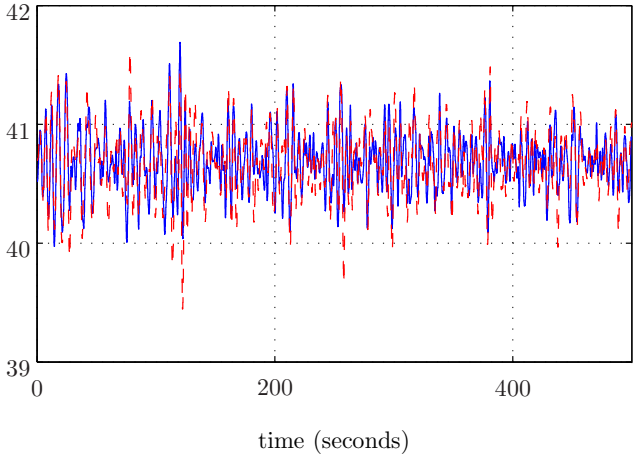


Fig. 4. Generator-torque reference (kNM, red-dashed line is linear MPC and solid blue line is the proposed approach)

REFERENCES

(2005). IEC 61400-1 wind turbines-part 1: Design requirements..

Apkarian, P., Gahinet, P., and Becker, G. (1995). Self-scheduled h_∞ control of linear parameter-varying systems: a design example. *Automatica*, 31(9), 1251–1261.

Baotic, M. (2005). *Optimal Control of Piecewise Affine Systems – a Multi-parametric Approach*. Ph.D. thesis.

Geyer, T. (2005). *Low Complexity Model Predictive Control in Power Electronics and Power Systems*. Ph.D. thesis.

Hansen, M.O.L. (2008). *Aerodynamics of Wind Turbines*. Earthscan.

Harris, M., Hand, M., and Wright, A. (2006). LIDAR for turbine control. Technical report, National Renewable Energy Laboratory, Golden, CO.

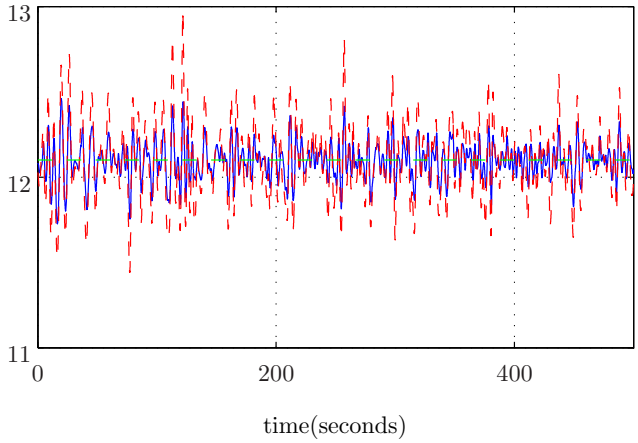


Fig. 5. Rotor rotational speed (ω_r, rpm, red-dashed line is linear MPC and solid blue line is the proposed approach)

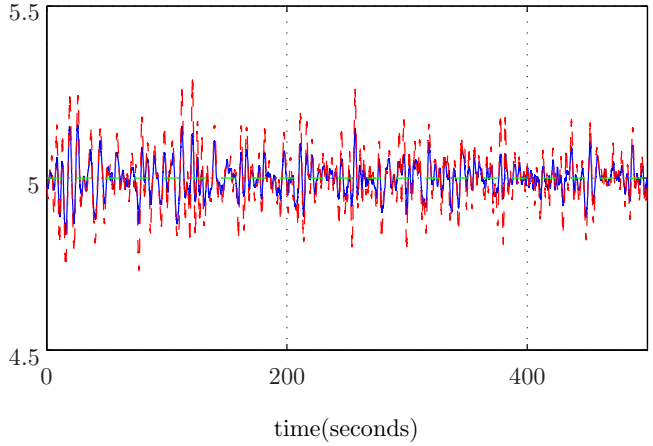


Fig. 6. Electrical power (mega watts, red-dashed line is linear MPC and solid blue line is the proposed approach)

Henriksen, L.C. (2007). *Model Predictive Control of a Wind Turbine*. Master's thesis, Technical University of Denmark, Informatics and Mathematical Modelling, Lyngby, Denmark.

Jonkman, B. (2009). Turbsim user's guide: Version 1.50. Technical report, National Renewable Energy Laboratory, 1617 Cole Boulevard, Golden, Colorado 80401-3393 303-275-3000.

Jonkman, J., Butterfield, S., Musial, W., and Scott, G. (2009). Definition of a 5MW reference wind turbine for offshore system development. Technical report, National Renewable Energy Laboratory, Golden, CO.

Jonkman, J.M. and Jr., M.L.B. (2005). Fast users guide. Technical Report NREL/EL-500-38230, National Renewable Energy Laboratory, Golden, CO.

Maciejowski, J. (2002). *Predictive control with constraints*. Pearson Education Lim., Essex.

Mirzaei, M., Niemann, H.H., and Poulsen, N.K. (2011). A μ-synthesis approach to robust control of a wind turbine. In *the 50th IEEE Conference on Decision and Control and European Control Conference*. Orlando, FL, USA.

Muske, K.R. and Badgwell, T.A. (2002). Disturbance modeling for offset-free linear model predictive control. *Journal of Process Control*, 12(5), 617–632.

Pannocchia, G. and Rawlings, J.B. (2003). Disturbance models for offset-free model-predictive control. *AIChE Journal*, 49(2), 426–437.

Qin, S.J. and Badgwell, T.A. (1996). An overview of industrial model predictive control technology. –.

Modelica library for simulation of bioprocesses

Jan Peter Axelsson, senior consultant,
Vascaia AB, 121 77 Stockholm, Sweden.

Developing models for bioprocesses requires a combination of knowledge from different fields like reactor dynamics, liquid-gas transfer, cell metabolism etc. Part of this knowledge is well established and can be re-used, while modelling of cell metabolism or product formation may be more unique. Further the same reactor may be operated in different ways: batch, fed-batch, continuous, or using perfusion. Finally it is of interest to evaluate various configurations of control systems.

Here a small library in Modelica is developed for simulation of bioprocesses based on text book know-how and simplified models of microbial cell growth and metabolism [1] and [2]. Despite its simplicity it is rich enough to require some thought on interfaces and connnections between different parts to facilitate re-use of models. The presentation will focus on this modelling or structuring aspectes. The library enables the user to configure systems based on well-proven components on a high-level.

The implementation of the library is made in JModelica which incorporates optimization to the Modelica language [3]. The optimization tools is used in two ways. In one example a certain critical parameter is estimated in the microbial model during an experiment in a continuous reactor, see Figure 1. In another example the same culture is run in fed-batch mode and the optimal feed-profile is determined.

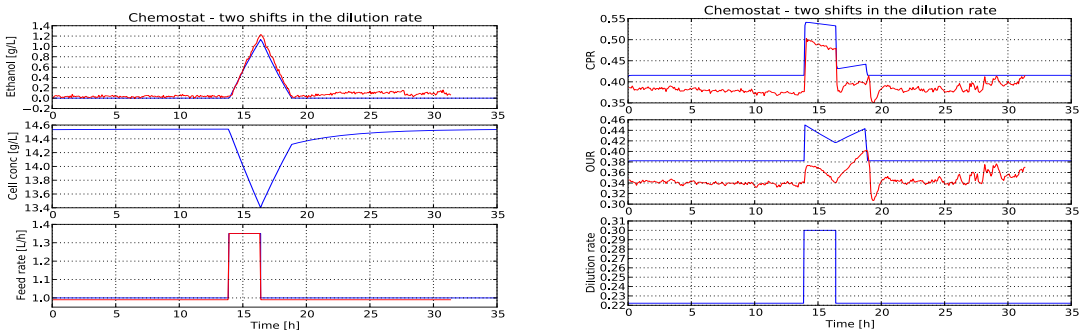


Figure 1. Shift-up/down of flow rate in continuous cultivation of yeast (noisy line) in comparison with simulation (solid line), see [1].

References:

[1] Axelsson, J. P., A. Hagman, and J. Mari, ”The Role of Simulation during Development of a Recombinant Yeast Process for Production of Human Growth Factors”, SIMS’94 Applied Simulation in Industry, Stockholm, 1994.

[2] Akesson, M., E. Nordberg Karlsson, P. Hagander, J. P. Axelsson, A. Tocaj, ”On-line Detection of Acetate Formation in Escherichia coli Cultures Using Dissolved Oxygen Responses to Feed Transients”, Biotechnol. Bioeng, 2001.

[3] Akesson J., K-E Årzén, M. Gäfvert, T. Bergdahl, H. Tummescheit, ”Modeling and OPtimization with Optimica and JModelica.org – Languages and Tools for Solving Larbe-Scale Dynamic Optimization Problems, Computers & Chemical Engineering, 2009.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Modeling and simulating an electrical grid subsystem for power balance analysis ⋆

Luminita C. Totu, John Leth, Rafael Wisniewski

Department of Electronic Systems, Aalborg University
(e-mail: lct,jjl,raf@es.aau.dk)

Abstract: We present an approach for power balance analysis in Smart Grids where the physical behavior of different electrical devices is modeled at unit level, and the collective load and generation curves can later be obtained by aggregation. In this way, new behaviors, flexibilities and intelligent strategies for power consumption and generation can be easily introduced at the user-level and the system-level impact analyzed on the aggregated profiles. The future aim is to investigate bottom-up balancing strategies, where units with a flexible energy band can react independently to power balance signals such as dynamic prices.

Keywords: power balance, smart grid, modeling

1. THE GRID SUBSYSTEM

In the following, the term grid subsytem is used to refer to a geographical region of a national electrical network composed of low and medium voltage lines and the end-users, with only one connection to the high voltage transmission system.

The grid subsystem will be modeled as an aggregation of individual electrical units interconnected by the electrical network. Only the active power is considered and the electrical network is simplified to a summation representation (Fig. 1). While clearly incomplete, this simplification captures the first essential property of the electrical power system: power consumption must be met in real-time by the power production.

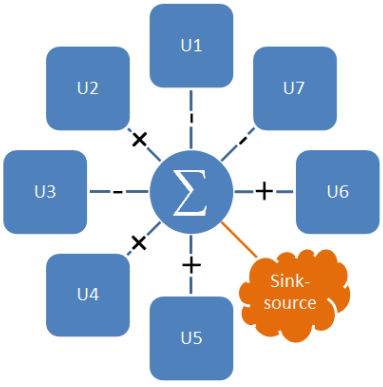


Fig. 1. Representation of a grid subsystem, where each U_i element is an electrical device and the summation element Σ represents the electrical network.

At every time instance, an individual unit connected to the grid is either a producer or a consumer of active power. The outputs of all units, producers (+) and consumers (-), need to be in balance. Whenever the sum is not zero, there is an instantaneous exchange through an external sink-source element that maintains balance. The sink-source has signification of an inter-regional electrical transmission connection. As anticipation, the external power exchange will be used as part of an optimization objective. For example, if the objective is to keep the subsystem independent of the interconnection link, then the optimal power exchange is 0.

Each electrical unit U_i operates independently and is composed of two systems: the service module and the physical device (Fig. 2).

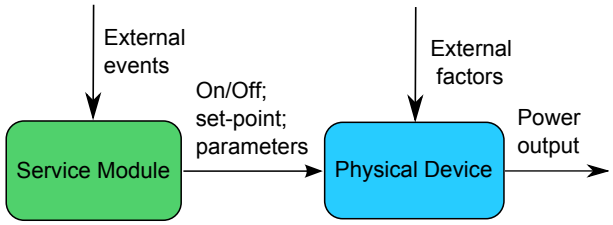


Fig. 2. Each electrical unit is composed of a physical device and a service module.

The physical device performs an energy converting task according to some configuration parameters and references, and operates under disturbances, such as weather and ambient factors. The output of the physical device is the electric power, either produced or consumed. The functionality will generally be modeled by employing first principles from physics, resulting in a mathematical description by differential equations.

The service module operates the "knobs and buttons" of the physical device, such as ON/OFF controls, parameters or reference settings. The service module is subject to external events but can also have internal logic or time plans. Modeling of the service module will include discrete event formalisms and stochastic elements.

⋆ This work is supported by the Southern Denmark Growth Forum and the European Regional Development Fund under the project "Smart & Cool".

In a traditional grid, the power balance is maintained at the global level by a number of large capacity power generators with highly controllable characteristics and which can quickly adapt their power output to new levels. The baseline production and availability of these generators is planed ahead of time based on consumption and generation predictions. In the Smart Grid, the power balance control needs to be distributed in the grid to cope with increasing intermittent generation and increasing peaks in the load, and more end-users of electricity will take an active part in this process.

Some form of coordination is needed among the independent units in the SmartGrid to maintain power balance at the system level. We consider a feedback mechanism that shares a power balance signal from the global level to the local units. In particular, the power flow measured in the sink-source can be shared with the service module (Fig. 3). Flexibility profiles in the service module will make it possible to react to the balance signal within limits of local operation constraints.

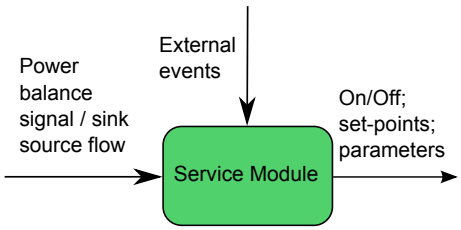


Fig. 3. The service module receives information about the global subsystem.

2. UNIT MODELS

This section presents simplified unit models for both power consuming and power producing devices with intermittent characteristics. It is mentioned that for the balance power problem only slow and mid-range dynamics are considered. Fast dynamics are assumed to be compensated by a grid frequency control scheme that is not discussed here.

2.1 Devices with externally driven ON/OFF states

Many types of electrical units such as lights, TVs, etc., have an ON/OFF operation that is directly driven by external events. This behavior can be represented as a continuous time stochastic process with a discrete state-space of two elements: $x = 1$ is the ON state and $x = 0$ is the OFF state.

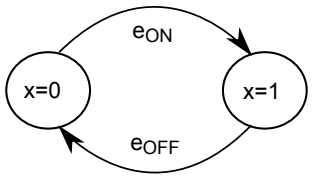


Fig. 4. The ON/OFF behavior

It is the occurrence in time of the events that is of particular interest. It can be considered that the ON and OFF events occur in a purely random manner, but with a varying frequency depending on the time of day. That is because, for example, it is more likely to switch a light in the evening than at midday. Two independent, nonhomogeneous Poisson processes will be used next to represent the arrival of the ON and OFF events. For an introduction to Poisson processes and the more general point process, we refer to Cox and Miller (1965).

Let Δt be a small time interval and $N(a, b)$ a variable counting the number of events occurring in the interval $(a, b]$. Looking in the small time interval Δt, there are three possibilities: no event, one event, or more than one event occurs. The probability of having more than one event occurring is very low, the probability of one event occurring is dependent on rate parameter varying with the time, and the probability remaining until 1 corresponds to the no event case. This is the definition of a nonhomogeneous Poisson process. The notation $o(\Delta t)$ below is used to denote terms which vanish with a small time interval, $\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0$.

$$\begin{cases} \text{prob}\Big(N(t, t+\Delta t) = 1\Big) & = \lambda(t)\Delta t + o(\Delta t) \\ \text{prob}\Big(N(t, t+\Delta t) \geq 2\Big) & = o(\Delta t) \end{cases} \tag{1}$$

We let $\lambda_{ON}(t)$ and $\lambda_{OFF}(t)$ denote the variable rates of the two types of events. These can be taken as piecewise constant functions along the hours of the day, with different profiles for different devices. The ON/OFF mechanism together with the rate functions is enough for simulation purposes.

It is also of immediate interest to derive the probability distribution of the states, i.e. $\text{prob}_{x=1}(t)$, and also to characterize a group of devices in terms of number of units in state $x = 1$ at a given time. These will be addressed in the future work.

About the physical device modeling, when active, $x = 1$, the units included in this category can be considered to consume a constant amount of power although some loads can vary slightly, e.g. a radio for different volume settings, or a computer in idle mode versus while performing intensive computations. In a simple approximation, the power consumption p of a unit U will be written as:

$$p(t) = -cx(t), \tag{2}$$

where c is a device specific constant.

2.2 Space Heating and Cooling devices

This category contains units that are always in operation and the ON/OFF power consumption cycle is decided internally in the service module. Space heating devices operating on electricity, such as heat-pumps, air-conditioning units used for cooling and refrigerators are examples of devices with a switched operation based on thermostat settings.

As example, a simple model for a compressor cooled room is described below based on Tahersima et al. (2010), Halvgaard et al. (2012) and Hovgaard et al. (2010). It can correspond to a storage room that is kept at a lower temperature than the ambient. The physical system is composed of the cold room and the refrigeration system, as shown in Fig. 5. The notation is described in Table 1 below.

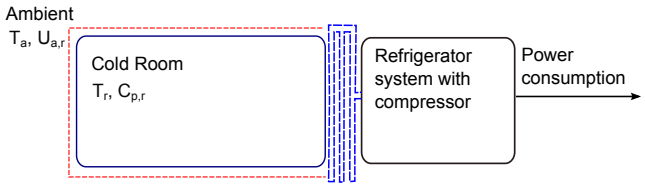


Fig. 5. A cooling system composed of a refrigeration system and a cold room in an ambient environment.

The physical system is described by the thermodynamic energy balance equation on the cold room.

$$C_{p,r}\dot{T}_r = Q_{a,r} - Q_{r,e} \tag{3}$$

The heat transfer between the ambient and the room $Q_{a,r}$ can be written as:

$$Q_{a,r} = \mathrm{UA}_{a,r}(T_a - T_r) \tag{4}$$

Table 1. Notation

T_a(t)	Ambient temperature, an uncontrollable but measurable, time varying value
T_r(t)	Cold room temperature
$C_{p,r}$	Heat capacity (at constant pressure) of the cold room
$Q_{a,r}$	Heat transfer from the ambient to the cold room
$\mathrm{UA}_{a,r}$	Heat tranfer coefficient (thermal resistance) between the ambient and the cold room

The assumptions here are that the temperature T_a and the heat transfer coefficient between the ambient and the room $\mathrm{UA}_{a,r}$ are known. Thus it is possible to calculate the evolution of the temperature in the cold room over time. By combining the equations, we can write the following description of the cool room in continuous-time variable T_r.

$$\dot{T}_r(t) = \frac{\mathrm{UA}_{a,r}}{C_{p,r}}\Big(T_a(t) - T_r(t)\Big) - \frac{Q_{r,e}}{C_{p,r}} \tag{5}$$

The decision logic module is related to the start and stop of the compressor. The ON/OFF behavior from Fig. 4 applies also to the cold room, where the events e_{ON} and e_{OFF} have the triggering mechanism

$$\begin{cases} T_r > T_{ref+} \rightarrow e_{ON} \\ T_r < T_{ref-} \rightarrow e_{OFF} \end{cases}, \tag{6}$$

with T_{ref+} and T_{ref-} temperature thresholds, parameters of the service module. When activated, the power consumption of the compressor is considered constant.

$$p(t) = -c\,x(t) \tag{7}$$

The only information missing is a relation that describes the amount of heat absorbed by the refrigerator system from the cold room $Q_{r,e}$. This relation depends on the properties (performance) of the refrigeration cycle, the compressor state, and a heat transfer constant with the cold room, but is not expanded at this point.

Additionally, a stochastic term can be added to the dynamic description (5) to account for other heat transfer processes occurring in the cold room, for example doors opening and small variations in the overall heat capacity of the room. Such a model, and also results for the aggregated consumption of a groups of similar devices, are considered in Malhame and Chong (1985) and Malhame (1990).

2.3 Wind turbines

There are different types of devices that generate electrical power from wind. In this section we consider the most common wind turbine design, the horizontal axis machines with 3 blades, operating independently. Wind-farms models will be addressed in future work.

Every type of turbine has a static power curve characterization. The power curve relates the electrical output of the turbine to the input wind-speed, as shown in Fig 6.

The power curves are poor predictors for the instantaneous power output as they do not contain any information on the dynamics of the turbine subsystems nor references about the fluctuations of the wind, but they are a good representation as an average behavior over longer periods of time. Power curves will be used here as an approximated model because they have the advantage of being easily available for different turbine models, and can be used for all types of turbine designs.

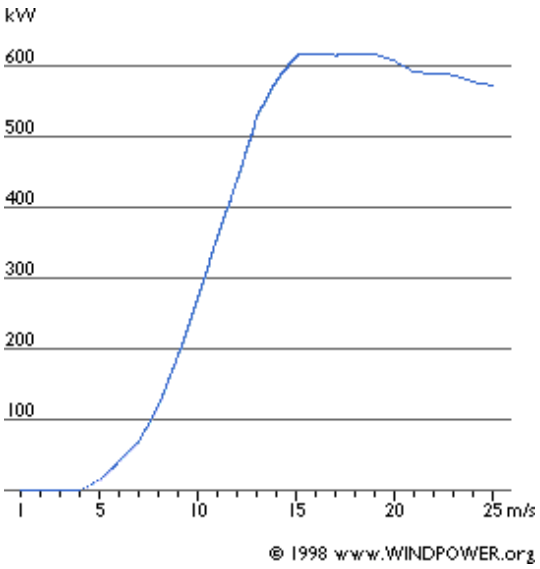


Fig. 6. Power curve for typical Danish 600 kW turbine with stall control, from www.windpower.org

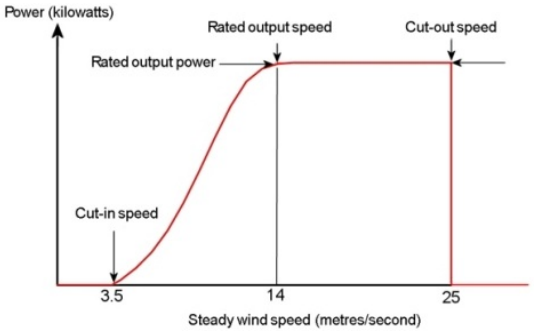


Fig. 7. Typical power output with steady wind speed for a pitched controlled turbine, www.wind-power-program.com

A simple service module for wind turbines is described next. The wind-turbine generator is OFF for wind-speeds

under 5m/s, ON and tracking the power curve for wind-speeds smaller than 23m/s. If the wind-speed exceeds this value, the turbine is stopped. Besides wind-farms, medium size CHPs (combined heat and power plants) and solar panels will be addressed in future work.

3. FUTURE WORK ON SCENARIOS AND SIMULATION

A scenario will contain multiple units of different types, e.g. 2000 light units, 100 ground heat pump devices, and 20 wind turbines. It possible to create a scenario using real data from a geographical region provided sufficient information is available on the number and characteristics of the consumers and producers in the area. Observational data on air temperature, solar irradiation and wind is available from different weather stations in the world, including Denmark. By running software simulations, daily and monthly aggregated load and generation curves are produced. The comparison of aggregated curves from the simulation with those from real data, in terms of main trends, will serve as model validation for the both unit models and the aggregation methods.

The purpose of the scenarios is to test different user-level energy strategies and investigate how the local behaviors scale up and affect the electric power balance of the subsytem.

For example, the temperature thresholds of the refrigerator thermostat can be changed in response to a power balance signal, within the limits permissible for the cold room operation. When the grid subsystem has an excess of power, the service module of the refrigerator system can lower the temperature of the cold room by reducing the threshold values T_{ref-} and T_{ref+}. The compressor will consume more power, but will do so at an advantageous time, when the costs are lower and energy can be stored locally for later use. When there is a deficit of power, the service module can choose to reduce the use of the compressor and increase the threshold temperatures. This simple strategy appears to be "smart" at user level, but needs to be validated against an aggregated scenario, as it can have pitfalls. When more than one devices react in real-time to the same signal, it is possible that the system over-reacts and an imbalance of the opposite sign is created. Also the local effect of the strategy over time can turn out not to be beneficial. Increasing the threshold temperature can be too costly over time if more power needs to be consumed at inappropriate times for recovery.

For wind turbine systems, the service module can run a delta control (nominal underproduction) which allows for a band of flexibility. The turbine will produce less in normal operation, but has the possibility of compensating in power deficit situation. By analyzing scenarios it will be possible to evaluate when this trade-off is sufficient to assure the stability of the subsystem.

REFERENCES

Cox, D., Miller, H., 1965. The Theory of Stochastic Processes. Chapman and Hall, London.

Halvgaard, R., Poulsen, N. K., Madsen, H., Joergensen, J. B., 2012. Economic model predictive control for climate control of buildings in a smart energy system. IEEE PES Conference on Innovative Smart Grid Technologies (ISGT).

Hovgaard, T. G., Edlund, K., Joergensen, J., 2010. The potential of economic mpc for power management. Proc. of 49th IEEE Conference on Decision and Control.

Malhame, R., 1990. A jump-driven markovian electric load model. Advances in Applied Probability.

Malhame, R., Chong, C., 1985. Electric load model synthesis by diffusion approximation of a high-order hybrid-state stochastic system. Automatic Control, IEEE Transactions on.

Tahersima, F., Stoustrup, J., Rasmussen, H., Nielsen, P. G., 2010. Thermal analysis of an hvac system with trv controlled hydronic radiator. Automation Science and Engineering (CASE), 2010 IEEE Conference on.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Modeling of the freezing process for fish in vertical plate freezers

Christoph Backi * **Jan Tommy Gravdahl** **

Norwegian University of Science and Technology
Department of Engineering Cybernetics

* *Christoph.Backi@itk.ntnu.no*
** *Jan.Tommy.Gravdahl@itk.ntnu.no*

Abstract: Energy-efficiency is one of the big issues of the 21st century. Due to the limited resources of primary energy carriers and their environmental load, a higher effectiveness for their use has to be achieved. Even small improvement in models, observers and/or control-strategies can have a large impact on consumption of energy carriers. The system we are looking at is a fishing vessel (trawler). The overall aim is therefore, to reduce energy consumption and, at the same time, preserve or even enhance fish quality.

The vessel is driven by a (diesel-) main engine, which produces electricity. All of the vessel's consumers, such as electric propulsion motors, freezing units, processing units, cooling pumps, ship operation equipment and other facilities are powered by electric energy from the main engine. Besides the propulsion motors, the freezing units are the biggest energy consumers on board. The caught fish shall, after having been processed (blooded, headed and gutted), be frozen as fast as possible. The process, that is used to freeze the fish, is a vapor-compression refrigeration circle process run with ammonia (NH_3). The ammonia flows through the freezer's plates, cools them down very strongly ($-38\,^{\circ}\mathrm{C}$) and due to direct contact, the fish gets frozen.

The model, that is taken for simulating the temperature distribution throughout the fish block with thickness L, is the one dimensional heat-equation, a linear partial differential equation: $\rho(T) \cdot c(T) \cdot \frac{\partial T(t,x)}{\partial t} = \lambda(T) \cdot \frac{\partial^2 T(t,x)}{\partial x^2}$.
It has to be noticed, that the parameters $\rho(T)$, $c(T)$ and $\lambda(T)$ change with temperature. Simplified, the fish can be considered as a thermodynamical alloy of many basic components, such as water/ice, protein, fat, carbohydrates and ash. The parameters can therefore be calculated by $\rho(T) = \sum_i \rho_i(T) \cdot x_i$, $c(T) = \sum_i c_i(T) \cdot x_i$ and $\lambda(T) = \sum_i \lambda_i(T) \cdot x_i$, where $\rho_i(T)$, $c_i(T)$ and $\lambda_i(T)$ correspond to component i and x_i represents the mass fraction of component i.
Note, that not 100% of the water in the fish gets frozen in the temperature range, we are looking at. At $-30\,^{\circ}\mathrm{C}$ only about 90% of the water is frozen. The reason for this is, that, when freezing, solutes remain in the liquid phase, lowering its freezing point. Just at about $-70\,^{\circ}\mathrm{C}$ all the water can be considered as frozen.

A phenomenon, that has to be considered due to the large fraction of water present in fish (60%–80%), is the latent heat of fusion. This means, that there will occur heat transfer at the freezing point without lowering the temperature. This can be explained by energy storage present in the formation of water molecules. This energy has to be removed in order to enable water molecules to nucleate and ice crystals to form and grow.
Ice crystal growth has a big influence on quality of the product. Dependent on the speed of freezing, ice crystal sizes differ. Slow freezing results in large ice crystals and fast freezing in small ice crystals, respectively. Large ice crystals emerge when extracellular freezing happens prior to intracellular freezing, which disrupts the thermodynamical equilibrium. As a consequence, fluid is drawn from the inside of a cell to the outside, causing destruction of cell walls and thus reducing the quality.

In this work, we present a mathematical model for the temperature of fish in vertical plate freezers. The temperature distribution will be linked to a quality measure. This model will be useful when designing control strategies for the on-board energy system and, at the same time, monitoring fish quality.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Modeling Smart Energy Systems for Model Predictive Control

Rasmus Halvgaard * **Niels Kjølstad Poulsen** *
Henrik Madsen * **John Bagterp Jørgensen** *

* *Department of Informatics and Mathematical Modelling, Technical University of Denmark, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, Denmark (e-mail: {rhal,nkp,hm,jbj}@imm.dtu.dk).*

Abstract: Integrating large amounts of renewable energy sources like wind and solar power introduces large fluctuations in the power production. Either this energy must be stored or consumed right away. Storage solutions are very expensive and not applicable everywhere. So utilizing all of this green energy as it is produced requires a very flexible and controllable power consumption. Examples of controllable electric loads are heat pumps in buildings and Electric Vehicles (EVs) that are expected to play a large role in the future danish energy system. These units in a smart energy system can potentially offer flexibility on a time scale ranging from seconds to several days by moving power consumption, exploiting thermal inertia or battery storage capacity, respectively. Using advanced control algorithms these systems are able to reduce their own electricity costs by planning ahead and moving consumption to periods with green and cheap electricity. This situation occurs when there is a lot of excess wind power in the system which is reflected in the electricity price and in turn creates an incentive to absorb the energy.

In this paper a decentralized control strategy is investigated where prices indirectly influence the total power consumption of the smart energy systems connected to the power grid. Compared to a direct control strategy the complexity of the problem is reduced and decreases both the computation efforts and the need for communication. However, not only the current price, but a forecast of the expected future price should also be available in order for the individual units to plan ahead in the most feasible way. This is necessary since Economic MPCs do not respond to the absolute cost of electricity, but to variations of the price over the prediction horizon. Economic MPC is ideal for price responsive units where the model is known very well. Constraints and disturbance forecasts are straight forward to implement in the controller. MPC relies on the receding horizon principle, where a new optimal control signal is calculated at each time step for the prediction horizon. Only the optimal control signal at the current time step is implemented and consequently closed loop feedback is obtained.

A generic model of an energy component is proposed in this paper, so the same Economic MPC framework can be used to design controllers for the different units. However, different signals and forecast, e.g. weather forecasts and usage patterns, are used depending on the unit. The generic state space will be a discrete time state space model with hard input constraints and soft output constraints. For the considered energy systems there is usually a strict limit on the maximum available power, but the output, e.g. a temperature or an EV battery state of charge, can often be relaxed. The output constraints thus define a band of operation, that can be time varying, and the controller must keep the output within these limits in the cheapest possible way.

In this paper the price forecast available by all units is assumed to be known and equal to the day-ahead elspot price from the Nordic electricity exchange market NordPool. The resulting electricity cost savings compared to an MPC with no price considerations are around 30-50% for the chosen units. In future work the price could be replaced by an intrahour price that is related to the deviation between the planned and the actual consumption. In this way all units are motivated to stick to the predicted consumption plan.

Keywords: Economic Model Predictive Control, Smart Grid, Heat pump, Electric Vehicle

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Modelling and control design for SHARON/Anammox reactor sequence

B. Valverde-Pérez*, M. Mauricio-Iglesias* and G. Sin*

** CAPEC, DTU Chemical Engineering, Technical University of Denmark, DK-2800 Lyngby, Denmark, (e-mail:bvape@kt.dtu.dk; mmmi@ kt.dtu.dk; gsi@kt.dtu.dk).*

Abstract: With the perspective of investigating a suitable control design for autotrophic nitrogen removal, this work presents a complete model of the SHARON/Anammox reactor sequence. The dynamics of the reactor were explored pointing out the different scales of the rates in the system: slow microbial metabolism against fast chemical reaction and mass transfer. Likewise, the analysis of the dynamics contributed to establish qualitatively the requirements for control of the reactors, both for regulation and for optimal operation. Work in progress on quantitatively analysing different control structure (pairing of controlled variables with manipulated variables) as well as exploring the feasibility of advanced process control including model predictive control.

Keywords: Autotrophic nitrogen removal, model predictive control, multivariable control, plantwide control, modelling

1. INTRODUCTION

Ammonium is one of the most abundant and important pollutants related with the waste water treatment, and should be removed before discharge, preventing problems for the biological media as eutrophication. As a consequence, ever more strict regulations set low nitrogen concentration thresholds for discharge. The first widely used process used for nitrogen removal consisted on nitrification and denitrification processes over nitrate. However, this process presents some disadvantage, namely high energetic consumption for oxygen supply, high tank volume and the need external carbon sources. A number of alternative operations for nitrogen removal have been developed in the last decades in order to reduce aeration costs, external carbon sources and sludge production the energy consumption, the footprint and the need of using carbon sources.

The SHARON process (Single reactor High activity Ammonia Removal Over Nitrite, Hellinga et al. 1998) was primarily based on the following sequences: i) the partial nitrification of the ammonium by aerobic oxidizing bacteria (AOB) and ii) the denitrification to nitrogen gas by heterotroph bacteria (HB).

Another option to this second step is the so-called anaerobic ammonium oxidation (Anammox, Murder et al.,1995). The process achieves a total ammonium conversion using equimolar amounts of ammonium and nitrite. This process presents some additional advantages, like the lowering of gases with greenhouse effect (CO_2 and NO_2) or the elimination of external carbon sources. Its main drawback is related to the low growth rate of Anammox bacteria, involving the use of sludge retention systems, e.g. membranes, granular systems. Besides, in order to achieve a high elimination of all the nitrogen sources, it must be ensured that the ammonium and nitrite are fed in stable, close to equimolar proportions .

The partial nitrification can be achieved by a previous SHARON reactor coupled whit the Anammox process or both reactions can be implemented in a single reactor (CANON). In this case the SHARON-Anammox process is studied. The partial nitrification step:

$$NH_4^+ + 1.5O_2 \rightarrow NO_2^- + H_2O + 2H^+$$

And the Anammox process follows the next equation:

$$NH_4^+ + NO_2^- \rightarrow N_2 + 2H_2O$$

It has been reported that that with constant parameters, in the supposed corrects values, the ratio nitrite:ammonium might deviate from the ideal ratio and endanger the operation. Some strategies were applied in this field in order to optimize the nitrogen removal costs (Volcke et al. 2005). These strategies are based in different control loops, feedback and cascade, that fix set points for the key variables in the process, as pH or dissolved oxygen (DO), but always operating with them as independent loops. In order to get a better optimization of the global process and advanced control strategy is going to be applied. MPC (model predictive control) use the models of the process and predicts how is going to evolve it, and optimize at each time the set points of the individual control loops. Another objective is establish a procedure to configure and implement a MPC controller in different process, due nowadays there is no unification criteria to use them.

A scheme of the global waste water treatment plant is included in Appendix 2.

2. MODEL DEVELOPMENT

Both SHARON and Anammox reactor were modeled as continuous stirred tank reactors (CSTR). Assuming that the

reactor hold-up and all the inflows and outflows have the same constant density, the total mass and energy balances are:

$$\frac{dV}{dt} = F_{in} + F_{acid} + F_{base} + F_{CH_3OH} - F_{out} = F_{in,net} - F_{out} \quad (1)$$

$$\frac{d(V \cdot T)}{dt} = F_{in} \cdot T_{in} - F_{out} \cdot T_{out} - \frac{U_w \cdot \pi \cdot d_R \cdot H_R \cdot (T_{out} - T_{env})}{\rho \cdot C_P \cdot V} + \sum_j \frac{(-\Delta_R H)_j}{\rho \cdot C_P} \quad (2)$$

Additionally, it has been assumed that the Cp is equal and constant for every stream (2).

The individual mass balances developed are set up for the lumped compounds, because every change in a component involved in a chemical equilibrium causes changes in all the concentrations of all components that take part in the equilibrium. The individual mass balance for a component i is:

$$\frac{d(V \cdot C_i)}{dt} = F_{in,net} \cdot C_{i,in,net} - F_{out} \cdot C_i + k_L a_i \cdot (C_i^* - C_i) \cdot V + r_i \cdot V \quad (3)$$

The components considered are: H+, NH4+, NH3, HNO2, NO2-, CO2, HCO3-, CO32-, H2PO4-, HPO42-, NO3-, O2, N2, ammonia oxidizing bacteria (AOB), nitrite oxidizing bacteria (NOB), heterotrophic bacteria (HB), CH3OH and Z (charge not involved in biological reactions). Since some of the previous chemical species are in chemical equilibrium, the model works with lumped components, namely:

$$C_{TNH} = C_{NH_3} + C_{NH_4^+}$$

$$C_{TNO2} = C_{HNO_2} + C_{NO_2^{2-}}$$

$$C_{TIC} = C_{CO_2} + C_{HCO_3^-} + C_{CO_3^{2-}}$$

$$C_{TIP} = C_{HPO_4^{2-}} + C_{H_2PO_4^-}$$

2.1 Reaction modelling

Five different biological reactions are included in the SHARON model. The nitrification process is divided in two different steps: the oxidation of the ammonia to nitrite, carrying out by AOB, and the oxidation of the nitrite to nitrate, carrying out by NOB. The denitrification of both compounds is achieved by heterotrophic bacteria, as the oxidation of the methanol. The stoichiometric coefficients are showed in the Petersen matrix (tables 1 and 3). In order to take account of the microbial growth in the mass balances, the biomass composition is fixed as CH1.8O0.5N0.2.

The volumetric conversion rate for a component i is defined as:

$$r_i = \sum_{j=1}^{5} A_{ij} \cdot \rho_j \quad (4)$$

where Aij is the corresponding stoichiometric coefficient of the Pedersen matrix, and ρj is the process rate considered. The Pedersen matrix and the expressions of the process rates for the two reactors appear in the Appendix A.

2.2 pH calculation.

The microbial activity affects to the pH because exists a production and consumption of protons. These pH changes must be modeled in order to obtain the correct values during the reactor operation. In this case it is used a model based in a charge balance. The nonlinear system of equations is solved with a multidimensional Newton-Raphson method.

2.3 Gas-liquid transport

Gases are transported between both phases following the model showed below:

$$TR_i = k_L a_i \cdot (C_i^* - C_i) \quad (5)$$

The main gas in the process is the oxygen, and it'll be also a control variable of the process, so is the only modeled gas compound. With the Henry law the oxygen equilibrium concentration is obtained:

$$C_i^* = \frac{C_{G,i}^o}{m_i} \quad (6)$$

$$m_i = -403 + 2.52 \cdot T - 3.56 \cdot 10^{-3} \cdot T^2 \quad (7)$$

3. SIMULATION RESULTS

The models were tested using different cases study. For the SHARON reactor the conditions established in the simulations of A.Galí et al. in 2006 were used. This reactor is used to achieve a suitable influent for the Anammox reactor, with a ratio NO2-/NH4+ close to 1. With a hydraulic retention time of 1 day a removal efficiency of the 50% is achieved.

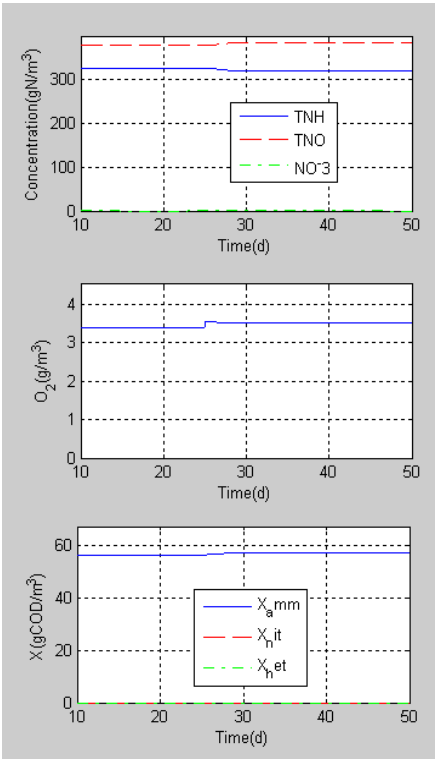


Figure 1: Simulation of $K_L a$ step of 5%

The response to-step inputs were also simulated to investigate the dynamics of the process. The first step consists in an increase of 5% in the $k_L a$ value. A higher growth rate is observed for AOB bacteria. In 5 days a new steady state is achieved. Two phenomena with different rates can be

observed. A peak is observed in the oxygen concentration because at the beginning of the step, corresponding to the higher oxygen mass transfer while the number of oxygen-consumming bacteria has not increased yet. As the bacteria concentration increases more ammonium is converted in nitrite and the concentration of oxygen decreases from the initial response.

With a decrease of 1% in pH the bacteria growth is lower, so they are in less concentration, while the oxygen level achieves higher values and the less ammonium is removed from the reject water. It is needed 11 days to achieve the new steady state.

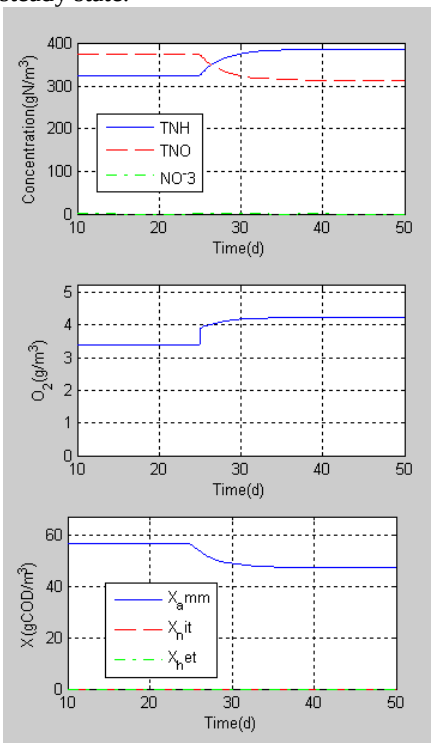


Figure 2: Simulation of pH step of -1%.

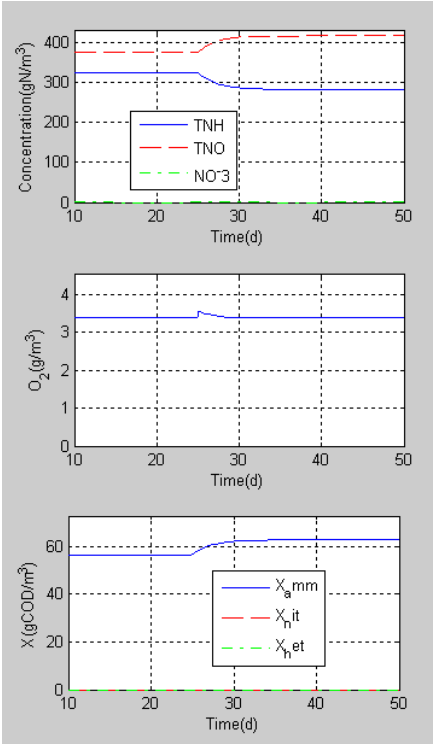


Figure 3: Simulation of HRTstep of +5%.

Finally, the decrease of 5% of the output flow implies an increase of 5% of the hydraulic retention time, which allows a higher microbial growth of AOB. In 8 days the new steady state is achieved, with a decrease of the dissolved oxygen, because there are more bacteria consuming oxygen, but a peak is observed. This means that the DO has a faster response to flow changes than the bacteria concentration. Also a higher ammonium conversion is achieved.

The Anammox process was tested with the conditions used in the experiment of Güven et al. in 2004 in a lab-scale reactor. The ratio NO_2^-/NH_4^+ should be higher than 1 for stoichiometric reasons. Whit a hydraulic retention time of 5 days the ammonium removal efficiency is 94% and nitrite removal efficiency is higher than 99%. The overall nitrogen removal in the process is 90%.

6. CONCLUSIONS AND PERSPECTIVES

The main variable that links both reactors is the ratio NO2-/NH4+, which should be close to 1. The inspection of some step responses led to a qualitative screening of potential manipulated and controlled variables. For instance, pH should be controlled protons are produced because in the biological reaction driving the process to conditions out of the optimum range for the microbial growth. Another key parameter is the dissolved oxygen that has to be kept in a level that allows the growth of the AOB but prevents the growth of NOB or a too large conversion of ammonium. The hydraulic retention time play a similar role to oxygen as an intermediate value is needed to wash out the NOB but no the AOB. Temperature is also important because with low values, between 5-20 C, NOB growth faster than AOB. All this parameters can be controlled separately with different

feedback loops and fixed set points, but also cascaded structures can be used to actualize the set points when the ratio NO2-/NH4+ changes. Other possibilities are multivariable control strategies such as MPC. The implementation of an advanced control structure based on the modeling results will be investigated during future activities in this project.

REFERENCES

Galí, A., Dosta, J., van Loosdrecht, M.C.M. and Mata-Alvarez, J. (2007). Two ways to achieve an Anammox influent from real reject water treatment at lab-scale: Partial SBR nitrification and SHARON process. *Process Biochemistry* **42**, 715-720.

Güven, D., van de Pas-Schoonen, K., Schmid, M.C., Strous, M., Jetten, M.S.M., Sözen, S., Orhon, D. and Schmidt I. (2004). Implementation of the Anammox Process for Improved Nitrogen Removal, *Journal of Environmental Science and Health*, **A39**(7), 1729-1738.

Hellinga,C., Schellen, A.A.J.C., Mulder, J.W., van Loosdrecht, M.C.M., and Heijnen, J.J. (1998). The SHARON process: an innovative method for nitrogen removal from ammonium-rich waste water. *Wat. Sci. Tech.* **37**(9),135-142.

Mulder, A., van der Graaf, A.A., Robertson, L.A. and Kuenen, J.G. (1995). Anaerobic ammonium oxidation discovered in a denitrifying fluidized bed reactor. *FEMS Microbial Ecol.* **16**,177-83.

Volcke, E.I.P., Van Hulle, S.W.H., Donckels, B.M.R., van Loosdrecht, M.C.M. and Vanrolleghem, P.A. (2005). Coupling the SHARON process with Anammox: Model-based scenario analysis with focus on operating cost. *Wat. Sci. Tech.* **54**(4),107-115.

Appendix A. Reaction stoichiometry and expressions

Appendix B. Waste Water Treatment Plant with a SHARON-Anammox process.

Appendix A. Reaction stoichiometry

Table1. Gujer matrix of the SHARON process

j process	Aij												
	TNH	**TNO2**	**CO2**	**TIP**	**NO3**	**O2**	**N2**	**AOB**	**NOB**	**Xhet**	**Methanol**	**H+**	**H2O**
Ammonium oxidation	$-\frac{1}{Y_1}$	$\frac{1}{Y_1}-n$	-1			$-(-1+\frac{1.5}{Y_1}-\frac{h}{4}+\frac{o}{2}-\frac{3}{4}\cdot n)$		1				$\frac{2}{Y_1}-n$	$\frac{1}{Y_1}-\frac{h}{2}+\frac{n}{2}$
Nitrite oxidation	-n	$-\frac{1}{Y_2}$	-1		$\frac{1}{Y_2}$	$-(-1+\frac{0.5}{Y_1}-\frac{h}{4}+\frac{o}{2}+\frac{3}{4}\cdot n)$			1			n	$\frac{3}{2}\cdot n-\frac{h}{2}$
Denitrification nitrite	-n	$-\frac{1}{Y_3}$	$-\frac{1}{3}+\frac{0.5}{Y_3}-\frac{h}{6}-\frac{o}{3}-\frac{n}{2}$				$\frac{0.5}{Y_3}$			1	$-(\frac{2}{3}+\frac{0.5}{Y_3}+\frac{h}{6}-\frac{o}{3}-\frac{n}{2})$	$-\frac{1}{Y_3}+n$	$\frac{4}{3}+\frac{13}{6\cdot Y_3}-\frac{h}{6}-\frac{2\cdot o}{3}+\frac{n}{2}$
Denitrification nitrate	-n		$-\frac{1}{3}+\frac{5}{6\cdot Y_4}+\frac{h}{6}-\frac{o}{3}-\frac{n}{2}$		$-\frac{1}{Y_4}$		$\frac{0.5}{Y_4}$			1	$-(\frac{2}{3}+\frac{5}{6\cdot Y_4}+\frac{h}{6}-\frac{o}{3}-\frac{n}{2})$	$-\frac{1}{Y_4}+n$	$\frac{4}{3}+\frac{13}{6\cdot Y_4}-\frac{h}{6}-\frac{2\cdot o}{3}+\frac{n}{2}$
Methanol oxidation	-n		$-1+\frac{1}{Y_5}$			$-(-1+\frac{1.5}{Y_5}-\frac{h}{4}+\frac{o}{2}+\frac{3}{4}\cdot n)$				1	$-\frac{1}{Y_5}$	n	$\frac{2}{Y_5}-\frac{h}{2}+\frac{3}{2}\cdot n$

Table 2: Biological growth rate expressions of the SHARON process.

$\rho_1 = \mu_{\max}^{AOB} \frac{C_{NH_3}}{K_{NH_3}^{AOB} + C_{NH_3}} \cdot \frac{C_{O_2}}{K_{O_2}^{AOB} + C_{O_2}} \cdot \frac{K_{I,HNO_2}^{AOB}}{K_{I,HNO_2}^{AOB} + C_{HNO_2}} \cdot X_{AOB}$
$\rho_2 = \mu_{\max}^{NOB} \frac{C_{HNO_2}}{K_{HNO_2}^{NOB} + C_{HNO_2}} \cdot \frac{C_{O_2}}{K_{O_2}^{NOB} + C_{O_2}} \cdot X_{NOB}$
$\rho_3 = \mu_{\max}^{dNO2} \frac{C_{TNO2}}{K_{NO_2}^{dNO2} + C_{TNO2}} \cdot \frac{K_{I,O_2}}{K_{I,O_2} + C_{O_2}} \cdot \frac{C_{TNO2}}{C_{TNO2} + C_{NO_3}} \cdot \frac{C_{CH_3OH}}{C_{CH_3OH} + K_{CH_3OH}^{het,an}} X_{het}$
$\rho_4 = \mu_{\max}^{dNO3} \frac{C_{NO_3}}{K_{NO_3}^{dNO3} + C_{NO_3}} \cdot \frac{K_{I,O_2}}{K_{I,O_2} + C_{O_2}} \cdot \frac{C_{NO_3}}{C_{TNO2} + C_{NO_3}} \cdot \frac{C_{CH_3OH}}{C_{CH_3OH} + K_{CH_3OH}^{het,an}} X_{het}$
$\rho_5 = \mu_{\max}^{met} \cdot \frac{C_{O_2}}{K_{O_2}^{het} + C_{O_2}} \cdot \frac{C_{CH_3OH}}{C_{CH_3OH} + K_{CH_3OH}^{het,ox}} X_{het}$

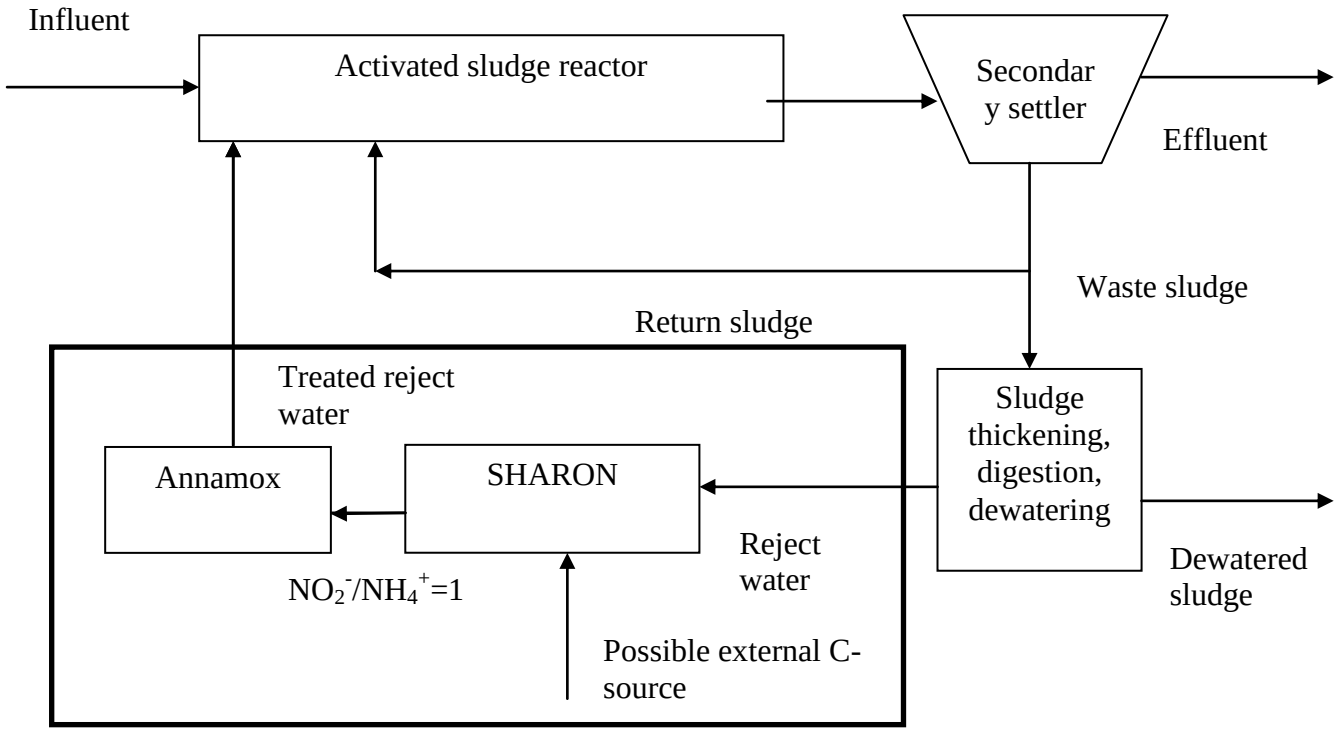
Table 3: Gujer matrix of the Anammox process.

j process	Aij												
	TNH	O2	TNO2	NO3	N2	alk	S	XAOB	XNOB	XAnAOB	Xhet	Xs	Xi
AOB growth	$-\frac{1}{Y_{AOB}} - i_{NXB}$	$-\frac{3.43-Y_{AOB}}{Y_{AOB}}$	$\frac{1}{Y_{AOB}}$			$-\frac{1}{14}(\frac{2}{Y_{AOB}}+i_{NXB})$		1					
NOB growth		$-\frac{1.14-Y_{NOB}}{Y_{NOB}}$	$-\frac{1}{Y_{NOB}}$	$\frac{1}{Y_{NOB}}$		$-\frac{i_{NXB}}{14}$			1				
AnAOB growth	$-\frac{1}{Y_{AnAOB}} - i_{NXB}$		$-\frac{1}{Y_{AnAOB}} - 1.52$	1.52	$\frac{2}{Y_{AnAOB}}$	$-\frac{i_{NXB}}{14}$				1			
AOB decay						$\frac{1}{14}(i_{NXB} - i_{NXI})$		-1				1-fi	fi
NOB decay						$\frac{1}{14}(i_{NXB} - i_{NXI})$			-1			1-fi	fi
AnAOB decay						$\frac{1}{14}(i_{NXB} - i_{NXI})$				-1		1-fi	fi
het1 growth	$-i_{NXB}$	$-\frac{1-Y_{het}}{Y_{het}}$					$-\frac{1}{Y_{het}}$				1		
het2 growth	$-i_{NXB}$		$-\frac{1-Y_{het}}{1.71Y_{het}}$		$\frac{1-Y_{het}}{1.71Y_{het}}$		$-\frac{1}{Y_{het}}$				1		
het3 growth	$-i_{NXB}$			$-\frac{1-Y_{het}}{2.86Y_{het}}$	$\frac{1-Y_{het}}{2.86Y_{het}}$		$-\frac{1}{Y_{het}}$				1		
het decay											-1	1-fi	fi
Hydrolysis	$-\frac{i_{NXB} - fi * i_{NXI}}{1-fi}$						1					-1	

Table 4: Biological growth rate expressions of the Anammox process.

$\rho_1 = \mu_{\max}^{AOB} \frac{C_{NH_3}}{K_{NH_3}^{AOB} + C_{NH_3}} \cdot \frac{C_{O_2}}{K_{O_2}^{AOB} + C_{O_2}} \cdot \frac{K_{I,HNO_2}^{AOB}}{K_{I,HNO_2}^{AOB} + C_{HNO_2}} \cdot X_{AOB}$
$\rho_2 = \mu_{\max}^{NOB} \frac{C_{HNO_2}}{K_{HNO_2}^{NOB} + C_{HNO_2}} \cdot \frac{C_{O_2}}{K_{O_2}^{NOB} + C_{O_2}} \cdot X_{NOB}$
$\rho_3 = \mu_{\max}^{AnAOB} \cdot \frac{C_{HNO_2}}{K_{HNO_2}^{AnAOB} + C_{HNO_2}} \cdot \frac{K_{I,O_2}^{AnAOB}}{K_{I,O_2}^{AnAOB} + C_{O_2}} \cdot \frac{C_{NH_3}}{K_{NH_3}^{AnAOB} + C_{NH_3}} \cdot X_{NOB}$
$\rho_4 = b_{AOB} \cdot X_{AOB}$
$\rho_5 = b_{NOB} \cdot X_{NOB}$
$\rho_6 = b_{AnAOB} \cdot X_{AnAOB}$
$\rho_7 = \mu_{\max}^{met} \cdot \frac{C_{O_2}}{K_{O_2}^{het} + C_{O_2}} \cdot \frac{C_S}{C_S + K_S^{het}} \cdot \frac{C_{TAN}}{C_{TAN} + K_{TAN}^{het}} \cdot X_{het}$
$\rho_8 = \mu_{\max}^{het} \cdot \eta_H \cdot \frac{C_{TNH}}{K_{TNH}^{het} + C_{TNH}} \cdot \frac{K_{I,O_2}^{het}}{K_{I,O_2}^{het} + C_{O_2}} \cdot \frac{C_{TNO2}}{C_{TNO2} + C_{NO_3}} \cdot \frac{C_S}{C_S + K_S^{het}} X_{het}$
$\rho_9 = \mu_{\max}^{dNO3} \cdot \eta_H \cdot \frac{C_{NO_3}}{K_{NO_3}^{het} + C_{NO_3}} \cdot \frac{K_{I,O_2}^{het}}{K_{I,O_2}^{het} + C_{O_2}} \cdot \frac{C_{TAN}}{K_{TAN}^{het} + C_{TAN}} \cdot \frac{C_S}{C_S + K_S^{het}} \cdot X_{het}$
$\rho_{10} = b_{AnAOB} \cdot X_{AnAOB}$
$\rho_{11} = k_H \cdot \frac{X_S}{K_X}$

Appendix B. Waste Water Treatment Plant with a SHARON-Anammox process:



Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Modelling Fungal Fermentations For Enzyme Production

Mads O. Albaek*, Krist V. Gernaey **,
Morten S. Hansen *, Stuart M. Stocks******

**Technical University of Denmark, Department of Chemical and Biochemical Engineering and Novozymes A/S; Bagsvaerd Denmark (Tel: +45 61 26 47 48; e-mail: maoa@novozymes.com)*
***Technical University of Denmark; Kgs. Lyngby Denmark (e-mail:kvg@kt.dtu.dk)*
****Novozymes A/S; Bagsvaerd Denmark (e-mail: mosh@novozymes.com)*
*****Novozymes A/S; Bagsvaerd Denmark (e-mail: stus@novozymes.com)*

Abstract: We have developed a process model of fungal fed-batch fermentations for enzyme production. In these processes, oxygen transfer rate is limiting and controls the substrate feeding rate. The model has been shown to describe cultivations of both *Aspergillus oryzae* and *Trichoderma reesei* strains in 550L stirred tank pilot plant reactors well. For each strain, 8 biological parameters are needed as well as a correlation of viscosity, as viscosity has a major influence on oxygen transfer. The parameters were measured averages of at least 9 batches for each strain. The model is successfully able to cover a wide range of process conditions (0.3-2 vvm of aeration, 0.2-10.0 kW/m^3 of specific agitation power input, and 0.1-1.3 barg head space pressure). Uncertainty and sensitivity analysis have shown that the uncertainty of the model is mainly due to difficulties surrounding the estimation of the biological parameters and to a lesser degree the uncertainty of the viscosity and mass transfer correlations. Until now, the model has been applied to evaluation of energy efficiency at different process conditions and bioreactor designs. Our goal is to expand the model to cover both pilot plant and production scale so that the model may assist downscaling operations as well as production optimization and production planning. Further developments of the model will enable more advanced applications such as model based control and simulated process optimization.

Keywords: process model, model based control, uncertainty analysis, sensitivity analysis, *Aspergillus oryzae*, *Trichoderma reesei*, pilot plant bioreactor, oxygen transfer, rheology

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Operation and Control of Enzymatic Biodiesel Production

Jason A. Price*, Jakob Kjøbsted Huusom*,
Mathias Nordblad *, John Woodley*

* *Department of Chemical and Biochemical Engineering,*

Technical University of Denmark, DK-2800 Kgs Lyngby, Denmark; (e-mail: {japr,jkh,matn,jw } @ kt.dtu.dk).

Abstract: This work explores the control of biodiesel production via an enzymatic catalyst. The process involves the transesterification of oils/fats with an alcohol (usually methanol or ethanol), using enzymatic catalysts to generate mono-alkyl esters (the basis of biodiesel) and glycerol as by-product. Current literature indicates that enzymatic processing of oils and fats to produce biodiesel is technically feasible and developments in immobilization technology indicate that enzyme catalysts can become cost effective compared to chemical processing. However, with very few exceptions, enzyme technology is not currently used in commercial-scale biodiesel production. This is mainly due to non-optimized process designs, which do not use the full potential of the catalysts in a cost-efficient way. Furthermore is it unclear what process variables need to be monitored and controlled to ensure optimal economics. Critical to the project is to develop a control methodology to optimize the productivity of biodiesel production (e.g. the dosing of alcohol to minimize catalyst deactivation, minimization of waste and delivering consistent product quality meeting specifications).

For production of biodiesel (BD) via an enzymatic route, batch operation is a straightforward and efficient means for producing BD with its main disadvantage being the downtime between batches. For large-scale production of biodiesel, continuous operation is an attractive alternative as it enables efficient use of manpower and capital assets including equipment and raw materials. Currently our group is evaluating various process configurations for continuous BD production in packed bed reactors (PBRs), continuous stirred tank reactors (CSTRs) and a combination of the aforementioned reactors in series. These configurations will be reviewed to identify the process variables that need to be monitored and controlled.

Keywords: Biodiesel, Enzymatic, transesterification, Modelling and Simulation, Process control

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Optimal Control of a Batch Reactor Using the Linearized Hamilton-Jacobi-Bellman Equation

Per Rutquist [**] **Torsten Wik** [*] **Claes Breitholtz** [*]

[*] *Department of Signals and Systems, Chalmers University of Technology, Göteborg, Sweden*
[**] *Tomlab Optimization AB, Västerås, Sweden*

AbstractIn this work we present an efficient method for solving an optimal control problem for a batch reactor, where a temperature dependent exothermic reaction takes place within a preset duration and within specified temperature bounds. The Hamilton-Jacobi-Bellman (HJB) equation corresponding to the optimal control problem is nonlinear and has infinite boundary conditions due to the state constraints (bounds on temperature and concentration), which makes it troublesome to solve. However, using a logarithmic transformation, the HJB-equation is transformed into a linear partial differential equation with zero boundary conditions. Furthermore, the problem can then be solved using variable separation such that the time-dependent part has an analytical solution and the state dependent part becomes a linear eigenvalue problem which can readily be solved using standard software.

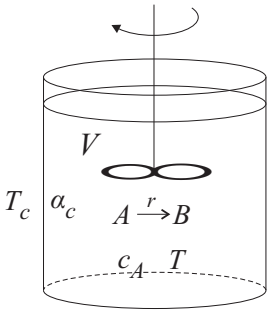


Figure 1. Illustration of the batch process.

1. INTRODUCTION

Consider the system illustrated in Figure 1, which is a batch process where a substance A reacts at a rate

$$r = k_0 e^{-\frac{E}{RT(t)}} c_A(t) \qquad \text{(mol/l/s)}$$

where E is the activation energy (J/mol), R is the ideal gas constant (J/mol/K), T is the temperature (K), k_0 is the reaction rate coefficient (1/s) and c_A is the concentration of A (mol/l). The reaction is exothermic, releasing ΔH_r Joule per mol A reacted.

The dynamics in the coolant system is ignored for simplicity and therefore we may consider the coolant temperature T_c (see Figure 1) to be the manipulated variable. The heat transfer coefficient α is assumed to be a constant parameter but in reality the transfer depends on local flows inside the reactor, coolant temperature fluctuations, flow fluctuations etc. These uncertainties and others, such as nonuniform reactions, are considered as one random disturbance that adds to the coolant temperature.

For control purposes we may write the mass and the energy balances for this system as

$$\begin{aligned} V\frac{d}{dt}c_A(t) &= -k_0 V e^{-\frac{E}{RT(t)}} c_A(t) \\ V\rho c_p \frac{d}{dt}T(t) &= \Delta H_r k_0 V e^{-\frac{E}{RT(t)}} c_A(t) \\ &\quad -\alpha_c(T(t) - T_c(t) - \Delta T_c) \end{aligned}$$

where V is the volume (l), ρ is the density (assumed unchanged by the reaction), c_p is the specific heat capacity (also assumed unchanged), and ΔT_c is the disturbance.

Selecting c_A to be the first state, T to be the second state and assuming the noise can be described by a Gaussian white noise with variance σ^2, we may write this on the form

$$\dot{x} = f(x) + G(x)(u + w), \tag{1}$$

where

$$f(x) = \begin{bmatrix} -k_0 x_1 e^{-\frac{a}{x_2}} \\ k_1 x_1 e^{-\frac{a}{x_2}} - k_2 x_2 \end{bmatrix} \quad \text{and} \quad G(x) = \begin{bmatrix} 0 \\ k_2 \end{bmatrix}$$

where $k_1 = \Delta H_r k_0/(\rho c_p)$ and $k_2 = \alpha_c/(\rho c_p V)$.

2. CONTROL PROBLEM

The batch process is operating with a cycle time t_f and at the end the concentration of A should have decreased to $c_{A,f}$ (with a corresponding produce of B). To maintain a sufficient rate of reaction the temperature should never go below $T_{\min}$ and to avoid problems of overheating it should never go above $T_{\max}$ (see Figure 2). The control problem can the be formulated as an optimization problem

$$\min_u V(x(t), t)$$

subject to Eq. (1) and

$$\begin{aligned} 0 \le x_1(t) &\le c_A(0) \\ T_{\min} \le x_2(t) &\le T_{\max} \end{aligned} \tag{2}$$

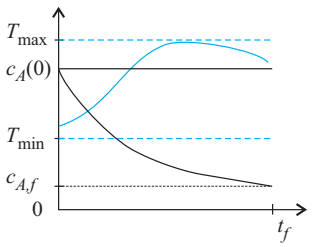


Figure 2. Constraints on the temperature and concentration trajectories.

where the cost function is

$$V(x(t),t) = \mathrm{E}\left\{\int_t^{t_f} u^2(\tau)d\tau + \gamma(x_1 - c_{A,f})^2\right\}$$

As can be seen there is no cost associated with the state trajectories although the method used would allow any such term on the form $l(x(\tau)) > 0$.

The optimal control law can be determined by solving the corresponding stochastic Hamilton-Jacobi-Bellman (HJB) equation (Dorato et al. [1995])

$$\begin{aligned} -\frac{\partial V}{\partial t} = &-\frac{\gamma}{4}(\nabla V)GG^T(\nabla V)^T + (\nabla V)f \\ &+ \frac{\sigma^2}{2}\mathrm{tr}\left[(\nabla^T\nabla V)GG^T\right] \end{aligned} \tag{3}$$

with $V \to \infty$ as $x \to \partial\Omega$, where the boundary $\partial\Omega$ is given by the state constraints (2).

3. METHOD

The nonlinearity and the infinite boundary conditions make this partial differential equation difficult to solve. However, by applying the transformation introduced by Rutquist et al. [2008] for stationary infinite horizon problems, i.e.

$$V = -\frac{2\sigma^2}{\gamma}\log(Z), \tag{4}$$

and separation of variables, i.e. $Z(t,x) = \Gamma(t)\Phi(x)$ we show that the optimal control is given by the solution to

$$\lambda\Gamma(t) = \frac{d}{dt}\Gamma(t), \tag{5}$$

which has the analytical solution $\Gamma(t) = ce^{\lambda t}$, and

$$\lambda\Phi = (\nabla\Phi)f - \frac{\sigma^2}{2}\mathrm{tr}\left[(\nabla^T\nabla\Phi)GG^T\right] \tag{6}$$

with the boundary condition

$$\Phi(x) = 0, \quad x \in \partial\Omega$$

In contrast to the difficult original HJB-equation this is a linear eigenvalue problem with zero boundary conditions that can readily be solved with standard software. The result is a family of solutions (λ_n, T_n, ϕ_n) from which the transformed cost is determined as a linear combination of the solutions for different eigenvalues, i.e.

$$Z(t,\mathbf{x}) = \sum_{n=1}^{\infty}\beta_n\exp(-\lambda_n(t_f - t))\phi_n(\mathbf{x}), \tag{7}$$

where the coefficients $\beta_n \in \mathbb{R}$ are given by the projection of the final condition

$$Z(t_f,\mathbf{x}) = \exp\left(-\frac{\gamma^2}{\sigma^2}(x_1 - c_{A,f})^2\right) \tag{8}$$

onto the space spanned by the eigenfunctions ϕ_n.

The optimal control policy can then be determined as

$$u = \frac{\sigma^2}{Z}G^T(\nabla Z)^T$$

This is illustrated by an application of the method to the reactor system described by Lagerberg and Breitholtz [1997].

REFERENCES

P. Dorato, T. C. Abdallah, and V. Cerone. *Linear Quadratic Control: An Introduction*. Prentice-Hall, 1995. ISBN 1575241560.

A. Lagerberg and C. Breitholtz. A study of gain scheduling control applied to an exothermic cstr. *Chem. Eng. Technol.*, 20:435–444, 1997.

P. Rutquist, C. Breitholtz, and T. Wik. On the infinite time solution to state-constrained stochastic optimal control problems. *Automatica*, 44:1800–1805, 2008.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Extended Abstract: Optimal Input Design for Parameter Identification in Dynamic Systems Using Nonlinear Programming ⋆

Tor Aksel N. Heirung * **B. Erik Ydstie** ** **Bjarne Foss** *

* *Department of Engineering Cybernetics, Norwegian University of Science and Technology, N-7491 Trondheim, Norway (e-mail: heirung@itk.ntnu.no).*
** *Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA.*

This contribution outlines work in progress on a novel approach to input design for parameter identification in dynamic systems. The goal is to find an input that is small in some sense, while minimizing the variances of the parameter estimates. The method in its present form does not aim to find an optimal input for a general class of systems; rather, the input will be optimal for a single system with one given set of parameter values. An objective function, the system equation, the identification equations, and bounds on the input will together form a nonlinear programming (NLP) problem, solvable with standard NLP solvers.

The proposed method will be outlined for linear time-invariant noise-free systems with constant parameters, one input, and one output. This class of systems can be formulated as autoregressive moving average (ARMA) processes

$$y(t) + a_1 y(t-1) + \cdots + a_n y(t-n) = b_0 u(t-1) + \cdots + b_{n-1} u(t-n) \quad (1)$$

where t is discrete time. A more compact notation is

$$y(t) = \varphi^\top(t)\theta(t) + e(t) \quad (2)$$

where

$$\theta = [b_0 \ \cdots \ b_{n-1} \ a_1 \ \cdots \ a_n]^\top$$

is a vector of all the parameters and

$$\varphi(t) = [u(t-1) \ \cdots \ u(t-n) \ -y(t-1) \ \cdots \ -y(t-n)]^\top$$

is a regression vector containing past inputs and outputs.

The two goals of minimizing the variances of the parameter estimates and using an input that is small in some sense are in general conflicting and must be balanced. A fairly general formulation of this objective is

$$\min_{u(t)} \sum_{t=0}^{N-1} \Big\{ f\big(P(t)\big) + g\big(u(t)\big) \Big\} \quad (3)$$

where N is the number of input samples used in the identification experiment. Specific examples include

$$\min_{u(t)} \sum_{t=0}^{N-1} \Big\{ w(t) \operatorname{trace} P(t) + u^2(t) \Big\} \quad (4)$$

and

$$\min_{u(t)} \sum_{t=0}^{N-1} \Big\{ w(t) \operatorname{trace} P(t) + (\Delta u(t))^2 \Big\} \quad (5)$$

In both (4) and (5), $w(t)$ is a possibly time-varying positive weight used for balancing the two goals. In both objectives the trace of the covariance matrix is minimized; in (4) this goal is weighted against keeping the amplitude of the input small, while in (5) it is weighted against keeping variations in the input small.

A Kalman filter or a recursive least-squares algorithm is used as the identification procedure. Then, the problem outlined above can be cast as the NLP problem

$$\min_{u(t)} \sum_{t=0}^{N-1} \Big\{ w(t) \operatorname{trace} P(t) + u^2(t) \Big\} \quad (6a)$$

$$\text{s.t. } y(t) = \varphi^\top(t)\theta \quad (6b)$$

$$\hat{\theta}(t+1) = \hat{\theta}(t) + K(t)\big(y(t) - \varphi^\top(t)\hat{\theta}(t)\big) \quad (6c)$$

$$K(t) = \varphi(t)\big(\varphi^\top(t)P(t)\varphi(t)\big)^{-1} \quad (6d)$$

$$P(t+1) = \big(I - K(t)\varphi^\top(t)\big)P(t) \quad (6e)$$

$$u_{\min} \le u(t) \le u_{\max} \quad (6f)$$

In this formulation, the Kalman filter is represented by (6c)-(6e) [Åström, 1970], where $\hat{\theta}$ is the estimate of the parameter vector, $K(t)$ is the Kalman gain, and $P(t+1)$ is the covariance matrix of the parameter estimates. In the case of hard constraints on the input, for instance due to limitations in the actuators, (6f) can be used to specify the bounds. When implemented in an NLP as in (6), the Kalman filter equations are nonlinear, despite the filter being linear. Hence, (6) is a nonconvex problem so meaning there can be several local solutions.

The NLP (6) can be solved using standard NLP solvers of interior-point or SQP type, for instance. The complexity of the NLP is not significantly influenced by the number of samples N. However, a large number of unknown system parameters does make the problem complex since the vector and matrix variables in the Kalman filter grow in size.

Despite the fact that the optimal input $u^*(t)$ is impossible to calculate for a physical system where the parameters are not known, this ideal input has several applications. Perhaps the most useful of these is providing a way of quantifying the quality of a suboptimal identification

⋆ This work was supported by the Center for Integrated Operations in the Petroleum Industry, Trondheim, Norway.

signal. For instance, assume some signal $u(t)$ was used to identify a set of parameter values; then, the expression

$$\sum_{t=1}^{N-1} \big(u(t) - u^*(t)\big)^2 \tag{7}$$

is a measure of the amount of unnecessary effort put into the identification experiment. Similarly, if the goal is to drive trace $P(t)$ to a minimum in the shortest time possible N^* while keeping $u(t)$ small, the expression $N - N^*$ measures the extra time spent on identifying the set of parameter values, where N is the time it takes for trace $P(t)$ to reach a minimum when $u(t)$ is used for identification. In a control setting, an interesting loss measure is reduced performance in closed-loop operation. Furthermore, analysis of the optimal input can give important insight on how to design suboptimal but more general inputs.

The method can be extended to state-space formulations rather than the input-output description used here without any major modifications. The ideas presented here also apply to systems with process noise or with time-varying parameters modeled as random walks. The problem formulation can also be modified so that the solution $u^*(t)$ in some sense is optimal for a range of parameter values. However, these modifications lead to stochastic or robust optimization problems which are significantly harder to solve.

Future work includes evaluating the performance of the method when applied repeatedly, i.e., at every time instant and with short horizon N. Parameter convergence will be prioritized. Applications within adaptive control will also be investigated, leading to a type of dual control structure [Fel'dbaum, 1961a,b,c,d] where the goal is to keep the output small while reducing parameter uncertainties.

REFERENCES

A. A. Fel'dbaum. Dual-control theory. I. *Automation and Remote Control*, 21(9):874–880, 1961a.

A. A. Fel'dbaum. Dual control theory. II. *Automation and Remote Control*, 21(11):1033–1039, 1961b.

A. A. Fel'dbaum. The theory of dual control. III. *Automation and Remote Control*, 22(1):1–12, 1961c.

A. A. Fel'dbaum. The theory of dual control. IV. *Automation and Remote Control*, 22(2):109–121, 1961d.

K. J. Åström. *Introduction to Stochastic Control Thory*. Academic Press, 1970.

Optimisation of Oil Production in Two – Phase Flow Reservoir Using Simultaneous Method and Interior Point Optimiser.

Dariusz Lerch, Carsten Völcker, Andrea Capolei, John Bagterp Jørgensen, Erling Halfdan Stenby

Natural petroleum reservoirs are characterised by 2-phase flow of oil and water in the porous media (e.g. rocks) which they are built of. Conventional methods of extracting oil from those fields, which utilise high initial pressure obtained from natural drive, leave more than 70 % of oil in the reservoir. A promising decrease of these remained resources can be provided by smart wells applying water injections to sustain satisfactory pressure level in the reservoir throughout the whole process of oil production. Basically to enhance secondary recovery of the remaining oil after drilling, water is injected at the injection wells of the down-hole pipes. This sustains the pressure in the reservoir and drives oil towards production wells. There are however, many factors contributing to the poor conventional secondary recovery methods e.g. strong surface tension, heterogeneity of the porous rock structure leading to change of permeability with position in the reservoir, or high oil viscosity. Therefore it is desired to take into account all these phenomena by implementing a realistic simulator of the 2-phase flow reservoir, which imposes the set of constraints on the state variables of optimisation problem. Then, thanks to optimal control, it is possible to adjust effectively injection valves to control 2 phase immiscible flow in every grid block of the reservoir and navigate oil to the production wells so it does not remain in the porous media. The use of such a smart technology known also as smart fields, or closed loop optimisation, can be used for optimising the reservoir performance in terms of net present value of oil recovery or another economic objective.

In order to solve an optimal control problem we use a direct collocation method where we translate a continuous problem into a discrete one by applying explicit and implicit Euler methods. A substantial challenge of finding optimal solution in a robust way comes along with handling the scale of the optimal control problem due to discretisation in time and space.

Consequently, an Ipopt(Interior Point Optimiser) open source software for large scale nonlinear optimisation was applied. Because of its versatile compatibility with programming technologies, a C++ programming language in Microsoft Visual Studio integrated development environment was used for modelling the optimal control problem. Thanks to object oriented features of the language, it was possible to approach the problem in a very modular way by automating the discretisation process and develop interfaces for retrieving information from a continuous problem.

When tackling this problem, we reduce approximation error made by discretising of the original problem, by increasing the number of simulation steps and therefore it is necessary to solve large instances of the reformulation. As a result, it is very suitable to use Ipopt algorithm which implements an interior-point line-search filter method making it very powerful for solving large problems with up to hundreds of millions of constraints and variables.

The paper is a based on some work done within master project written at Center for Energy Resources Engineering, under the supervision of John Bagterp Jørgensen and support of phd students Andrea Capolei and Carsten Völcker.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Data Analysis and Monitoring of Thickness Sensor Fouling Using Self-Organizing Maps

Vesa-Matti Tikkala, * **Tommi Myller,** ** **Tomi Kulomaa,** **
Veikko Hämäläinen, *** **Sirkka-Liisa Jämsä-Jounela** *

* *Aalto University School of Chemical Technology,*
Department of Biotechnology and Chemical Technology,
P.O. Box 16100, 00076 Aalto, Finland
(email: vesa-matti.tikkala@aalto.fi)
** *Stora Enso Oyj, Imatra, Finland*
*** *Efora Oy, Imatra, Finland*

Abstract: This paper presents a case study of utilizing self-organizing maps (SOM) for data analysis and process monitoring with an application to a board machine process. The SOM is used to analyse data in order to find the causes for a thickness sensor fouling fault. Furthermore, it is used for monitoring the state of the process in order to estimate when fouling occurs. The results show that by means of the data analysis it is possible to identify the variables affecting the fouling problem. Also, the process state could be estimated with satisfactory accuracy using the SOM.

Keywords: Process monitoring, self-organizing map, board machine, quality control system, thickness sensor, fouling

1. INTRODUCTION

Increasing complexity of the modern production processes, tightening global competition and environmental regulations are posing enormous challenges to the operation of the production plants. The plants must run safely and efficiently, without any disturbances. To this end, new solutions to tackle the abnormal events in the processes have been searched by means of fault diagnosis and process monitoring.

According to Isermann (2006), process monitoring is considered as a countinuous on-line task of determination of the conditions of the process. Typically, this is achieved by process history data -based methods (see e.g. Venkatasubramanian et al. (2003)).

One of these methods is the self-organizing map (SOM). It was introduced by Kohonen (1982) and it has a wide variety of applications in different fields (Kangas and Kaski (1998)), such as visualization or voice and image analysis. It has also been successfully applied to process monitoring tasks, such as monitoring of fluidized bed combustion process (Liukkonen et al. (2011)), Internet-based remote supervision (Domínquez et al. (2007)), fault diagnosis of ethylene cracking process (Kämpjärvi et al. (2008)) and monitoring of flash smelting furnace (Jämsä-Jounela et al. (2003)), for instance.

In this paper, the SOM is utilized in the data analysis and monitoring of thickness sensor fouling at a board machine. Thus, the objectives of this case study are (1) to find out the variables affecting thickness sensor fouling; (2) to study in which process conditions the fouling occurs; and (3) to utilize self-organizing maps to monitor the process and to estimate when fouling occurs.

This paper is organized as follows. In Section 2, the description of the self-organized maps and its implementation is given. Section 3 describes the case process and the thickness sensor fault in question. Testing procedure, the results of the data analysis and the monitoring tests are presented in Section 4, followed by the conclusions in Section 5.

2. DESCRIPTION OF SELF-ORGANIZING MAP AND ITS IMPLEMENTATION

2.1 Mathematical description of the SOM

A self-organizing map is a type of artificial neural network that is trained using unsupervised training to produce a low-dimensional representation of the input space of the training samples, called a map. SOM produces a similarity graph of the input data by converting the nonlinear statistical relationships between high-dimensional data into simple geometric relationships on a low-dimensional display, usually a two-dimensional grid of nodes. Therefore, SOM compresses the data, but preserves the topological properties. (Kohonen (1998))

A SOM consists of a number of neurons or nodes that are described with a d-dimensional weight vector (sometimes referred as a codebook vector)

$$\mathbf{w} = [w_1, w_2, \ldots, w_d], \tag{1}$$

where d denotes the dimension of the input data vectors (number of variables). The nodes are organized in the map according to a specific topology. Typically, the SOMs are presented as two-dimensional sheets, where the nodes are arranged into a rectangular or hexagonal lattice.

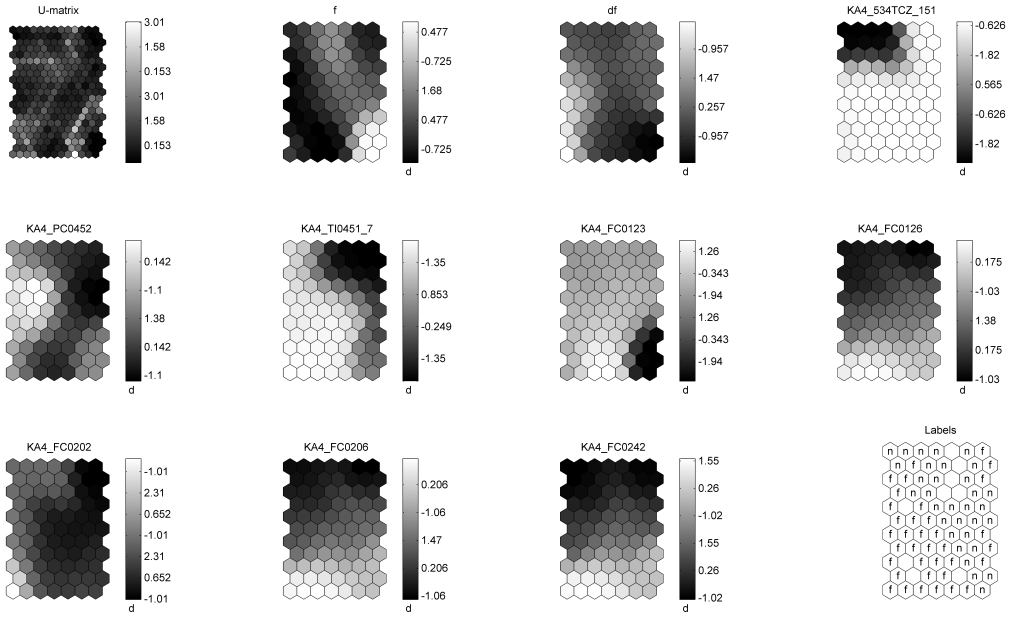


Fig. 4. Overview of the SOM analysis for the training data: U-matrix, single variables maps and distribution of the faulty and normal operation samples.

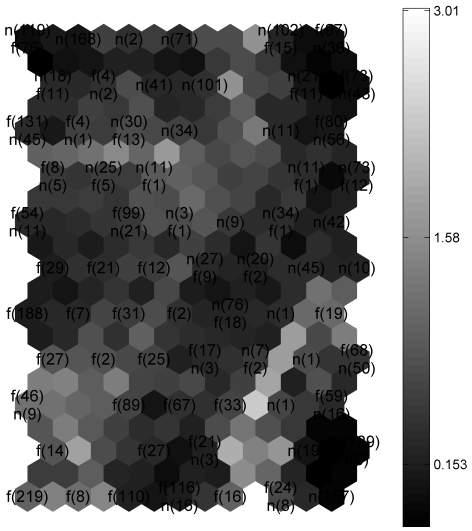


Fig. 5. Unified distance matrix of the training set showing the frequency of the samples on each map node.

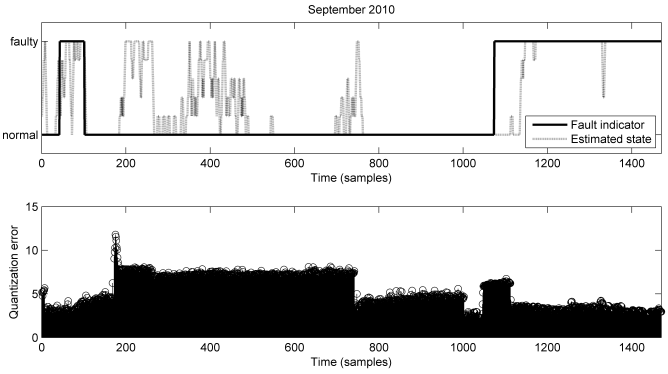


Fig. 6. Monitoring results using SOM: September data

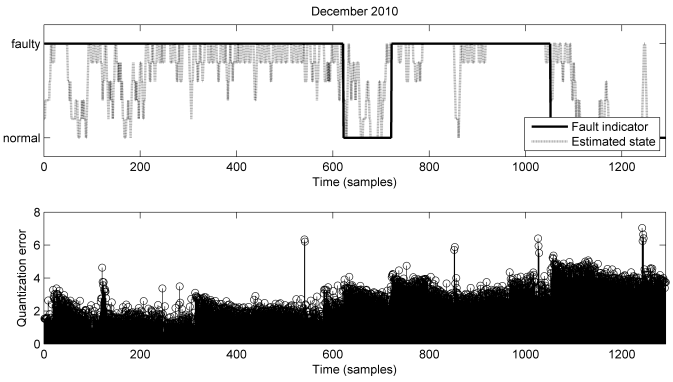


Fig. 7. Monitoring results using SOM: December data

process state was determined if the BMU matches with a faulty map node. The testing of the SOM was performed using the data sets of September and December.

The monitoring results of the SOM are presented in Fig. 6 and 7. In both figures, the upper panel compares the estimated process state and the fault indicator. The bottom panels show the errors related to the calculation of the BMUs. To reduce noise and false alarms, the estimated state has been filtered using a moving average filter with a window length of 5 samples.

It can be confirmed from both figures that the SOM gives a rather good estimate for the actual process condition. In the September data (Fig. 6), the SOM can detect the faulty periods at the beginning of the month as well as after $t = 1100$. However, the process state is falsely estimated to be faulty after $t = 200$ and around $t = 400$. In the December data (Fig. 7), the process state is estimated satisfactorily during the first 600 samples except minor fluctuations in the estimation around $t = 100$ and $t = 200$. The non-faulty period after $t = 600$ is estimated successfully as well as the period in the end of the month. Table 3 summarizes the performance of the SOM by showing the rates of correctly estimated states, falsely estimated states and uncertain states.

The self-organizing map is trained by adapting the weights of the nodes to match the input data. The training can be performed in two ways: by a sequential training or a batch training. Basically, training includes first the search of the closest map units, called the best-matching units (BMU), of the data samples and then the update of the weight vector of the BMU and its neighbouring nodes. The fundamental difference between the different training procedures is that in sequential training the weights are updated after each sample and in batch training the weights are updated once per epoch. The distance to the BMU c is determined for a data sample $\mathbf{x} \in \mathbb{R}^d$ as follows

$$||\mathbf{x} - \mathbf{w}_c|| = \min_i\{||\mathbf{x} - \mathbf{w}_i||\}, \quad i = 1, \ldots, m, \qquad (2)$$

where $||\cdot||$ is a distance measure, typically Euclidean, and m is the number of map nodes.

After finding the BMU, the weight vector of the BMU and the neighbouring nodes are updated according to an update rule, which typically has the following form

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + \alpha(t)h_{ci}(t)[\mathbf{x}(t) - \mathbf{w}_i(t)], \qquad (3)$$

where t denotes time, $h_{ci}(t)$ is the neighbourhood kernel around the BMU and $\alpha(t)$ is the learning rate. The neighbourhood kernel defines the effect of the update on the neighbouring nodes. The shape of the neighbourhood kernel and its radius determines the change of the weights of the nodes to be updated on each training step. Typically, a Gaussian neighbourhood function

$$h_{ci}(t) = e^{-d_{ci}^2/2\sigma_t^2}, \qquad (4)$$

where $d_{ci} = ||\mathbf{r}_c - \mathbf{r}_i||$ is the distance between map nodes in the grid and σ_t is the neighbourhood radius defining the width of the function, is used.

In the batch training procedure the BMUs are calculated first for the whole data set, and then the weights of the nodes are updated at once. The data set is partitioned according to the Voronoi regions of the map units. That is, each input data vector $\mathbf{x}$ belongs to the data set of the map unit which it is closest to. Then the sum of vectors in each Voronoi set is calculated as follows

$$\mathbf{s}_i(t) = \sum_{j=1}^{n_{V_i}} \mathbf{x}_j \qquad (5)$$

where n_{V_i} is the number of samples in the Voronoi set of the node i. The new values of the weight vectors are calculated as

$$\mathbf{w}_i(t+1) = \frac{\sum_{j=1}^m h_{ij}(t)\mathbf{s}_j(t)}{\sum_{j=1}^m n_{V_j}(t)h_{ij}(t)}. \qquad (6)$$

2.2 SOM algorithm implementation

The SOM algorithm implementation (SOM Toolbox, Vesanto et al. (2000)) used in this case study uses the batch training algorithm presented in the previous section. However, in order to reduce the memory consumption, the best-matching units for each input data sample are calculated, not at once, but in several batches. Nevertheless, the BMUs corresponding to each input data sample are found before any adjustments to the weights are made.

The SOM is trained in two phases. The rough training phase is performed first where a smaller number of training epochs and a larger neighbourhood radius are used than in the following fine-tuning phase.

3. DESCRIPTION OF THE PROCESS AND THE THICKNESS SENSOR FOULING FAULT

3.1 Board Machine Process

The board making process begins with the preparation of raw materials in the stock preparation section. Different types of pulp are refined and blended according to a specific recipe in order to achieve the desired composition and properties for the board grade to be produced. The consistency of the stock is controlled with dilution water.

The blended stock is pumped from the stock preparation to the short circulation by a pump that controls the basis weight of the board. In the short circulation, the stock is first diluted in the wire pit to the correct consistency for web formation. The diluted stock is then cleaned and screened before passing to the head box, from where it is sprayed onto the wire in order to form a solid board web.

The excess water is first drained through the wire and later by pressing the board web between rolls in the press section. The remaining water is evaporated off in the drying section using steam-heated drying cylinders. After the drying, the board is calendered in two phases in order to achieve the desired surface properties.

The important quality variables, such as basis weight, moisture and thickness, are measured after the calender section with a measurement scanner traversing across the board web.

3.2 Thickness Sensor Fouling Fault

At the case board machine, thickness sensor fouling is a difficult problem. In order to maintain the functionality of the quality control system, the sensor must be cleaned on a regular basis. Otherwise, the fouling impedes the online control of the thickness profile and the monitoring of board thickness in machine direction.

The thickness sensor is located in the measurement scanner. It consists of two plates that are in contact to each side of the web, see Fig. 1. The board travels between the plates and the distance between the plates can be measured by measuring the magnetic resistance caused by the board thickness. Due to this contacting sensor construction, the sensor is subject to fouling. The dirt builds up on the sensor plates and disturbs the measurement.

The thickness measurement will drift significantly when the fouling occurs. In Fig. 2, the measurement data of

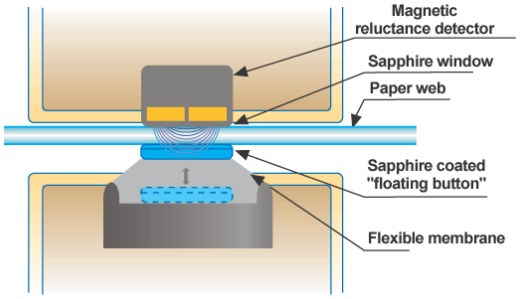


Fig. 1. Cross-sectional diagram of the thickness sensor. Modified from Anon. (2009).

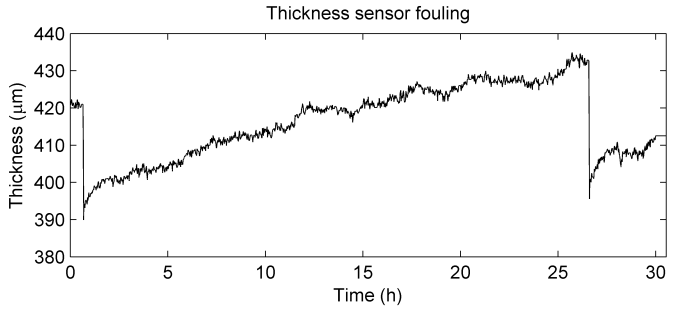


Fig. 2. Effect of the thickness sensor fouling. Data of 30 hours showing the evolution of the thickness sensor reading between two cleanings (at 1 hour and 27 hours).

30 hours are shown during a faulty period. It is clearly displayed that the measured value increases in time, while the setpoint of the thickness was kept constant over the whole period. Between the cleanings of the sensor at 1 hour and 27 hours, the slope of the drift is approximately 1.2 μm/h.

According to the maintenance experts, additional problems are caused due to the flaking of the dirt. It typically takes place, when the dirt layer reaches approximately 20 μm. This will cause serious disturbances to the measurement.

To facilitate the mathematical analysis of the thickness sensor fault, a qualitative representation of the fault was needed. To this end, a binary variable, called the fault indicator, was developed based on the maintenance data of the board machine. According to the recorded maintenance actions on the thickness sensor, the fault indicator was given value 1 or 0, indicating faulty and normal operation, respectively.

4. DATA ANALYSIS AND MONITORING RESULTS

4.1 Testing Procedure

In order to reach the objectives of the study, the analyses were conducted according to the following procedure:

First, a correlation analysis was carried out in order to find out which variables are affecting the sensor fouling. The list of correlated variables was augmented with variables selected based on the process knowledge. Then, the data of the selected variables were analysed using self-organizing maps in order to confirm the interactions between the variables and the sensor fault. Simultaneously, a SOM was trained for testing purposes.

Finally, the SOM was used to detect the conditions in which the thickness sensor fouling. A testing data set containing data from the faulty operation periods was introduced to the SOM. Best-matching units were calculated for each sample in the test data. The process conditions in which the fouling occurs were detected, when the faulty samples of the test data were located similarly than the faulty samples of the training data.

Table 1. List of variables for the SOM analysis

#	Tag	Description
1	`f`	Thickness control error (cv-sp)
2	`df`	Filtered derivative of `f`
3	`534TCZ_151`	1st calender thermo roll temperature
4	`PC0452`	Zero-pressure level of the secondary hood
5	`TI0451_7`	Hood ventilation air temperature 7
6	`FC0123`	Wet strength size flow (top/bottom)
7	`FC0126`	Starch flow (top/bottom)
8	`FC0202`	Neutral size flow (top)
9	`FC0206`	Retention starch flow (top)
10	`FC0242`	Retention agent flow (top)

4.2 Variable selection

Variable selection is the most critical task in the analysis. The list of variables was formed based on a correlation analysis and on further process and engineering knowledge.

By means of the correlation analysis, an indication that the temperature of the web (affected by calendar temperature, etc.) or the temperature around the measurement scanner had an effect on the fouling. Also, certain chemicals used in the board production seemed to correlate with the fault. In the data analysis using SOM, these variables were further studied in order to confirm the results of the correlation analysis.

The list of variables for the SOM analysis consisted of thickness control error and its filtered derivative, temperature of the 1st calendar, zero-pressure level of the secondary hood, hood ventilation air temperature and some flows related to the chemicals used in the board production, see Table 1.

4.3 Data Preparation

When using SOM for the analyses, there are two data sets needed: training data and testing data. Total of four months of data was divided into two, such that October and November data were to be used as the training data and the tests were to be run with September and December data sets. The sampling interval used in the SOM analysis was 5 minutes. The training data set contained 8 variables and 3775 samples (approx. 314 hours), see Fig. 3. The testing data sets contained 1370 and 1210 samples for September and December data sets, respectively.

The data was prepared by removing the non-production data, i.e. shut-downs and web breaks, removing the outliers and by normalizing the data. The non-production data was removed by excluding data segments, when the production rate has zero value. Outliers were removed manually by observing the data and replacing the diverging values with an average of its neighbours. Finally, the data was normalized to reduce the effect of different magnitudes of the measurement signals.

The data samples were also labelled in order to separate the faulty operation from the normal operation. The labels for each sample were determined based on the fault indicator variable. Samples falling into faulty periods (fault indicator = 1) were labelled with 'f' and the samples representing normal operation were labelled with 'n'.

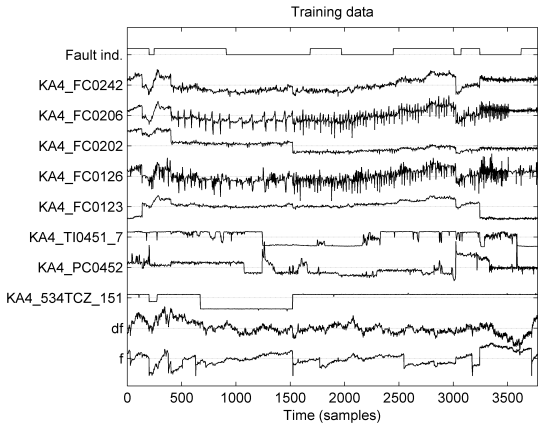


Fig. 3. Training and analysis data for the SOM

4.4 SOM training parameters

SOM analysis was performed using an algorithm implemented for this purpose. The parameters used in the training of the SOM models are listed in Table 2. The map size m was determined according to suggestions given in literature (Vesanto et al. (2000)). Also, the number of training epochs and neighbourhood function radii were based on default values in literature (Vesanto et al. (2000)) and calculated according to the data and map size. Map size was calculated as follows

$$m = \text{ceil}(5\sqrt{n}), \tag{7}$$

where n is the number of samples in the training data. The map dimensions are then calculated using eigenvalue decomposition of the training data matrix. The ratio of height m_h and width m_w of the map is determined as a ratio of largest and second largest eigenvalue. For the SOM used in this analysis the map was found to consist of $m = 77$ nodes organized in an 11-by-7 rectangular grid.

The SOM was initialized by using linear initialization in which the map weight vectors are given initial values that lay in the subspace spanned by the two largest eigenvectors of the input data. The training length is determined for training phases, rough training and finetuning, based on the size of the map and size of the data. A suggestion given in (Vesanto et al. (2000)) first defines a ratio of number of map units m and length of the data n:

$$\xi = m/n. \tag{8}$$

Then the used training length $20 \cdot \xi$ from which $4 \cdot \xi$ is for rough training and $16 \cdot \xi$ for finetuning. In this case, the data length was large ($n = 3775$) compared to the number of map units ($m = 77$) resulting in $\xi < 1$. Therefore, the used training lengths were 1 epoch for rough training and 1 epoch for finetuning.

Neighbourhood function used by the algorithm is a Gaussian kernel function, see (4). Its radius was defined for both training phases separately. Typically, a larger neighbourhood radius is used in the rough training than in the finetuning phase. For the rough training phase the neighbourhood radius is determined as follows:

$$\sigma_t = \max(1, \max([m_h \; m_w])/8). \tag{9}$$

For the case study, the rough training radius was determined using (9) resulting $\sigma_{tr} = 1.375$ and for the finetuning phase radius $\sigma_{tf} = 1$ was used.

Table 2. Parameters of the SOM training

Parameter	Description	Value
m	Map size	77
m_y, m_x	Map dimensions	[11, 7]
	Initialization	Linear
	Training algorithm	Batch
	Number of training epochs, rough training	1
	Number of training epochs, finetuning	1
h_{ci}	Neighbourhood kernel function	Gaussian
σ_{tr}	Neighbourhood radius, rough training	1.375
σ_{tf}	Neighbourhood radius, finetuning	1

4.5 Training data analysis

The SOM analysis was performed on the training data set comprising October and November data. In Fig. 4 are presented the results of the training data analysis.

The first panel (top left) of the figure shows the unified distance matrix (later U-matrix), which describes the clustering of the data. Dark colours denote a short distance between the map nodes, thus representing closely clustered data samples. On the contrary, light colours denote long distance.

By studying the plot of data labels and their distribution on the map (see the bottom left panel of Fig. 4), it is noticed that most of the faulty samples are distributed below the diagonal of the map. The distribution of the samples is illustrated more in detail in Fig. 5, where the U-matrix is presented with the frequency of each sample class ('f' or 'n') on each node of the map. The figure confirms the foregoing analysis of the distribution of faulty samples mostly on the bottom right part of the map.

For the monitoring purposes the nodes belonging to the faulty clusters located in the bottom right part of the U-matrix were labelled as faulty nodes, i.e., when the process is operating in those conditions, there is a risk of fouling of the thickness sensor.

By studying the single variable maps presented in Fig. 4, the effect of the temperature, especially the hood ventilation air temperature (`TI0451_7`), is obvious; the light areas indicating high values are concentrated in the same part of the SOM than the faulty observations in the U-matrix (Fig. 5). Calender temperature (`534TCZ_151`) has not such clear distribution of high and low values, but it is notable that when the fault is occurring, the temperature of the calendar is high. There are hardly any faulty observations located to areas where the temperature is low. From the chemicals, the neutral size seems to have the most explicit effect on the fouling, which occurs when the neutral size is used the most.

4.6 Monitoring of sensor fouling

The objective of the monitoring tests was to find out whether it is possible to use the trained SOM to detect when the process is operating in such conditions that fouling of the thickness sensor might occur. The tests were carried out by introducing the SOM a new data set containing normal operation data as well as faulty data and the best-matching units for each data sample were found out. The calculated BMUs were compared to map nodes labelled as faulty in the training phase. A faulty

Table 3. Results of the monitoring tests using SOM

	Sep	Dec
Rate of correct process states	78.0%	72.9%
Rate of false process states	11.9%	9.7%
Rate of uncertain process states	10.1%	17.4%

Based on the monitoring tests, the SOM is able to estimate the state of the process correctly in over 70% of time. The rate of falsely estimated states is rather low, approximately 10% on average. The perceived errors may result from the fault indicator, which has been developed based on the dates of the fault reports and therefore it might not be exactly aligned with actual fouling.

5. CONCLUSION

In this study, self-organizing maps were used to analyse the causes of a persistent thickness sensor fault at the case board machine. SOM was also tested for monitoring the process conditions in which fouling occurs.

Based on the preceding analyses, the main causes leading to the fouling of the thickness sensor are the temperature in the surroundings of the measurement frame and the temperature of the web, and the chemicals used in the board production, particularly, the neutral size.

Both factors seem reasonable in terms of the process operation. Neutral size is used in the process to improve the hydrophobicity of the board. It is a resin based size, which activates in high temperatures (drying section) and creates a hydrophobic layer on the fibres. Therefore it is possible that the size and high temperature is forming sticky substance that attach to the thickness sensor plates and disturb the measurement. Process experts at the plant say that neutral size also causes problems for instance in pipes by blocking them.

By utilizing SOM the process conditions in which the fouling occurs could be detected with reasonable accuracy. Therefore, an indicator, which informs the operators that the process is under condition where it is likely to have problems with the thickness measurements, can be developed. Then, the operators can take counteractions by schedling maintenance or by changing the operating point if possible.

ACKNOWLEDGEMENT

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007–2013) under grant agreement no. 257580. The research consortium is acknowledged for the support. In particular, the authors wish to thank Stora Enso Oyj and Efora Oy for providing the data and the expert knowledge for the analyses.

REFERENCES

Anon. (2009). Knowpap 11.0 — learning environment for papermaking and automation.

Domínquez, M., Fuertes, J., Reguera, P., Díaz, I., and Cuadrado, A. (2007). Internet-based remote supervision of industrial processes using self-organizing maps. *Engineering Applications of Artificial Intelligence*, 20, 757–765.

Isermann, R. (2006). *Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance*. Springer.

Jämsä-Jounela, S.L., Vermasvuori, M., Endén, P., and Haavisto, S. (2003). A process monitoring system based on the kohonen self-organizing maps. *Control Engineering Practice*, 11(1), 83–92.

Kämpjärvi, P., Sourander, M., Komulainen, T., Vatanski, N., Nikus, M., and Jämsä-Jounela, S.L. (2008). Fault detection and isolation of an on-line analyzer for an ethylene cracking process. *Control Engineering Practice*, 16, 1–13.

Kangas, J. and Kaski, S. (1998). 3043 works that have been based on the self-organizing map (SOM) method developed by kohonen. Technical Report A49.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43, 59–69.

Kohonen, T. (1998). The self-organizing map. *Neurocomputing*, 21, 1–6.

Liukkonen, M., Hiltunen, T., Hälikkä, E., and Hiltunen, Y. (2011). Modeling of fluidized bed combustion process and NO_x emissions using self-organizing maps: An application to the diagnosis of process states. *Environmental Modelling and Software*, 26, 605–614.

Venkatasubramanian, V., Rengaswamy, R., Kavuri, S.N., and Yin, K. (2003). A review of process fault detection and diagnosis, part iii: Process history based methods. *Computers and Chemical Engineering*, 27, 327–346.

Vesanto, J., Himberg, J., Alhoniemi, E., and Parhankangas, J. (2000). Som toolbox for matlab 5. Technical Report A57.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Production Optimization for Two-Phase Flow in an Oil Reservoir

Carsten Völcker, John Bagterp Jørgensen, Per Grove Thomsen

Department of Informatics and Mathematical Modeling
Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

Erling H. Stenby

Department of Chemical and Biochemical Engineering
Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

Keywords : Reservoir simulation/management, Runge-Kutta, ESDIRK, optimal control, nonlinear model predictive control, adjoint sensitivity.

Petroleum reservoirs are subsurface formations of porous rocks with hydrocarbons trapped in the pores. Initially, the reservoir pressure may be sufficiently large to push the fluids to the production facilities. However, as the fluids are produced the pressure declines and production reduces over time. When the natural pressure becomes insufficient, the pressure must be maintained artificially by injection of water. Conventional technologies for recovery leaves more than 50% of the oil in the reservoir. Wells with adjustable downhole flow control devices coupled with modern control technology offer the potential to increase the oil recovery significantly. In optimal control of smart wells, downhole sensor equipment and remotely controlled valves are used in combination with large-scale subsurface flow models and gradient based optimization methods in a Nonlinear Model Predictive Control framework to increase the production and economic value of an oil reservoir. Wether the objective is to maximize recovery or some financial measure like Net Present Value, the increased production is achieved by manipulation of the well rates and bottom-hole pressures of the injection and production wells. The optimal water injection rates and production well bottom-hole pressures are computed by solution of a large-scale constrained optimal control problem.

The objective is to maximize production by manipulating the well rates and bottom hole pressures of injection and production wells. Optimal control settings of injection and production wells are computed by solution of a large scale constrained optimal control problem. We describe a gradient based method to compute the optimal control strategy of the water flooding process. An explicit singly diagonally implicit Runge-Kutta (ESDIRK) method with adaptive stepsize control is used for computationally efficient solution of the model. The gradients are computed by the adjoint method. The adjoint equations associated with the ESDIRK method are solved by integrating backwards in time. The necessary information for the adjoint computation is calculated and stored during the forward solution of the model. The backward adjoint computation then only requires the assembly of this information to compute the gradients.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Reducing Revenue Loss due to Utility Disturbances using Buffer Tanks – A Case Study at Perstorp

Anna Lindholm
Automatic Control, Lund

Charlotta Johnsson
Automatic Control, Lund

Tore Hägglund
Automatic Control, Lund

Hampus Carlsson
Perstorp AB

I. Introduction

In the chemical process industry, disturbances in utilities such as steam and cooling water often cause large losses. Earlier studies have been performed on the synthesis of utilities to satisfy the demand, for example in [1] and [2]. The present study focuses on how disturbances in utilities affect production. Since utilities often are used plant-wide, disturbances in the supply of utilities may affect large parts of a site, either directly or indirectly because of the connections of production areas via the product flow. A general method for handling disturbances in utilities has recently been proposed in [3]. The method is called the utility disturbance management (UDM) method. In the present study, this framework is applied to an industrial site at Perstorp. The objectives are to obtain an indication of which utilities that cause the greatest revenue losses at the site, and to suggest strategies for reducing these losses. To complete all steps of the UDM method, a model of the production site is needed. Chemical plants are often complex, and thus difficult and time-consuming to model in detail. Here, a simple modeling approach is used, in which production areas at a site are modeled as either 'on' or 'off', i.e. either producing at maximum production rate or not at all. Buffer tanks between areas are also included. If a production area has to be shut down due to a utility disturbance, buffer tanks will allow production to continue for a certain period in downstream areas, before it is necessary to shut down these areas as well. This coarse model will not capture all the variability, but has shown to be useful in providing indications of the effects of disturbances in utilities on production.

II. The UDM method

The UDM method, introduced in [3], aims to reduce the economic effects of disturbances in utilities. The method consists of four steps:

1) Get information on site-structure and utilities
2) Compute utility and area availabilities
3) Estimate revenue loss due to disturbances in utilities
4) Reduce revenue loss due to future disturbances in utilities

Each of the steps have a number of sub-steps, which are defined in [3]. In the case study at Perstorp, the method is applied using an on/off production modeling approach including buffer tanks between production areas. A case study has previously been performed at the same production site using on/off production modeling without including buffer tanks. The results from this study are presented in [4].

III. Main results of case study

The site that is studied is a site owned by Perstorp that produces specialty chemicals, and is located in Stenungsund, Sweden. The site consists of 10 production areas, producing products 1-10. Internal buffer tanks exist for products 1-5. A flowchart ot the product flow at the site is shown in Figure 1.

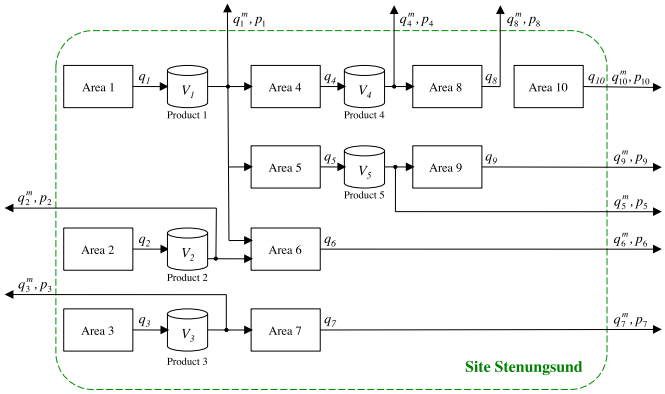


Fig. 1. Product flow at site Stenungsund.

The site is modeled by on/off production including buffer tanks. This gives ordering of utilities according to an estimate of the revenue loss they cause, which could be of great help for proactive disturbance management, i.e. when trying to reduce the number of disturbances in the future. At site Stenungsund, use of the UDM method showed that the cooling water utility seems to cause the greatest loss of all utilities at the site. In addition to this, two strategies for decreasing the revenue loss due to utilities are obtained. The first is suggestions on how buffer tank levels should be chosen to minimize the effects of disturbances in utilities (proactive disturbance management), and the second is suggestions on how to control the production at the occurrence of a utility disturbance (reactive disturbance management). Below, these two strategies are discussed.

A. Choice of buffer tank levels

Good choices of stationary buffer tank levels can ensure that the site can run even at a failure in one or more areas. In this case study, it has been chosen to only consider downstream effects of a disturbance upstream of a buffer tank. Thus, only lower constraints on the buffer tank levels will be imposed, and there will be a trade-off between handling as many failures as possible and minimizing inventory at the

site. This work does not focus on computing the costs of the inventories to achieve the optimal trade-off between utility disturbance management and cost of inventory. Optimal choice of inventory is discussed in e.g. [5] and [6].

Choosing the buffer tank levels to handle the longest disturbance durations for utilities will often give unneccesarily high buffer tank levels at normal operation, since disturbances of such long durations often are very uncommon. A suggestion is to choose the levels so that a certain percentage of all disturbances in utilities are handled. In Figure 2, the levels that correspond to handling 90 % of all disturbances in utilities at site Stenungsund are given, based on measurement data from Aug 1, 2007 to July 1, 2010. As a comparison, the average buffer tank levels over the considered time period are shown in the figure. It can be seen that the average buffer

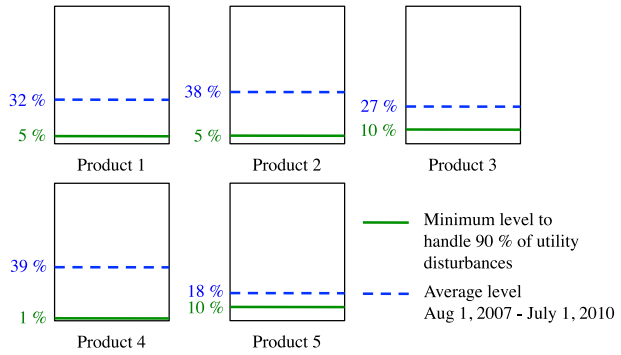


Fig. 2. Buffer tank levels at site Stenungsund.

tank levels over the selected time period are well above the levels required to handle 90 % of all disturbances in utilities. However, the buffer levels are not chosen only to handle disturbances in utilities, but to handle all disturbances at the site and to provide inventory of products to be sold to the market. This must be taken into account to evaluate if the buffer tank levels are appropriately chosen. The constraints from disturbances in utilities give one piece that has to be taken into account when choosing desired buffer tank levels.

If upstream disturbances also are taken into account, disturbances that affect a downstream area of a buffer tank, but not all upstream areas, will impose high-level constraints on some buffer tanks.

B. Control of the product flow

At the occurrence of a disturbance, a decision must be taken on how to control the product flow if the area that suffers a failure has more than one downstream area. Guidelines for how to control the product flow at a utility disturbance that affect an area, but not all its downstream areas, are obtained when using on/off production modeling including buffer tanks. Given the estimated disturbance duration, suggestions for the time that each downstream area should be run during the failure are obtained. The prioritization order of areas in the guidelines are determined from the profitability of the downstream areas. The suggestion is to let the operators at the site estimate the disturbance duration at the occurrence of a disturbance, and use this to compute the guidelines. The guidelines can be recomputed if the estimate of the disturbance time changes. Over time, contribution margins for different products could change, which makes it necessary to change the prioritization order of areas.

IV. Conclusions and future work

The case study at Perstorp presented in this paper gives ordering of utilities at the site according to an estimate of the loss of revenue they cause, using an on/off modeling approach with buffer tanks between areas. This list can be used to determine on which utilities improvement efforts should be focused. The case study also resulted in suggestions on how to choose the buffer tank levels and how to control the production at utility disturbances. It should be noted that only disturbances in utilities have been considered. This is only one piece of the entire picture, where also market conditions, cost of inventories and other disturbances must be taken into account. This case study shows which constraints disturbances in utilities place on buffer tank levels and product flow control.

The on/off production modeling approach including buffer tanks should give more accurate estimates of the losses that are caused by utilities at a site than the on/off model without buffer tanks. However, areas are still modeled as on or off, and thus the site model does not adequately reflect the actual production. To catch more of the variability, the site should be modeled using a continuous production model. Continuous production modeling of a site is currently being investigated, and will also be applied to the Perstorp site in Stenungsund. With continuous production, more elaborate reactive disturbance management strategies may be obtained, that gives real-time advise to operators on how to control the product flow at the occurrence of a disturbance.

Acknowledgments

This research was performed within the framework of the Process Industrial Centre at Lund University (PIC-LU), which is supported by the Swedish Foundation for Strategic Research (SSF).

References

[1] S. A. Papoulias and I. E. Grossmann, “A structural optimization approach in process synthesis–ı : Utility systems,” *Computers & Chemical Engineering*, vol. 7, no. 6, pp. 695–706, 1983.

[2] L. O. A. Maia, L. A. V. de Carvalho, and R. Y. Qassim, “Synthesis of utility systems by simulated annealing,” *Computers & Chemical Engineering*, vol. 19, no. 4, pp. 481–488, 1995.

[3] A. Lindholm, H. Carlsson, and C. Johnsson, “A general method for handling disturbances on utilities in the process industry,” in *proceedings of the 18^{th} World Congress of the International Federation of Automatic Control (IFAC), Milano, Italy*, pp. 2761–2766, 2011.

[4] A. Lindholm, H. Carlsson, and C. Johnsson, “Estimation of revenue loss due to disturbances on utilities in the process industry,” in *proceedings of the 22^{nd} Annual Conference of the Production and Operations Management Society (POMS), Reno, NV, USA*, 2011.

[5] D. D. Newhart, K. L. Stott Jr., and F. J. Vasko, “Consolidating product sizes to minimize inventory levels for a multi-stage production and distribution system,” *The Journal of the Operational Research Society*, vol. 44, no. 7, pp. 637–644, 1993.

[6] W. J. Hopp, N. Pati, and P. C. Jones, “Optimal inventory control in a production flow system with failures,” *International Journal of Production Research*, vol. 27, no. 8, pp. 1367–1384, 1989.

Regulatory Control of 4-product Kaibel column

Deeptanshu Dwivedi, [b] Ivar J. Halvorsen, [a] Sigurd Skogestad [a,b]
[a]*Department of Chemical Engineering, Norwegian University of Science and Technology, Trondheim, 7491, Norway*
[b]*Applied Cybernectics, SINTEF, Trondheim, 7491, Norway*

Keywords: Energy Efficient Distillation, Thermally Coupled Distillation, Control Structure Design, Divided wall column

Abstract

Conventional distillation remains the most widely used separation technique, but it is an energy intensive process. The energy consumption as well as capital costs can be greatly reduced by using direct material coupling. This leads to complex arrangements and difficult to operate systems like Kaibel column (Kaibel 1987) and Petlyuk arrangements (Petlyuk 1965). Three product divided wall column have been in use in industries and numerous applications have been reported. However, no works have been reported on separation of four product separations using such arrangements, which may offer further energy savings.

We have demonstrated the stabilizing operation of four-product divided wall column done on an experimental set up. We validate a decentralized four-point temperature inferential regulatory control (Strandberg 2006) structure experimentally. We test this control structure for steady state operation and subject it to several disturbances like feed rate change and test the individual temperature loops for setpoint changes. We also highlight some problems we faced in operating our system. Another challenge in the acceptability of divided wall applications is setting an optimal split of vapor between the two sides of the diving wall during operation. The vapor split is traditionally set by the natural pressure drop in the column. We provide a solution to this using a very simple vapor split valve to control a temperature in prefractionator and demonstrate the feasibility of active vapor split operation.

References

Kaibel, G. (1987). "Distillation columns with vertical partitions." Chemical Engineering & Technology 10(1): 92-98.

Petlyuk, F. B., Platonov, V.M. and Slavinskii, D.M. (1965). "Thermodynamically optimal method for separating multicomponent mixtures." International Chemical Engineering 5(3): 555-561.

Strandberg, J. S., Skogestad (2006). Stabilizing operation of a 4-product integrated Kaibel column. Institution of Chemical Engineers; 1999. 152: 636-647.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Robust implementation of optimal control policies for transient processes

Vinicius de Oliveira, Johannes Jäschke and Sigurd Skogestad

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

Abstract: We consider batch process optimization and robust implementation of optimal control policies. The dynamic optimization of such processes is in most cases model based, and therefore subject to uncertainties. This may lead to sub-optimal control trajectories with significant economical losses. Our goal is the development of simple rules which guarantee near-optimal operation under all conditions using feedback. Here, 'under all conditions' means for the defined disturbances, plant changes and implementation errors.

For processes whose economics are defined by the steady-state behaviour the concept of self-optimizing control was developed by Skogestad (2000). Self-optimizing control focus on selecting a set of controlled variables c that, when kept at constant setpoints, indirectly result in optimal economic operation in spite of disturbances. Diverse systematic methods are available to find the right variables to control for steady-state problems. Skogestad and Postlethwaite (2005) proposed the Maximum Gain Rule to select individual measurements. Alstad and Skogestad (2007) presented the Null Space method to select optimal linear combinations of measurements to be controlled. The Null Space method is very simple and yet gives zero economical loss if enough measurements are available and there is no noise.

In this paper we extended the Null Space method to optimal control of batch processes. The main idea is to compute a linear time-varying combination of measurements that should be controlled to zero by feedback. By doing so, the nominal control trajectories are optimally updated in case disturbances occur. The proposed method was tested in a semi-batch reactor case study and the performance compared with that of the neighbouring-extremal controller given in Gros et al. (2009). We found that both controllers give near-optimal performance for various disturbances. However, our approach is much simpler and intuitive and, most importantly, does not require complete state information.

REFERENCES

Alstad, V. and Skogestad, S. (2007). Null space method for selecting optimal measurement combinations as controlled variables. *Industrial & Engineering Chemistry Research*, 46, 846–853.

Gros, S., Srinivasan, B., Chachuat, B., and Bonvin, D. (2009). Neighboring-extremal control for singular dynamic optimization problems. I-single-input systems. *International Journal of Control*, 82(6), 1099–1112.

Skogestad, S. (2000). Plantwide control: The search for the self-optimizing control structure. *Journal of Process Control*, 10, 487–507.

Skogestad, S. and Postlethwaite, I. (2005). *Multivariable feedback control - analysis and design*. Wiley, Chichester, 2 edition.

[1] Corresponding author: Sigurd Skogestad *skoge@chemeng.ntnu.no*

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Simulation, Control and Optimization of Single Cell Protein Production in a U-Loop Reactor

Franck Guillaume Engoulevent * **John Bagterp Jørgensen** *

* *Department of Informatics and Mathematical Modeling, Technical University of Denmark, DK 2800 Kgs Lyngby, Denmark. (e-mail: jbj@imm.dtu.dk).*

Abstract: In 2011, the world population passed 7 billions inhabitants. While this number witnesses the success of humankind on earth, it also rises among other things questions about food supply. Declining live stock in the wild, rising price of energy combined with climatic change give a new economic potential for alternative sources of protein production.
Single cell protein (SCP) is protein produced by growth of micro organisms. Among these micro organisms, Methylococcus Capsulatus is particular interesting as it can grow on either methane or methanol and contains 70% protein. The U-Loop reactor is particular useful for production of SCP by M. Capsulatus as it has good gas-liquid mass transfer capabilities and also the capability to remove the significant amount of heat developed by the reaction.
In this paper we describe an implementation of a model to simulate SCP production in the U-Loop reactor. We report simulation results. In addition we design and compare different regulatory control systems for regulation of SCP production in the U-Loop reactor. The purpose of the regulatory control systems is to keep the process at a steady state and to reject disturbances. We design and implement such control systems based upon PID and MPC technology. In particular, we design these control systems such that they can be used as the regulatory layer in a process control hierarchy and enable resilient transition from one operating point to another. The optimal operating points are determined by the real-time optimization (RTO) part of the control system.

Keywords: Single Cell Protein, Simulation, Control, Optimization, PID, MPC, Real-Time Optimization

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

State Estimation for the Automotive SCR Process

Guofeng Zhou*, Jakob K. Huusom*,
John Bagterp Jørgensen**

**Department of Chemical and Biochemical Engineering, Technical University of Denmark 2800 Lyngby, Denmark (e-mail: guzho@student.dtu.dk; jkh@kt.dtu.dk).*
*** Department of Informatics and Mathematical Modeling, Technical University of Denmark 2800 Lyngby, Denmark (e-mail: jbj@imm.dtu.dk)}*

Abstract: Selective catalytic reduction (SCR) of NO_x is a widely applied diesel engine exhaust gas aftertreatment technology. For advanced SCR process control, like model predictive control, full state information of the process is required. The ammonia coverage ratio inside the catalyst is difficult to measure. Therefore we design an ordinary Kalman filter as well as an extended Kalman filter to estimate the ammonia coverage. The filters are built over a first principle model with four states. Among the four states, NO, NO_2 and NH_3 concentration are measured by the sensors, while the ammonia coverage ratio is left to be unknown. The performance of the filters is shown by simulation with the World Harmonized transient cycle. In particular, during transient operations the extended Kalman filter performs significantly better than the ordinary Kalman filter. Since such operation regimes are always present for SCR in engine applications, we recommend to estimating the ammonia coverage using the extended Kalman filter.

Keywords: Kalman filter, Extended Kalman filter, Selective catalytic reduction

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

State Estimators for a Pilot Anaerobic Digestion Reactor

Finn Haugen*. Rune Bakke. Bernt Lie***.**

**Telemark University College, Porsgrunn, Norway. (Tel: +47 35575000; e-mail: finn.haugen@hit.no).*
***Telemark University College, Porsgrunn, Norway. (Tel: +47 35575000; e-mail: rune.bakke@hit.no).*
****Telemark University College, Porsgrunn, Norway. (Tel: +47 35575000; e-mail: bernt.lie@hit.no).*

Abstract: A continuous flow pilot anaerobic digestion (AD) reactor of 220 L, located at Foss dairy farm in Skien, Norway, a is fed filtered cow manure as substrate for production of energy-rich biogas which contains mainly methane. The liquid effluent from the reactor is used as feed to a nitrification reactor to produce high quality fertilizer. The AD reactor study is presented here.

It is desirable to have a computer-based state estimator for the AD reactor for several reasons: (1) To have a methane gas flow estimate which is less noisy than what is obtained from raw measurements. (2) To have a methane gas flow estimate in situations where biogas concentration sensors and/or biogas flow sensors have an outage. (3) To have estimates of non-measured state variables for the purpose of monitoring and control. The most important state in this respect is the (total) concentration of volatile fatty acids (VFA). High VFA concentrations are inhibitory to methane generating microbes ("methanogens").

In this study a number of state estimators are applied to simulated data and measured data from the physical reactor. The estimators are ordinary and augmented Extended Kalman Filters and ditto Unscented Kalman Filters, and a ballistic (open-loop) estimator. All estimators are based on a modified version of a first principles model of the reactor with four state variables originally developed by D. T. Hill (1983). The model has been adapted to the AD reactor using a data set from online sensors and lab analysis. The estimators are compared with respect to accuracy of the means of the estimates, noise suppression in the estimates, robustness against model errors, estimator tuning, and implementation issues.

Keywords: State estimation; Kalman Filter; Unscented Kalman Filter; Noise; Anaerobic Digestion Reactor.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Stochastic Model Predictive Control with Applications in Smart Energy Systems

Leo Emil Sokoler [*,**] **Kristian Edlund** [**] **Tommy Mølbak** [**]
Niels Kjølstad Poulsen [*] **Henrik Madsen** [*]
John Bagterp Jørgensen [*]

[*] *Department of Informatics and Mathematical Modelling, Technical University of Denmark (DTU), 2800 Lyngby, Denmark (e-mail emil@sokoler.net, {nkp,hm,jbj} @imm.dtu.dk).*
[**] *DONG Energy, Overgade 45, Skaerbaek, 7000 Fredericia, Denmark (e-mail {kried,tommo}@dongenergy.dk)*

Abstract: In response to growing concerns related to environmental issues, limited resources and security of supply, the energy industry is changing. One of the most significant developments has been the penetration of renewable energy sources. In Denmark, the share of wind power generation is expected to cover more than 50% of the total consumption by 2050.
Energy systems based on significant amounts of renewable energy sources are subject to uncertainties. To accommodate the need for model predictive control (MPC) of such systems, the effect of the stochastic effects on the constraints must be accounted for. In conventional MPC, the stochastic effects on the constraints is handled by constraint back-off and the MPC problem can still be solved by solution of either a linear program or a quadratic program. Treating the constraints as probabilistic constraints provides a more systematic approach to handle the stochastic effects on constraints. In this formulation, the MPC may be represented by a chance constrained mathematical program. The chance constraints allow a direct tradeoff between a certain (low) frequency of violating the constraints and a performance function (e.g. an economic loss function). This is convenient for energy systems, since some constraints are very important to satisfy with a high probability, whereas violation of others are less prone to have a large economic penalty.
In MPC applications the control action is obtained by solving an optimization problem at each sampling instant. To make the controller applicable in real-time efficient and reliable algorithms are required. If the uncertainty is assumed to be Gaussian, the optimization problems associated with chance constrained (linear) MPC can be expressed as second order cone programming (SOCP) problems. In this paper, we show that tailored interior point algorithms are well suited to handle this type of problems. Namely, by utilizing structure-exploiting methods, we implement a special-purpose solver for control of smart energy systems. The solver is compared against general-purpose implementations. As a case study, we consider a system consisting of fuel-fired thermal power plants, wind farms and electric vehicles.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

On controllability of an integrated bioreactor and periodically operated membrane separation process

Prado-Rubio O.A. , Jørgensen, S.B. and Jonsson, G.

Department of Chemical and Biochemical Engineering, Technical University of Denmark (DTU), Lyngby, Denmark (e-mail: oap@kt.dtu.dk, sbj@kt.dtu.dk, gj@kt.dtu.dk)

Abstract:

Investigation of integrated processes involves challenges at both design and control levels, these can mainly be associated with different dynamic behaviors of the individual units plus their interaction. Therefore, the design and operation of the integrated system constitutes a key issue. In order to understand the controlled operation of the integrated process, it is convenient to use a model based approach supported by experimental evidence.

Recently, an integrated bioreactor and electrically driven membrane separation process (Reverse Electro-Enhanced Dialysis - REED) has been proposed as a method for intensification of lactic acid fermentation (Rype, 2003). This fermentation has been studied extensively driven by an increasing number of applications of the potential fermentation products. The main limitation of lactic acid bioproduction is that lactic acid bacteria normally are impaired by product inhibition at a certain lactate concentration level. Hence, productivity can be enhanced by the *in situ* lactate removal from the cultivation broth during pH controlled fermentation. This can be done by means of ion exchange membranes and electrical potential gradients. The novelty of the integrated process lies on the innovative REED technology, where lactate ions are exchanged by hydroxide ions. This allows the lactate removal and simultaneously facilitates the pH control in the fermenter. Long operation time is achieved by reversing periodically the polarity of the imposed electrical field to significantly reduce the influence of membrane fouling.

Previously, the REED and fermentation processes have been modeled and investigated separately (Prado-Rubio *et al.*, 2011a; Boonmee, 2003). Additionally, a simple quasi-sequential strategy for integrated process design and control structure development has been proposed (Prado-Rubio *et al.*, 2011b). The main purpose of this first attempt of process integration was to predict the productivity improvements and to reveal to which extend the REED module can facilitate the pH control in the fermenter. There, the membrane and reactor unit interactions are exploited to substantially increase the lactate productivity and substrate utilization compared to a conventional fermentation with a crude control of pH. Nevertheless, the proposed pH control structure is unable to tightly control the pH in the fermenter, which may result in a loss of productivity. The purpose of this contribution is to discuss the controllability of the integrated system, focused on the role of the REED module within the process. Interestingly, there are potential solutions either from process and control structure design such as: i. Account for the productivity enhancement earlier in the integrated process design, ii. Use multiple REED units activated sequentially or iii. Try to avoid the controllers fighting by a more appropriate control structure design. Hopefully merging those ideas, an improved strategy for the integrated process design and control development can be proposed.

References:

Rype, J. (2003). Modelling of Electrically Driven Processes. Ph.D. thesis, Technical University of Denmark.

Prado-Rubio, O.A., Jørgensen, S.B. and Jonsson, G. (2011a). Reverse Electro-Enhanced Dialysis for lactate recovery from a fermentation broth. *Journal of Membrane Science*, 374, 20–32.

Prado-Rubio, O.A., Jørgensen, S.B. and Jonsson, G. (2011b). Systematic Procedure for Integrated Process Operation: Reverse Electro-Enhanced Dialysis during Lactic Acid Fermentation. ESCAPE21, volume 29, pages 1406-1410. Great Britain: Elsevier.

Boonmee, M.; Leksawasdi, N.; Bridge, W. and Rogers, P. (2003). Batch and Continuous Culture of Lactococcus lactis NZ133: Experimental Data and Model Development. *Biochemical Engineering Journal*, 14, 127-135.

Proceedings of the 17th Nordic Process Control Workshop
Technical University of Denmark, Kgs Lyngby, Denmark
January 25-27, 2012

Model Predictive Control for an Industrial SAG Mill

Valeriu Ohan [*,**] **Florian Steinke** [**] **Michael Metzger** [**]
Thomas Runkler [**] **John Bagterp Jørgensen** [*]

[*] *DTU Informatics*
Technical University of Denmark, 2800 Kgs Lyngby, Denmark
(e-mail: jbj@imm.dtu.dk).
[**] *Siemens AG, Corporate Technology, Intelligent Systems & Control,*
80200 Munich, Germany
(e-mail:
{Florian.Steinke,Michael.Metzger,Thomas.Runkler}@siemens.com)

Abstract: We discuss Model Predictive Control (MPC) based on ARX models and a simple lower order disturbance model. The advantage of this MPC formulation is that it has few tuning parameters and is based on an ARX prediction model that can readily be identified using standard technologies from system identification. When applied to MIMO systems we call this controller a MIMO-ARX based MPC.
We use an industrial Semi-Autogenous Grinding (SAG) mill to illustrate the performance of this controller. SAG mills are the primary units in a grinding chain and also the most power consuming units. Therefore, improved control of SAG mills has the potential to significantly improve efficiency and reduce the specific energy consumption for mineral processes. Grinding circuits involving SAG mills are multivariate processes. Commissioning of a control system based on a classical single-loop controllers with logic is time consuming, while MPC has the potential to both improve the control performance and the commissioning time and expertise required.
The simulation results demonstrate that the MPC based on a MIMO-ARX model is able to provide nice control performance measured by its ability to track an output reference and reject unknown disturbances. Furthermore, the method used to design the controller represents a systematic method that can be automatized for wide-spread deployment in industrial environments.

Keywords: Model Predictive Control, ARX Model, SAG Mill, Mineral Processes, Industrial Process Control