Mariano Martín *Editor*

Alternative Energy Sources and Technologies

Process Design and Operation

Springer

Alternative Energy Sources and Technologies

Mariano Martín
Editor

Alternative Energy Sources and Technologies

Process Design and Operation

Springer

Editor
Mariano Martín
Department of Chemical Engineering
University of Salamanca
Salamanca
Spain

ISBN 978-3-319-28750-8 ISBN 978-3-319-28752-2 (eBook)
DOI 10.1007/978-3-319-28752-2

Library of Congress Control Number: 2016930057

Printed on acid-free paper

This Springer imprint is published by SpringerNature
The registered company is Springer International Publishing AG Switzerland

Prologue

In this book we aim to present the use of alternative sources of energy and technologies to produce fuels and power. We describe the value chain from harvesting the raw material (i.e solar, wind, biomass or shale gas), followed by the analysis of the processing steps into power, fuels and/or chemicals and finally the distribution of the products. We focus on the challenges of the supply, the variability of the source and its prediction, as well as the uncertainties related to it, the description of novel processes that are being developed and evaluated for their transformation, and the current results on the techno-economic analysis that are being reported in the literature. Moreover, in terms of the operation of such systems, energy distribution to the consumer and usage, and how we can integrate the new chemicals, fuels and power within the current system and infrastructure is evaluated. An example of the operation of a real renewable-based system, El Hierro island (Spain) is also discussed.

The view of the book is given from the process perspective and how a process systems approach can help in the use and integration of these sources of energy and novel technologies.

I thank all the authors for without their effort and commitment, this work would have not been possible. Special thanks are due to the reviewers of the chapters, who have provided useful ideas and suggestions and a different perspective to the various topics covered in this work.

Salamanca
November 2015

Mariano Martín

Contents

Part I Alternative Energy Sources

Nonconventional Fossil Energy Sources: Shale Gas and Methane Hydrates 3
Mariano Martín

Renewable Energy Sector 17
Leandro Real, Esperanza Sierra and Alberto Almena

Part II Infrastructure Design for Various Energy Sources

Development Planning of Offshore Oilfield Infrastructure 33
Vijay Gupta and Ignacio E. Grossmann

Emerging Optimal Control Models and Solvers for Interconnected Natural Gas and Electricity Networks 89
Nai-Yuan Chiang and Victor M. Zavala

Part III Processing of Alternatives Raw Materials

Equation-Based Design, Integration, and Optimization of Oxycombustion Power Systems 119
Alexander W. Dowling, John P. Eason, Jinliang Ma, David C. Miller and Lorenz T. Biegler

Wind Energy 159
Jason Ganley, Jie Zhang and Bri-Mathias Hodge

Solar Energy as Source for Power and Chemicals 181
Lidia Martín, Borja Hernández and Mariano Martín

Biomass as Source for Chemicals, Power, and Fuels 207
Mariano Martín and Ignacio E. Grossmann

CO_2 Carbon Capture, Storage, and Uses 235
Miguel Ángel Delgado and Fabrice Del Corso

Optimal Design of Macroscopic Water and Energy Networks 267
Ramón González-Bravo, Fabricio Nápoles-Rivera
and José María Ponce-Ortega

Part IV Operations

Retrofit of Total Site Heat Exchanger Networks by Mathematical Programming Approach 297
Lidija Čuček and Zdravko Kravanja

Improving Energy Efficiency in Batch Plants Through Direct Heat Integration 341
Pedro M. Castro

Life Cycle Algal Biorefinery Design 363
Jian Gong and Fengqi You

Planning and Scheduling for Industrial Demand Side Management: Advances and Challenges 383
Qi Zhang and Ignacio E. Grossmann

Industrial Tools and Needs 415
Iiro Harjunkoski and Hubert Hadera

Renewable-Based Self-sustainable Operation of Isolated Islands 439
María Cristina Rodríguez-Rivero

Part V Energy Distribution

Multi-objective Optimisation Incorporating Life Cycle Assessment. A Case Study of Biofuels Supply Chain Design. 465
María Augusta Páez, Fernando D. Mele and Gonzalo Guillén-Gosálbez

Large-Scale Stochastic Mixed-Integer Programming Algorithms for Power Generation Scheduling. 493
Kibaek Kim and Victor M. Zavala

Contributors

Alberto Almena Department of Chemical Engineering, University of Salamanca, Salamanca, Spain

Lorenz T. Biegler Carnegie Mellon University, Pittsburgh, PA, USA

Pedro M. Castro Faculdade de Ciências, Centro de Matemática Aplicações Fundamentais e Investigação Operacional, Universidade de Lisboa, Lisbon, Portugal

Nai-Yuan Chiang Argonne National Laboratory, Mathematics and Computer Science Division, Argonne, USA

Fabrice Del Corso Air Liquide, Jouy en Josas Cedex, France

Lidija Čuček Faculty of Chemistry and Chemical Engineering, University of Maribor, Maribor, Slovenia

Miguel Ángel Delgado Fundación Ciudad de la Energia, Cubillos del Sil (León), Spain

Alexander W. Dowling Carnegie Mellon University, Pittsburgh, PA, USA

John P. Eason Carnegie Mellon University, Pittsburgh, PA, USA

Jason Ganley Department of Chemical and Biological Engineering, Colorado School of Mines, Golden, CO, USA

Jian Gong Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL, USA

Ramón González-Bravo Chemical Engineering Department, Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Michoacán, Mexico

Ignacio E. Grossmann Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA, USA

Gonzalo Guillén-Gosálbez School of Chemical Engineering and Analytical Science, The University of Manchester, Manchester, UK

Vijay Gupta Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA, USA

Hubert Hadera BASF SE, Ludwigshafen, Germany

Iiro Harjunkoski ABB Corporate Research, Ladenburg, Germany

Borja Hernández Department of Chemical Engineering, University of Salamanca, Salamanca, Spain

Bri-Mathias Hodge Power Systems Engineering Center, National Renewable Energy Laboratory, Golden, CO, USA

Kibaek Kim Argonne National Laboratory, Mathematics and Computer Science Division, Lemont, IL, USA

Zdravko Kravanja Faculty of Chemistry and Chemical Engineering, University of Maribor, Maribor, Slovenia

Jinliang Ma AECOM & National Energy Technology Laboratory, Morgantown, WV, USA

Lidia Martín Department of Chemical Engineering, University of Salamanca, Salamanca, Spain

Mariano Martín Department of Chemical Engineering, University of Salamanca, Salamanca, Spain

Fernando D. Mele Facultad de Ciencias Exactas y Tecnología, Universidad Nacional de Tucumán, San Miguel de Tucumán, Tucumán, Argentina

David C. Miller National Energy Technology Laboratory, Pittsburgh, PA, USA

Fabricio Nápoles-Rivera Chemical Engineering Department, Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Michoacán, Mexico

José María Ponce-Ortega Chemical Engineering Department, Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Michoacán, Mexico

María Augusta Páez School of Chemical Engineering and Analytical Science, The University of Manchester, Manchester, UK

Leandro Real Energy Efficiency Department Director (KPMG—Inabensa (Abengoa)), Sevilla, Spain

María Cristina Rodríguez-Rivero Department of Engineering, University of Cambridge, Cambridge, UK

Esperanza Sierra Energy Efficiency Department Director (KPMG—Inabensa (Abengoa)), Sevilla, Spain

Fengqi You Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL, USA

Victor M. Zavala Department of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI, USA

Jie Zhang Department of Mechanical Engineering, University of Texas at Dallas, Richardson, TX, USA

Qi Zhang Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA, USA

Part I
Alternative Energy Sources

Nonconventional Fossil Energy Sources: Shale Gas and Methane Hydrates

Mariano Martín

Abstract The world has run on fossil fuels for the past 50 years. However, the current rates of demand for energy and the limited reserves have led to focus our attention on renewable sources, in search for sustainability, and nonconventional fossil fuels, shale gas and methane hydrates. Shale gas is already a reality, but hydrates represent an even larger reserve, accounting for 50 % of the carbon fossil sources. The advantage of unconventional fuels is that they are virtually the same as conventional ones. Therefore, the technology to process them into energy and chemicals is already well known and developed. On the other hand, the exploitation of these sources faces important technical challenges in terms of safe extraction and to avoid any environmental burden as a result of its recovery from the ground. In this chapter we present the availabilities of shale gas and methane hydrates, the technologies for its recovery, and the effect on the energy market.

1 Introduction

In spite of the current effort towards more sustainable production of energy, until 1800, biomass supplied more than 90 % of the energy needs of mankind. This position of dominance declined over 120 years due to the growth in the population and the industrial development. From the 1920s on, the easy access to crude oil and its simple transformation into ready-to-use fuels displaced the use of renewables. Lately, social concern has driven the effort toward development of renewable-based energy; see Fig. 1 (ExxonMobil 2013).

Over the past years there has been a steady increase in energy consumption. Although there is a strong commitment for renewables development to increase their share in the energy mix, the technology and infrastructure for fossil sources has improved over decades and thus, renewables are at a disadvantage. Therefore, it

M. Martín (✉)
Department of Chemical Engineering, University of Salamanca,
Plz. Caídos 1–5, 37008 Salamanca, Spain
e-mail: mariano.m3@usal.es

M. Martín (ed.), *Alternative Energy Sources and Technologies*,
DOI 10.1007/978-3-319-28752-2_1

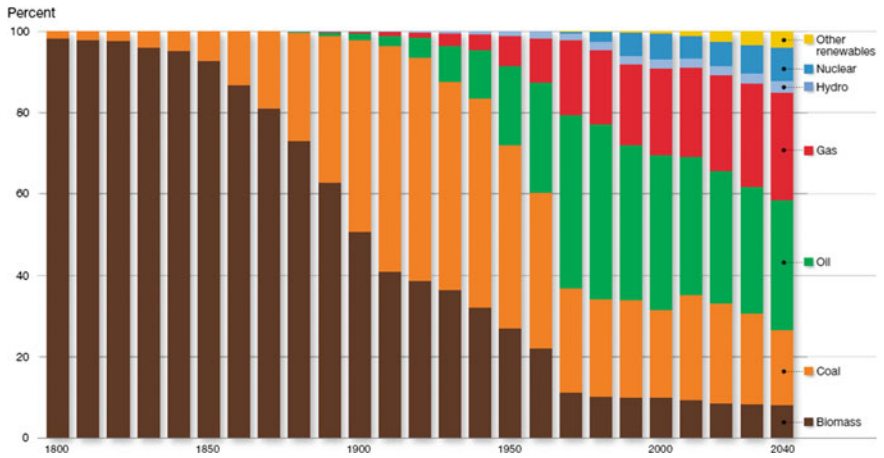


Fig. 1 2013 The outlook for energy a view to 2040. *Source* ExxonMobile (2013)

is convenient for the industry to continue using the infrastructure available and the transformation processes for power and chemicals production. Recently, two sources of nonconventional fossil fuels have focused the attention, shale gas and methane hydrates (MacDonald 1990). Actually, both represent alternative sources for natural gas, which can be easily integrated within the current supply chain.

2 Shale Gas

Shale gas is natural gas that is trapped within fine-grained sedimentary rocks known as shale formations. In spite of the recent media attention, shale gas has been extracted for more than a century. As early as 1821 in Fredonia, NY, shale gas was extracted from fractures in the ground. It was not until the 1920s that a technological breakthrough, the development of horizontal drilling and hydraulic fracking, accelerated the extraction of shale gas. The use of microseismic monitoring techniques certainly also contributed (Siirola 2014). The first fracking in the US can be dated back to 1947. However, the industrial exploitation of shale gas is more recent, around the 1970s. By that time, shale gas drilling was not profitable. Since then and together with the development in oil extraction, the Department of Energy and some companies were able to complete a successful air-drilled multi-fracture horizontal well in shale back in 1986. The first economically viable production of shale gas was due to Mitchell Energy, which in 1998 developed the slick water fracturing and shale gas started its contribution to the national economy of the U.S. (EIA 2012; Krauss 2009; USEIA 2011; Stevens 2012; White House 2009; Huges 2011; Wang et al. 2014). In Fig. 2 we see the availability of shale gas across the world. China, the US, Argentina, and Mexico have the largest reserves.

Fig. 2 Main shale gas reserves worldwide (US. Energy Information)

2.1 Gas Extraction

The extraction of shale gas is carried out using the technology known as fracking. It consists of horizontal drilling and hydraulic fracturing so that the rock is broken down using a fluid. The development of the two techniques, actually within the oil and gas industry, and their application to the particular case of the shale formations is what has enabled the extraction of shale gas. Figure 3 shows a 3-D image of the design of the drilling system. In a first step, vertical perforation is used to reach the shale formations. Once at that level, horizontal drilling begins. Next, water with a number of different compounds is injected. Thus, the rock is fractured and the gas can be extracted.

The aim of injecting a fluid at high pressure is to extend the fractures. This fluid also carries proppant when injected into the shale formation. The purpose of the proppant is to maintain a fracture open, avoiding damage in the formation or the well. Silica sand, resin-coated sand, or man-made ceramics are the most used. Two alternatives to transport the proppant can be used, either a high viscosity fluid or a high flow rate. The first one generates large fractures while the second one causes small micro fractures in the formation. The latter has become the method of choice. The fluid is not pure water, but a number of chemicals are added to efficiently deliver the proppant. The typical composition is 90 % water, 9.5 % is sand, and chemical additives up to 0.5 % (Ground Water Protection Council, GWPC 2009). Chemicals (3–13) (GWPC 2009; API 2010) are typically added including acids, NaCl, polyacrylamide, ethylene glycol, borate salts, sodium and potassium carbonates, glutaraldehyde, guar gum, citric acid, and isopropanol. Table 1 shows the purpose of each of the chemicals used. The injection of water containing these chemicals has arisen concern on the pollution of drinking water and reservoirs (Wang et al. 2014). Darrah et al. (2014) found that the fracturing of the rocks was not responsible for leaking but the construction of the wells that allows gas or oil leaks. The reservoirs can be found at around 3000 m while water resources are at

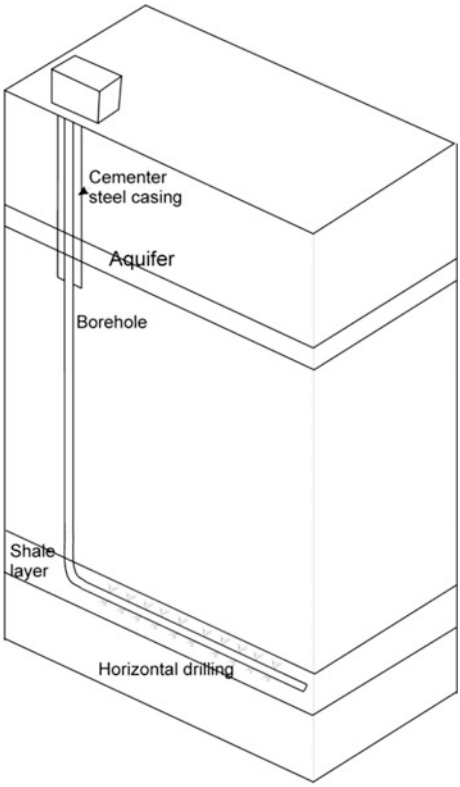


Fig. 3 Shale gas drilling and extraction (adapted from: MEPL 2015; BlogPreston 2011; Wang et al. 2014)

Table 1 Purpose of the chemicals

Acids	Cleaning
NaCl	Delays breakdown of the gel polymer chains
Polyacrylamide	Reduce friction pipe—fluids
Ethylene glycol	Prevents deposits in the pipe
Borate salts	Maintain fluid viscosity when the temperature increases
Na_2CO_3, K_2CO_3	Maintaining effectiveness of cross linkers
Glutaraldehyde	Disinfectants
Guar gum	Increases fluid viscosity
Citric acid	Corrosion prevention
Isopropanol	Increases fluid viscosity

about 200 m and the piping system typically reaches 350 m deep to avoid such contamination. However, Llewellyn et al. (2015) have reported a case where contaminants have migrated laterally through kilometers of rock affecting a potable water aquifer. Furthermore, there is another concern regarding water issues. A large amount of water is used to extract the gas. Ninety-seven billion gallons of water

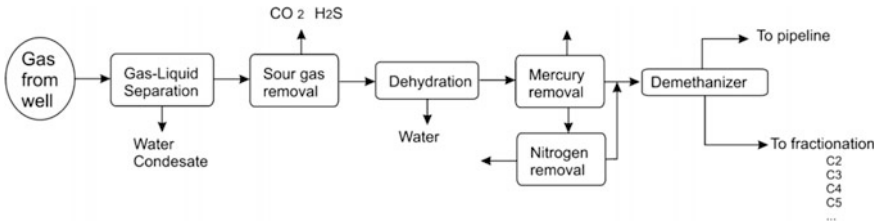


Fig. 4 Gas treatment at the well (adapted from Goellner 2012)

were used to frack the 39,294 oil and gas wells drilled from 2011 to 2014, most of them in regions with water scarcity (Ceres 2014).

The gas extracted contains a number of impurities that must be removed before injecting it into the natural gas pipeline. Figure 4 shows the treatment stages. First, the gas is separated from the liquid used for extraction. Next, sour gases such as CO_2 and H_2S are removed. Subsequently, the gas is dehydrated to remove the remaining water. Finally, mercury and nitrogen are removed. After this point, the natural gas is separated from the heavier fractions, the wet fraction. If the wet fractionation is economically feasible, then the shale gas can be used as a source of different chemicals. In particular, ethane is very interesting as a raw material for the production of ethylene, the major building block for polymers, such as polyethylene, polystyrene, PVC, etc., as well as intermediate for other chemicals like ethylene oxide, ethylene glycol, etc. (Siirola 2014).

2.2 *Composition*

The composition of the shale gas (mainly methane, ethane, propane, CO_2, and nitrogen) varies significantly from site to site of the actual location of the well, not to mention between different allocations. In Table 2 we present the composition

Table 2 Gas composition of different formations across the U.S.

Component	Marcellus	Appalachia	Haynesville	Eagle Ford
Methane	97.131	79.084	96.323	74.595
Ethane	2.441	17.705	1.084	13.824
Propane	0.095	0.566	0.205	5.425
C4+	0.014	0.034	0.203	4.462
Hexanes+	0.001	0.000	0.061	0.478
CO_2	0.040	0.073	1.816	1.536
N_2	0.279	2.537	0.369	0.157
Total	100.0	100.0	100.1	100.5
HHV (BTU/SCF)	1031.8	1133.2	1009.8	1307.1
Wobbe number	1367.1	1397.0	1320.1	1490.0

from different plays. In particular, the compositions were taken from shale plays that supply gas to the transmission network and end-users, including the Antrim, Barnett, Fayetteville, Haynesville, Marcellus, and New Albany plays as measured in 2009 (George and Bowles 2011; Bullin and Krouskop 2009). Wet shale gas is defined as the one that contains C2+ fractions, and dry the one that does not.

The use of this shale gas as natural gas is measured by comparing certain characteristic parameters such as the Wobbe number, the energy content, which is limited to 1,400 Btu/scf, the heating value, which should be around 1,100 Btu/scf, the upper limit on C4+ to be in 1.5 mol% and upper limits for inerts, 4 %. Apart from these specifications, the operator of the pipeline can decline to transport a certain gas if there is found risk to cause corrosion, declaring it to be "out-of-tolerance." Otherwise, the natural gases from shale formations are included in the natural gas supply chain. There are mainly four customers for the natural gas, the power sector, the industrial sector, commercial, and residential. In the US, they represent 34, 31, 14, and 21 % respectively (The Breakthrough Institute 2011).

Apart from a source of natural gas, the wet fraction provides ethane, propane, butane, etc. The separation of these components requires additional investment. Of particular interest is the ethane, which can be used in steam cracking units to produce olefins, ethylene, and propylene. These chemicals are the monomer of plastics such as polyethylene (PE) and polypropylene (PP) that have a wide range of applications (Siirola 2014). The production cost is the main advantage of ethane products in comparison to the production of the same chemicals from naphtha (Goellner 2012).

2.3 Shale Gas Production Cost

The costs per well range from $2 to $9 million in the US or Argentina versus $5 to $20 million in Europe (Philippe 2013) and the cost has decreased from the original $50 million (Cué 2015). For comparison, a regular well for crude oil costs $2 million. Further development over the next 5–10 years is expected to cut these prices by half. In terms of the price of the shale gas, in Europe it is expected to be between $6 and $15.5/MMBTU (Growitch et al. 2013), in the US values from $3.4 to 8.75 per MMBTU can be obtained depending on the place of production (Berman and Pittinger 2011; MIT 2009; Glickman 2014), while in China values of $11.2/MMBTU are reported (Glickman 2014).

3 Methane Hydrates

3.1 Availability of Methane Hydrates

Methane hydrates, or clathrates, are a cage-like lattice of ice inside of which molecules of methane are trapped, see Fig. 5. If the hydrate is either warmed or depressurized, it becomes a source of water and natural gas. Typically, one cubic meter of gas hydrate releases 172 m^3 of natural gas, when brought to Earth's surface (Ichikawa et al. 2007). The methane that has been trapped into ice has two main origins, biogenic, created by biological activity in sediments, and thermogenic, created by geological processes deeper within the Earth. Hydrate deposits may be several hundred meters thick and generally occur in two types of settings: under Arctic permafrost, and beneath the ocean floor on the continental margins at 500 m deep where nutrient rich waters are used by bacteria to produce methane, see Figs. 6 and 7. The interest in hydrates is not new. The petrochemical industry has been interested in them from the 1930s when they realized that the hydrates were blocking the oil pipes. Later in the 1960s, Russian drillers discovered the presence of hydrates in gas fields in Siberia. The first production field, Messoyakha, was discovered in 1968 just above the free gas reservoirs. In the 1970s methane hydrates were also found in deep waters. Actually the hydrates represent 50 % of all the hydrocarbon resources known in the world, see Fig. 8 (Kvenvolden 1993; Dillon 1992; Trehu et al. 2006; Boswell and Collett 2011).

3.2 Methane Hydrates Extraction

Hydrates occur at temperatures below 295 K and above 3000 kPa. However, the dissociation of the methane hexahydrate requires external energy. Therefore, three alternative technologies have been proposed either thermal recovery,

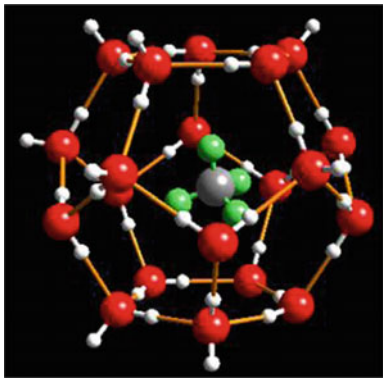


Fig. 5 Ice lattice for methane hydrates. http://woodshole.er.usgs.gov/project-pages/hydrates/primer.html

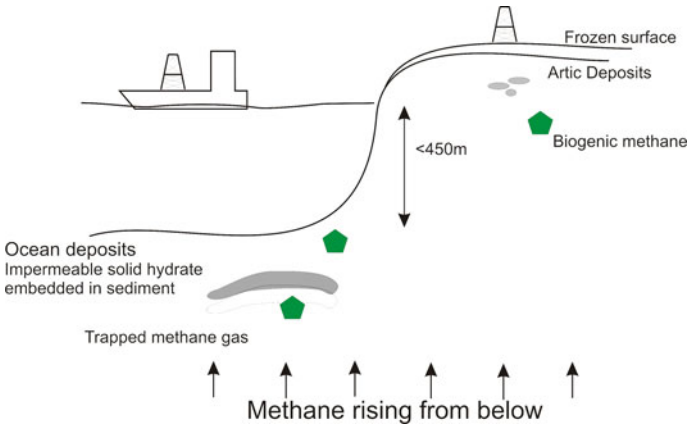


Fig. 6 Typical allocation of methane hydrates (adapted from http://energy.gov/fe/science-innovation/oil-gas-research/methane-hydrate)

Fig. 7 Allocation of the methane hydrates reservoirs (*Source* http://www.utexas.edu/news/files/Map-hydrates-canada1.jpg)

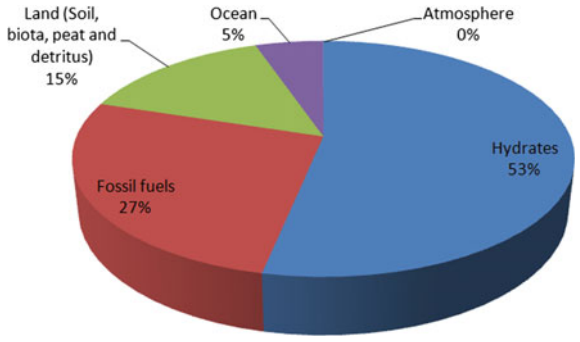


Fig. 8 Organic carbon availability by type

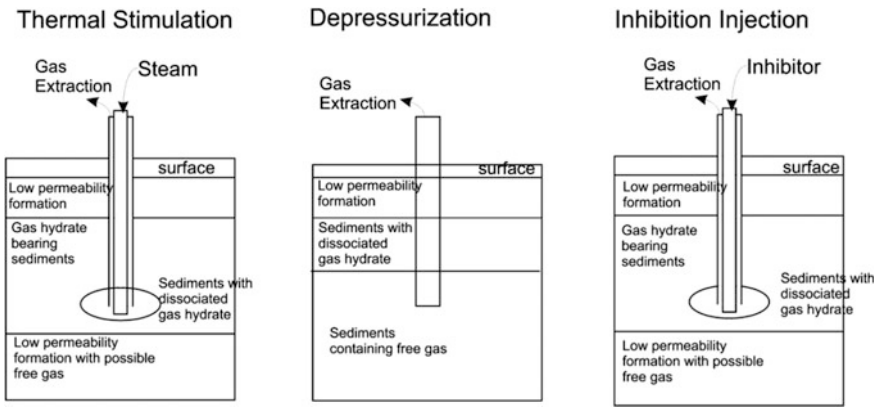


Fig. 9 Methods for methane from hydrates extraction (adapted from Ruppel 2007)

depressurization, or the use of chemicals (Xu and Ruppel 1999; Ruppel 2007, 2011), see Fig. 9 for a scheme representation of the three options.

The first technology consists of injecting steam, hot brine, or water to melt the ice so that the gas is released. The liberated gas ascends through the pipe. Although the method is simple, heating up the fluids to be injected is expensive and the fluid temperature can drop before reaching the reservoir due to the high- energy losses.

The second alternative is to relieve the pressure so that the gas flows out. This is carried out by drilling into the beds where gas methane can exist before forming hydrates. By changing the local pressure gradient within the sediment beds, the gases will flow to a wellhead. Three mechanisms govern the depressurization of the gas hydrates: (1) kinetics of dissociation, (2) conductive heat transfer, and (3) convective flow of fluids like gas and water. The gases are self-driven but it is less predictable than other methods. Furthermore, the dissociation of the hydrate is an endothermic process that can freeze the sediments and the machinery.

Finally we can inject chemicals such as methanol, monoethylene glycol and diethylene glycol, or brines. By altering the composition of the water, the hydrate becomes unstable and thus the gas can be released. Furthermore, the chemicals could lower the freezing point of the water in the surroundings, and the gases would be collected by the same wellhead. Actually, instead of methanol or other chemical we could use CO_2 to replace the methane in the hydrates. In this way, CO_2 is not only captured but is useful to recover the valuable methane. It is possible to exchange up to 64 % of the methane by CO_2. The process is exothermic, since the heat of formation of the CO_2 hydrate is higher than the heat of dissociation of the CH_4 hydrate. Thus, the exchange is thermodynamically favorable but the kinetics is slow (Ohgaki et al. 1996).

Although the technology is feasible, it is still not ready at industrial scale and it faces some important challenges such as leakage control and methane release to the atmosphere. For ease of exploitation of the various allocations where hydrates are found, the order could be (i) sub-permafrost hydrates, (ii) offshore accumulations

close to consumers, and (iii) remote offshore accumulations. Currently, the countries that are actively working on the development of technologies to extract this methane are Japan, South Corea, India, China, and the United States (Collect et al. 2008, 2009, 2011).

3.3 *Production Cost of Natural Gas from Hydrates*

The current cost of gas produced from methane hydrates is estimated to be US$30 to US$50 per million British thermal units (MMBTUs), which is an order of magnitude larger than the current Henry Hub price, around US$6 per MMBTU. The International Energy Agency estimates that once efficient practices and processes are developed, natural gas produced from methane hydrates will cost between US$4.70 and US$8.60 per MMBTU (Spalding and Fox 2014) becoming competitive and attractive.

4 Natural Gas Market. Effect of Unconventional Gas

Over the past years, the price of natural gas has oscillated within a wide range (George and Bowles 2011) as can be seen in Fig. 10. Note that the exploitation of shale gas began in 2009 resulting in a sharp decrease in the cost of natural gas followed by a cut in power prices.

As a result of the shale gas availability, the production of shale gas is projected to represent almost half of the total production of natural gas in the next 15 years, as presented in Fig. 11 (George and Bowles Liss 2012; Business Insider 2012). In this figure the contribution of the hydrates of methane is not included. Hydrates are

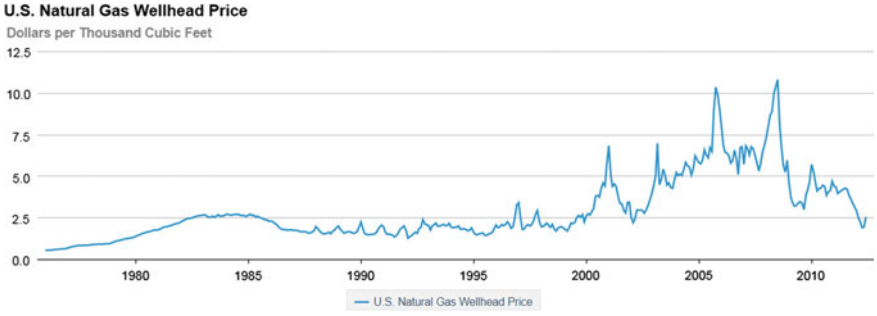


Fig. 10 Profile of cost of natural gas. *Source* U.S. Energy Information Administration (Oct. 2008)

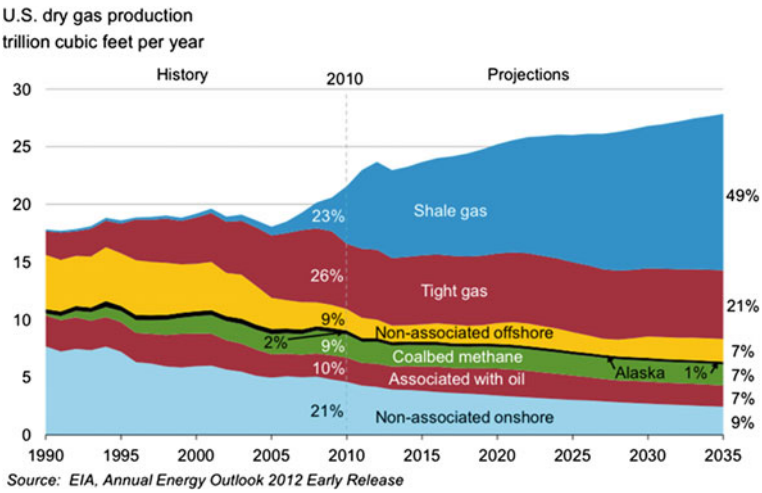


Fig. 11 Expected contribution of shale gas to the natural gas market (EIA 2012). *Source* U.S. Energy Information Administration (Oct. 2008)

beginning to be exploited commercially by Japan due to the need for energy sources, but no future estimation of the contribution to the market is available.

Furthermore, in the following years the delivered price is expected to increase from \$6.5/thous cu ft to \$9.5 thous cu ft in 2035 (EIA 2011), see Fig. 12.

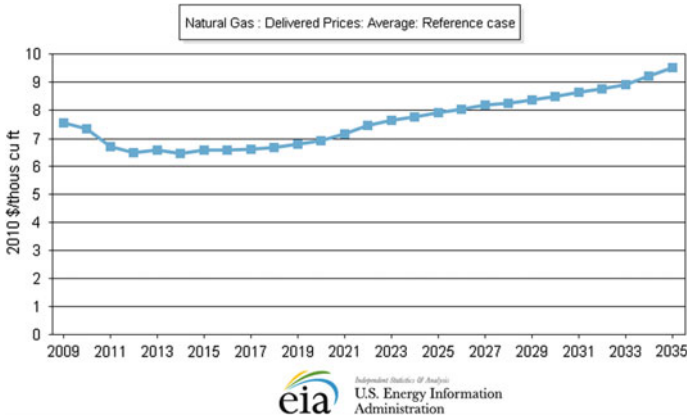


Fig. 12 Expected delivered prices for natural gas. *Source* U.S. Energy Information Administration (Oct. 2008)

References

American Petroleum Institute (2010) *Freeing Up Energy. Hydraulic Fracturing: Unlocking America's Natural Gas Resources*. http://www.api.org/policy/exploration/hydraulicfracturing/upload/HYDRAULIC_FRACTURING_PRIMER.pdf

Berman AE, Pittinger L (2011) U.S. Shale Gas.: Less Abundance, Higher cost. From The oil Drum http://www.theoildrum.com/node/8212

BlogPreston (2011) Environment vs. jobs: Lancashire's shale gas drilling investigated. http://blogpreston.co.uk/2011/02/environment-vs-jobs-lancashire%E2%80%99s-shale-gas-drilling-investigated/

Boswell R, Collett TS (2011) Current perspectives on gas hydrate resources. Energy and Environmental Science, 4: 1206–1215.

Bullin KA, Krouskop PE (2009) Compositional variety complicates processing plans for US shale gas, Oil and Gas Journal. 107, 10. http://www.ogj.com/articles/print/volume-107/issue-10/special-report/compositional-variety-complicates-processing-plans-for-us-shale-gas.html (last accessed March 2013)

Business Insider (2012) http://www.businessinsider.com/eia-shale-gas-2035-2012-1

Ceres (2014) Hydraulic Fracturing & Water Stress: Water Demand by the Numbers**. Boston Massachussets**

Collett T S, Johnson A, Knapp C. Boswell R (2008) Natural gas hydrates—a review *Natural Gas Hydrates—Energy Resource Potential and Associated Geologic Hazards* ed T S Collett *et al* (Tulsa, OK: American Association of Petroleum Geologists)

Collett T S, Johnson A, Knapp C. Boswell R (2009). Natural Gas Hydrates: A Review, in T. Collett, A. Johnson, C. Knapp, and R. Boswell, eds., Natural gas hydrates—Energy resource potential and associated geologic hazards: AAPG Memoir 89: 146– 219.

Collett TS, Lee MW, Agena WF, Miller JJ, Lewis KA, Zyrianova MV, Boswell R, Inks TL (2011). Permafrost associated natural gas hydrate occurrences on the Alaskan North Slope. Marine and Petroleum Geology, 28: 279–294.

Cue, C.E: (2015) Vaca Muerta, el sueño del oro negro argentino. EL PAIS. 16 June 2015.

Darrak TH, Vengosh A, Jackson RB, Warmer NR, Poreda RJ (2014) Nobel gases identify the mechanisms of fugitive gas gontamination in drinkin-water wells overlying the Marcellus and Barnett Sahes. PNAS, 111 (39): 14076–14081

Dillon W (1992) Gas (methane) hydrates. A new Frontier. United States Geological Survey. Sept 1002. http://marine.usgs.gov/Fac_sheets/gas-hydrates/title.html

Exxonmobile (2013) Outlook for energy 2013–2040 http://www.clingendaelenergy.com/inc/upload/files/2013_ExxonMobil_Outlook_for_Energy_.pdf

EIA (2011) Natural gas supply, disposition an prices, reference case (2011) http://www.eia.gov/oiaf/aeo/tablebrowser/#release=AEO2012&subject=0-AEO2012&table=13-AEO2012&region=0-0&cases=ref2012-d020112c

EIA (2012) Annual energy outlook 2012. Early Release http://www.eia.gov/forecasts/aeo/pdf/0383(2012).pdf

George DL, Bowles EB (2011) Shale gas measurement and associated issues. Pipe and Gas J. 238 (7) "http://www.pipelineandgasjournal.com/shale-gas-measurement-and-associated-issues?page=show" (last accessed March 2013)

Glickman N (2014) China's shale gas costs are at least double those in the us, but rising output will aid its bargaining position in world markets. Blomber. New Energy Finance. 29 May 2014

Ground Water Protection Council; ALL Consulting (April 2009) (PDF). Modern Shale Gas Development in the United States: A Primer (Report). DOE Office of Fossil Energy and National Energy Technology Laboratory pp. 56–66. DE-FG26-04NT15455. http://www.netl.doe.gov/technologies/oilgas/publications/EPreports/Shale_Gas_Primer_2009.pdf.

Growitsch C, Hecking H, John C, Panke T (2013) Unkonventionelles erdgas in Europa – Effekte auf Versorgung, Nachfrage und Preise bis 2035, EWI, Cologne, December 2013 http://www.ewi.uni-koeln.de/fileadmin/user_upload/Publikationen/Studien/Politik_und_Gesellschaft/2013/Unkonventionelles_Erdgas_in_Europa.pdf

Goellner JE (2012) Expanding the shale gas Infrastructure. CEP August 2012: 49–59

Hughes D (2011). "Will Natural Gas Fuel America in the 21st Century?" Post Carbon Institute http://www.postcarbon.org/report/331901-will-natural-gas-fuel-america-in May 2011

Ichikawa Y, Yonezawa T (2007): The Outline of the MH21 Program and the R&D Plan of Methane Hydrate Development System for Offshore Japan. http://www.jdc.co.jp/MH21Outline.pdf.

Krauss C (2009) New way to tap gas may expand global supplies. *New York Times*, 9 October 2009.

Kvenvolden K. (1993) Gas Hydrates_geological perspective and global change. Reviews of geophysics 31, 2: 173–187

Liss W (2012) Demand Outlook, A golden Age of natural gas. CEP August. 35–40

Llewellyn, G., Dorman, F., Westlland, J.L., Yoxtheimer, D., Grieve, P., Sowers, T., Humston_Fulmer, E., Branthley, S.L., (2015) Evaluating a groundwater supply contamination incident attributed to Marcellus Shale gas development. PNAS, 112 (20) 6325–6330

MacDonald G J 1990 The future of methane as an energy resource *Annu. Rev. Energy* **15** 53–83

MIT energy initiative (2009) MIT STUDY on the future of natural gas. https://mitei.mit.edu/system/files/NaturalGas_Appendix2D.pdf

Ministry of Energy of the Republic of lithuania (MERL) State budgetary institution (2015) http://www.enmin.lt/en/news/detail.php?ID=2180

Ohgaki K, Takano K, Sangawa H, Matsubara T, Nakano S, (1996) Methane exploitation by carbon dioxide from gas hydrates—Phase equilibria for CO2–CH4 mixed hydrate system," *Journal of Chemical Engineering of Japan*, 29 (3): 478–483

Philippe Q (2013) Europe's Comparative Disadvantage in Energy Intensive Industries: A Comparison of Shale Gas Production Costs and Break-Even Prices in Europe and the US, 2013:259–261.

Ruppel C, (2007) Tapping methane hydrates for unconventional natural gas, Elements, 3(3): 193–199.

Ruppel C, (2011) Methane hydrates and the future of natural gas, Supplementary Paper #4, The Future of Natural Gas, MIT Energy Initiative study, 25 pp. Staff (5 April 2011) World Shale Gas Resources: An Initial Assessment of 14 Regions Outside the United States US Energy Information Administration, Analysis and Projections, Retrieved 26 August 2012

Siirola J (2014) The Impact of Shale Gas in the Chemical Industry AIChE J. 60, 3: 810–819

Spalding D, Fox L (2014) Challenges of methane hydrate. Oil and Gas Financial Journal. May 7.2014

Stevens P (2012). "The 'Shale Gas Revolution': Developments and Changes". Chatham House. Retrieved 2012-08-15.

Trehu AM, Ruppel C, Holland M, Dickens GR, Torres ME, Collectt, TS, Goldberg D, Riedel M, Schultheiss P (2006), Gas hydrates in marine sediments: lessons from ocean drilling, Oceanography, 19: 124–143.

The Breakthrough Institute. Interview with Dan Steward, former Mitchell Energy Vice President. December 2011. http://thebreakthrough.org/archive/interview_with_dan_steward_for

White House, Office of the Press Secretary, *Statement on U.S.-China shale gas resource initiative*, 17 November 2009.

Xu, W. and C. Ruppel, 1999, Predicting the occurrence, distribution, and evolution of methane gas hydrate in porous marine sediments from analytical models, Journal of Geophysical Research, 104: 5081–5096.

Wang, Q., Chen, X., Jhan, A.N., Rogers, H., (2014) Natural gas from share formation- The evolution, evidences and challenges of shale gas revolution in the United States. Renew. Sust. Revs. 30, 1–28

http://www.ogfj.com/articles/print/volume-11/issue-5/features/challenges-of-methane-hydrates.html
http://energy.gov/fe/science-innovation/oil-gas-research/methane-hydrate
http://www.utexas.edu/news/files/Map-hydrates-canada1.jpg

Renewable Energy Sector

Leandro Real, Esperanza Sierra and Alberto Almena

1 Introduction

Some renewable energy sources have been used by humanity from the beginning of its existence, especially biomass, solar energy, wind energy, and hydraulic energy. We can find some examples for the traditional use of energy resources in sailing, in the windmills and watermills, or in the constructive dispositions of buildings to harness the solar power. However, the invention of the steam engine, at a time when there was no concern about the depletion of natural resources, supposed the end of these energetic uses.

Nevertheless, in the decade of the 70s, renewable energies had an upturn as an alternative to fossil fuels. Besides, the concern about the limited quantity of fossil fuels and their environmental impact, together with the energetic crisis of this decade, led to the appreciation of the renewable energies potential by both governments and researchers, in order to support the energetic demand.

This chapter contains a review of the historic evolution of each technology, from their rudimentary use in the past until nowadays. It also analyzes the improvement and the future potentialities of each kind of energy production.

L. Real (✉) · E. Sierra
Energy Efficiency Department Director (KPMG – Inabensa (Abengoa)),
Sevilla, Spain
e-mail: leadudo@gmail.com

A. Almena
Department of Chemical Engineering, University of Salamanca,
Plz. Caídos 1–5, 37008 Salamanca, Spain
e-mail: almena@usal.es

M. Martín (ed.), *Alternative Energy Sources and Technologies*,
DOI 10.1007/978-3-319-28752-2_2

2 Solar Thermal Energy

2.1 Low and Medium/High Solar Thermal Energy

At this point we are focusing on the historic development and the improvement in the efficiency of the current technology, used commercially for heating fluids up to low (<100 °C) and medium temperature (100 °C < T < 300 °C): the solar collectors.

It is difficult to specify exactly at what point humans started to harness the solar thermal energy to heat fluids, unlike the photovoltaic solar energy of which we actually know the date of discovery. We can say that solar thermal energy has always been used, to a greater or lesser degree. However, we can know exactly when this technology started to be developed for that purpose and to improve the efficiency of these processes has been improved.

The design of the first solar collector was found in 1767, when Horace de Saussure, a Swiss naturist, invented the self-named "hot box", comprising a box with its inside colored in black and insulation in all its walls, with the exception of the superior face that was made of glass. This device made it possible to achieve temperatures up to 109 °C. However, it was not until 1891 that the first patent of this invention was made, the water heater Climax, capable of heating water with solar power in an efficient enough way to ensure a wide use, especially in the sunny regions of the U.S.A. Later in 1909, an independent accumulation device was included into the design to allow the storage of the heat overnight, constituting the first thermosiphon system.

Unfortunately, the discovery of fossil fuels and the fall in prices of their production stopped the development of solar technology which, just like the other renewable energies, peaked coinciding with the subsequent energy crisis. It can be affirmed that the 70s' energy crisis marked a turning point in the evolution of the water heating technologies by solar power.

However, although the crisis was overcome and the prices of fossil fuels were controlled, the investment and development of new solar energy technologies continued, but in a slower way. Thus, this time the technological development was not suspended again.

Below, the most extensive technological solutions for fluid heating by solar energy at low, medium, and high temperatures are shown.

2.1.1 Low Temperature Systems

Flat Plate Collector

This kind of device is primarily used worldwide. Its technological development can be considered as optimal, but the research in this field is currently looking for new materials which could improve its efficiency, by providing better properties such a better heat insulation, an increase of greenhouse effect by using glass materials, a better design in the regulation and control systems, or by reducing the manufacturing costs.

The evolution of the solar collectors yield has been historically bonded to their market growth. This affirmation can be supported by two facts: the evolution from a craftsmanship to a mass production, causing a cost reduction and the application of quality systems; and the increase in the direct investment from different manufacturers, which allowed the product to be improved, become more competitive, and increase their market share.

In addition to this performance improvement of the solar collectors, also important progresses have been achieved in the auxiliary systems, such as an upgrade in the insulation of the storage systems, a more efficient pumping, the introduction of regulation and control systems, and the use of warning devices about failure and malfunction.

2.1.2 Medium/High Temperature Systems

Vacuum Tube Collector

This kind of collector has a less market share than the flat plate ones, due to their higher manufacturing cost and selling price. Despite the vacuum tube collectors have a performance range from 50 to 200 °C, their optimal efficiency is far from the standard use for domestic hot water production and heating applications, so they are used for this purpose only in those regions with low solar radiation, according to their greater uptake. Thus, these collectors are more useful in industrial processes, where they have to compete with other technologies that achieve higher temperature and maintain a better cost/yield ratio.

The efficiency improvement of this technology follows the same trend and constraints as those set for the solar collectors.

Cylindrical-Parabolic Collector

Talking about the historic development of fluid heating by the use of solar power, it is worth mentioning Mouchot's steam engine. This technology consists of a large parabolic receiver covered with mirrors, which concentrates the radiation at only one point. The heat generated by this device is able to activate a steam engine. This invention can be considered as the precursor to the solar thermal energy concentration technology, in which the parabolic collectors are based. This device can heat a thermal fluid by concentrating the solar energy at one point (as the Mouchot's machine does) or in one line.

The first installation of a parabolic collector was in 1912, being its purpose the generation of steam for the operation of a water irrigation pump. Furthermore, there are commercial plants using this technology since the 80s.

Efforts on increasing the performance of this technology focuses on the improvement of the reflection and refraction of the collectors and the search for the internal fluid with the best features. In addition, there are also auxiliary fields that

can be improved for this purpose, such as the control systems, the performance of the electric energy production system, the solar tracking, etc.

2.2 High Temperature Solar Energy. Solar Thermal Power Plants

This kind of technology has only been developed in the modern age, existing solar energy plants are based on in this knowledge since the 60s. However, the first commercial plant in Spain came out in 2007. The PS10 was built in Seville, designed with a power of 11 MW. Thereafter, solar thermal power plants have seen an exponential growth and they have been developed with improved attached technologies to enhance their yield, becoming a big contender for conventional plants, thanks to their versatility.

Unlike low and medium temperature applications, which have a simpler process of heating fluids up to moderate temperatures, high temperature solar systems are based on more complex processes to heat the working fluid up to higher temperatures to produce energy; hence there are more elements involved in the whole process, namely turbines, engines for heliostats, steam generators, etc. This technical complexity, together with having more determining factors, has hampered innovation in this sector, while the low and medium temperature applications may have led to the development of different technologies like flat plate collectors, vacuum tube collectors, parabolic-cylinders, thermosiphon, or forced systems. Thus, the architectural designs of these plants and the original building solution have not been altered significantly, so that the successful technological innovation is focused on the improvement of the process itself, and not in the development of attached technologies. Some examples of technological improvement of these plants are shown below:

- *Storage of molten salts* the first high temperature solar thermal power plants just heated a fluid with the radiation of the sun, stopping the electric power generation when the solar radiation ceased. This supposes a nonmanageable energy production, which means a competitiveness loss related to conventional power plants.

  The first molten salt thermal storage plant in Spain was built in Seville, in 2011. Thus, it supposes the first substitution of the circulating fluid to a molten salt fluid on a commercial scale. The new fluid, besides having a circulation behavior similar to the substituted one, also has higher specific heat and allows easy storage of the salts. Thanks to this innovation, the energy production becomes manageable and can therefore adapt the production to the demand during sunny hours, as well as keeping the energy production when there is no sunlight. This fact notably improves its versatility and ability to compete with conventional technologies.

- *Improvement of the constructive quality for the heliostats* the search for better efficiency of the heliostats focuses on the improvement of their structural design. In the same way as photovoltaic solar energy, the research in this field has served and serves as a spearhead for this sector, but oriented toward aerospace applications.
- *Improvement of the heliostats disposition and monitoring* thanks to the tests carried out at pilot solar stations, a specific software is currently available for designing and calculating different configurations for heliostat solar fields, determining the most efficient disposition for each case. Thus, it is energy efficiency improvement of the plant achieved by the optimization of the heliostat field configuration.
- *Control system upgrade* due to the process complexity and the amount of constraints and elements involved, the improvement of the control systems design must increase the process efficiency. However, current solar power plants are far from excellent in this area. Nowadays, it is not possible to design a solar power plant that is completely automated, with the resulting loss of efficiency due to manual control for such complex technological applications. It is looking for innovation in this regard, but there is still much room for improvement.
- *Other improvements* test of materials and constructive methods have allowed achieving higher temperatures in the central receiver without damaging the material. In addition, other improvements not directly related to the process of capturing solar energy, but to the whole process, such as those related to the steam generation or to the turbine performance, have contributed to the development of these technologies efficiency.

2.3 Other Solar Thermal Technologies

Apart from the technologies mentioned above, there are other technologies without extensive development and/or commercial use like parabolic disk systems, individual units functioning as solar concentrators comprising a parabolic disk that concentrate solar radiation in a receiver to heat a fluid or a gas up to 750 °C and producing electricity by using a turbine. Another technology that fits in this section is the system of compact linear Fresnel reflectors, which is based on a set of almost flat reflectors that concentrate the solar radiation onto elevated inverted lineal receivers, generating steam.

Literature and Sources

Web pages:

- Renewable energies companies association (In Spanish). http://www.appa.es/
- Solar site. Renewable energies portal (In Spanish). http://www.sitiosolar.com
- I+D Energy (In Spanish). http://idenergia.com

- Ministry of Industry, Energy and Tourism. Govern of Spain (In Spanish). http://www.minetur.gob.es/
- Renewal Energy National Center (In Spanish). http://www.cener.com/
- Centrales Termosolares. http://www.centralestermosolares.com

3 Photovoltaic Solar Energy

The photovoltaic effect was discovered by the French physicist Alexandre Edmond Becquerel in 1838. Becquerel was experimenting with an electrolytic cell of platinum electrodes, when he found that the electric current rises in one of the electrodes due to sunlight exposure.

Later, in 1873, the electric engineer Willoughby Smith found the photovoltaic effect on selenium. Then, the first selenium photovoltaic cell was created in 1877 by William Grylls Adams, an English professor of natural philosophy in the King College of London, working together with his pupil Richard Evans Day.

However, it was not until 1953 that the first electrolytic cell, with an adequate yield to consider a commercial viability for this technology, was manufactured. Investigators D.M. Chaplin, C.S. Fuller, and G.L. Pearson from Bell Laboratory created the first silicon photovoltaic cell and made its official presentation in that year. Nevertheless, despite the technological progress achieved in increasing the photovoltaic cell performance, the cost of this technology at the historical moment was too high and prevented their practical application. An outstanding data is that the cost of an electricity watt produced in conventional power plants in 1956 was set around 50 cents, while that produced by photovoltaic panels reached 300 dollars. This fact avoided the use of this technology as a great quantity electric power supplier.

Since then, the research has been focused on the reduction of the production cost of photovoltaic energy (dollar/kWh), achieving parity and even advantage over the price of conventional energy sources, which were six times cheaper than electricity that came from photovoltaic technology at its early stages, as seen above. This is, thanks to the technological progress and the scale of economy that this technology has undergone, taking into account that the value of the cost also depends on the generation conditions and the country of production today.

The main application that boosted, without any doubt, the technological improvement of this sector was the powering of the space satellites equipment in the incipient space race at the time. In 1955, the United States ordered its industry the task of producing photovoltaic panels to space applications, where the reliable supply of electric power in areas of difficult access is essential. Thus, this year the Hoffman electronic company provided 14 mW photovoltaic cells with a yield of 3 % and a price of 1500 $/W. Two years later, this company developed solar cells with a performance up to 8 %.

In this way, despite the initial misgivings from some directors of NASA, the use of photovoltaic solar technology ended, surpassing other alternative mechanisms

for powering satellites such as chemical batteries and nuclear energy. Then, the first photovoltaic powered satellite was launched in 1958. The Vanguard I counted with solar panels as backup power, but eventually became the main power source when batteries, first designed as the main power supplier, were sold out after only 20 days. The device was operational with this configuration for 5 years.

We conclude there was a feedback between the photovoltaic solar energy and the space race. It is true that this energy source had a development boost due to its use in the space race, but the last one could not be possible, in terms we know it nowadays, without the support of the photovoltaic powering.

However, even at this point, solar panels were still too expensive in terrestrial environment to be competitive with conventional energy sources. This issue changed when, in the 70s, Dr. Elliot Berman created a much cheaper solar cell, reducing the cost per watt. The consequent fall in prices allowed using the photovoltaic panel in facilities connected to the mains, starting to be the installation of this device more profitable than tracing an entire cable line or making the periodical maintenance and replacement of batteries. Thus, the practical applications of this technology multiplied, such as communication signal repeaters, lighting railways, marine lighting buoys, lighthouses, etc.

Today, the technological diversity is very wide and varies depending on the material of the panel, the connection type, regulation, etc. We focus on conventional flat panels and electronic inverters, the more widespread current commercial solutions. These modules can be classified according to their comprising cells as monocrystalline, polycrystalline, and amorphous. Their efficiency is greater due to the bigger size of comprising crystals, just like their price, weight, and thickness. The current yield average varies from 20 % of the monocrystalline modules to 7 % of the amorphous ones.

Important progress has been achieved in the manufacture of photovoltaic panels in the recent years. The research in this field has been focused mainly on the reduction of the panel manufacturing cost, the efficiency improvement, and the study of other materials. A summary of the achieved milestones is given below:

- *Reduction of the panel manufacturing cost* thanks to the economy of scale and the improvement in the production methods, it has been possible to reach parity in the mains of more than 100 countries. New materials showing applicable performances and reducing manufacturing costs are currently researched.
- *High efficiency of the photovoltaic effect* in 2014, in laboratory scale, a new record in the conversion efficiency was achieved, reaching 40.4 %. Today, it is not possible at a commercial scale yet.
- *Improvement of electronic invertors performance* the progress in their constructive elements and managing software have allowed to get, among other advances, higher frequencies, the capacity of working in a higher range of temperature with less losses, increase of efficiency at partial shadowing, or a better use of the received radiation when it substantially differs from one module to another at first and last day hours.

- *Study of other materials* for example, organic solar cells, carbon nanotubes, pyrite, magnetite, molybdenite, or graphene. The current use of chemical elements like gallium or silicon in solar panels increases the investment, causing multiple parallel researches looking for substitute materials.
- *Monitoring systems* the inclusion and development of these systems improve the process efficiency. Being able to measure and compute production and solar power consumption data at real-time supposes an advantage and a breakthrough, helping to prevent possible failures in the system by remote warnings about possible anomalies.
- *Other progresses* for example, an efficiency enhancement in adverse weather conditions, the use of flexible materials, the inclusion of control systems for the use in charge mode, and many more research areas.

Investment in photovoltaic solar power development and future forecast are promising. An application of this technology can be found in many areas, such as solar roads capable to charge vehicles that circulate over them, solar photovoltaic cell sprays with application in every surface, printable solar cells on clothing, the use of nanotechnology to improve yield, the synthesis of materials for reducing cost over an estimated 80 %, the enhancement of efficiency by holography solutions, etc. As it can be seen, there is a growing range of possibilities that only need a strong commitment by the regulatory entities. See Chap. 7 for further discussion on the use of solar energy to power and fuels production.

Literature and Sources

Books:

- **ELIAS CASTELS, Xavier; BORDAS ALSINA, Santiago: Energy, water, environment, territoriality and sustainability. In Spanish. Díaz de Santos, 2011.**

Web pages:

- Renewable energies companies association. In Spanish. http://www.appa.es/
- Eurelectric. Electricity for Europe. http://www.eurelectric.org/
- European Climate Foundation. http://europeanclimate.org/
- Renewal Energy National Center. In Spanish. http://www.cener.com/
- National Energy Club. In Spanish. http://www.enerclub.es/
- Solar site. Renewable energies portal. In Spanish. http://www.sitiosolar.com

4 Biomass

4.1 The Use of Biomass

Biomass has been the first source of power for humans and the main fuel until the Industrial Revolution and the beginning of the fossil fuels development. The last ones, counting with a greater calorific value and together with the need for

increasingly smaller spaces for the use of fuels, had reduced the uses of biomass to record lows. See Chap. 8 for further discussion on current uses to biofuels production.

4.2 Biogas

This is an alternative energy that has been used for centuries, just like the low temperature solar energy. Biogas is currently used widely all over the world, especially in those areas that do not have reserves of fossil fuels.

The first evidence of the use of biogas came from the tenth BC, in Mesopotamia, where it was used for heating water, although there are historical overviews that show the use of biogas for residue cleaning by Sumerians.

In our era, the Roman scholar Pliny the Elder described the brightness of some lights that appeared below the marshland surface, around the year 50 AC. A long time after Benjamin Franklin referred to biogas as that responsible for a fire in New Jersey and then in November of 1776, the Italian scientist Alejandro Volta published, in a letter called "Aria inflammable native delle Paludi," how an explosive gas was formed in Lake Como when the sediments were stirred. Volta also concluded there was a direct correlation between the amount of decaying organic material found in the mass of water and the volume of flammable gas produced, and he established methane as the main compound of natural gas. The importance of these results was recognized by the scientific community of the time, reflected in the fact that this letter was translated into German only 2 years after its publication. Later in 1804, John Dalton described the chemical structure of methane and related it with biogas, while in 1821 Amedeo Avogadro set the final chemical structure of methane for the first time.

In the mid-nineteenth century, it is important to mention the research carried out in France about removing bad smell from wastewater, which was supposed to be a milestone in biogas technology development. These experiments allowed discovering some of the microorganisms that today are considered essential for the fermentation processes.

In modern times, we have to go to 1859 to review the construction of the first digester in the asylum-hospital of leprosy in Matunga, near Mumbai. The plant worked at purifying wastewater and supplying energy in emergency cases. At the end of the century, Louis Pasteur carried out several researches on biogas formation from animal wastes and suggested its use for streetlights. Then, between 1895 and 1896 in Exeter, United Kingdom, street lamps started to be powered by the gas collected from digesters that fermented sewage sludge. At this time, the first biogas plants, similar to the current ones, were built in China.

Until then, biogas was considered as a process with poor energy performance and without any parallel technological development. A biogas plant comprised a storage tank where the material was introduced without further elaboration than a

closure to ensure certain tightness, just like a covering fabric, avoiding the entire leak of the produced gas when this material decomposed.

In the 30s of the twentieth century, the search for the energy efficiency of these applications started when the development of microbiology allowed identifying anaerobic bacteria and better conditions for the enhancement of methane production. The operating plants at this time were kept rudimentary or nonexistent, but the scientists began to understand being understood the reaction performance of the biogas production process and the use of catalyst was started, improving the efficiency by producing more volume of biogas in less time.

Thus, since then it is difficult to talk about innovation in this field, looking for an increase in the energy efficiency in this field. The studies and breakthroughs carried out in the past years are aimed at finding out which kind of substances can be used and in what quantity, to maximize the production, the search for the optimal mixtures and the best catalyst for the process, the study of material resistance against gas corrosion, etc. This is due to the relatively simple technology and construction of the biogas production, that makes unattractive the development of new technologies. The rest of the progress is therefore aimed at reducing the construction cost of the plants by achieving better constructive processes or the development of an economy of scale.

The most important improvements in the production of biogas are shown below:

- *Efficiency enhancement of the chemical process* maximizing the production by using the best substances in their optimal quantity, achieving the most productive mixtures, using the catalyst with the best performance, determining the optimal temperature in each step of the process, etc.
- *Improvement of the accumulation system* commercial biogas plants include hermetic accumulators to produce biogas in optimal conditions and without any leak. Some of the added breakthroughs are the heating and cooling systems for achieving the optimal temperature in each process step, the installation of agitators/aerators to move the decanted substances and maximize the biogas production or the equipment of inlet and outlet gates to assure a continuous process.
- *Improvement of the control system* today we cannot talk about totally automated biogas plants, but integrated services are increasingly included in the design with the consequent reduction of manual control and an efficiency improvement.

On the other hand, it should be pointed out that there is a difference between agricultural biogas source and waste biogas source.

4.2.1 Biogas from Agricultural Holding

There is still a lot of work for the expansion of this technology, which today has more interest from the point of view of sustainability than from achieving parity with the mains. Although there are many facilities all over the world, most need subsidies for viability and their objective is not mainly electricity generation. This

kind of plants takes into account other purposes such as having economic benefits by including residues with inerting cost, getting environmental benefits, the zero waste policies or the lack of other energy sources in rural areas and developing countries.

4.2.2 Biogas from Wastes

Here are included rubbish dumps, wastewater treatment plants, and any regulated source of residues. In this case, there is neither movement nor elaboration costs; these existing processes already produced biogas by themselves and thanks to the increasing electric power price, together with the strict regulations for the gases released to the atmosphere, has made the implementation of this technology viable. For example, in the US, biogas plants have only been installed in 3 % of agricultural facilities considered as viable, while 33 % of wastewater treatment plants and 58 % of rubbish dumps have been covered.

Literature and Sources

Books:

- **LOBERA J. History of biogas. In Spanish. Metabioresor, 2011.**
- **CERDÁ, E. Biomass in Spain: A renewable energy source with a great future. In Spanish. Foundation Ideas, 2012**
- **SEPA, Guidance on gas treatment technologies for landfill gas engines**

Web pages:

- European Climate Foundation. http://europeanclimate.org/
- National Energy Club. In Spanish. http://www.enerclub.es/
- Renewal Energy National Center. In Spanish. http://www.cener.com/
- Ministry of Industry, Energy and Tourism. Govern of Spain. In Spanish. http://www.minetur.gob.es/

5 Wind Energy

We can consider wind power as a variant of solar energy as it comes from the warming of the atmosphere and landform variation, so its use dates back to several millennia. First, wind power was essential in sailing, evidence of which exists in Egyptian drawings up to 5,000-years old showing the sailing boats used to cruise the Nile River. Later, 4,000 years ago in Babylonia and China, windmills were used to lift water for irrigation, which were the first wind turbine models.

The first wind machines date from the sixth century, which were of vertical axis type used for grinding and water pumping in the region of Scythia, situated between Iran and Afghanistan. After that, especially in the Mediterranean Greek islands,

horizontal axis windmills were developed, whose main feature was their blades were triangle sails. They are still used today to grind grain in the Greek island of Mikonos.

In the eleventh century windmills were widely used in the Middle East, but they were not introduced in Europe until the thirteenth century, as a consequence of the Crusades. Many of them were built during the Middle Ages and were of great importance, reaching to the extent that feudal lords reserved to themselves the authorization of their construction, as a way of forcing the subjects to grind grain in their own windmills and taking a part. Even planting trees near them was forbidden to ensure a good incidence of the wind.

Next, in the fourteenth century, the most important progresses in windmill technology were achieved in Holland and they started to be used extensively for draining water from the marshy area of the Rhine Delta. Subsequently, at the end of fifteenth century, the first windmills were built for producing oil, paper, and processing wood in sawmills. The number of these facilities kept growing, achieving in the middle of the nineteenth century over 9000 windmills operating for different purposes. However, the introduction of the steam engine after the Industrial Revolution supposed a negative effect for this technology and less than a thousand windmills kept working during the middle of the twentieth century.

Another country where wind power was widely used is Denmark, where about 3000 windmills were used for industry purposes and nearly 30,000 for domestic use in houses and farms. Nevertheless, as in other regions, the emergence of cheaper alternatives for power supply made them to be gradually replaced by thermal machines or electric engines fed by the mains.

As noted above, this technology disappeared from the energy supply issue and was replaced in navigation by steam machines and post technologies. However, in 1802, Lord Kelvin coupled an electric generator and a wind machine and later in 1850, once the dynamo was discovered, the wind turbine was created. Then, Charles F. Brush designed in 1888 the first wind turbine for electricity generation and, around 1890, the Danish meteorologist Paul la Cour started up the first machine specifically designed for electricity generation from wind power.

Despite the drop of this technology in the nineteenth century, as has been observed in previous cases, the energy crisis in the 70s took into account the wind power potential and boosted its progress. Thus, German and Danish engineers developed the first wind turbines with a profitable commercial use for this kind of technology.

The main component of these facilities is the wind turbine or aerogenerator, which we take as a whole due to its complexity. It is designed to receive the impact of wind and is connected to a dynamo that provides electricity. The innovation efforts in this technology focus on these two points, together with achieving more efficient constructive processes. The most widely traded model is the upwind three-bladed horizontal axis turbine, although there are some alternatives and the engineers keep looking for other configurations. We show some of them below:

- *Darrieus wind turbine* it was patented by the French aeronautical engineer George Darrieus in 1931 and commercialized by the American company

FloWind until its bankruptcy in 1997. This turbine is a vertical axis type and consists of a number of curved aerofoil blades mounted on a vertical rotating shaft or framework. The main advantage of this technology is that guidance systems are not needed, but it is less efficient than a horizontal axis turbine, requires starting assistance, and receives less wind as a consequence of being low to the ground.

- *Single-blade, two-blade, three-blade, and multi-blade wind turbines* The first aerogenerators counted with a big number of blades. It has been reduced until the three they generally have, being the least number of blades that provides more stability and simultaneously allows material and weight savings. Some models use two-blade and single-blade rotors, achieving even greater savings, but resulting less efficient and more complex control systems for stability improving must be included.
- *Leeward rotor turbine* Wind turbines usually have their rotor located upwind, in front of the nacelle, avoiding that any element could stop the wind or create turbulences. However, there are also models whose rotors are located leeward, where its blades are at the rear of the nacelle. This design could be interesting in small machines to make the housing of the nacelle work as a vane, in order to guide the turbine to the wind direction without requiring other devices.

The research in this technology is focused on reducing the size of the engines, improving the reliability of the machines, increasing their power and achieving a better integration to the mains. In 1980, the average power of a wind turbine was 50 kW developed by a rotor of 15 m diameter, while in the year 2003, the average power was a hundred times greater with an average diameter of 124 m.

Today, despite the innovation in this technology is still growing, the research is more focused on offshore wind facilities, which still represent a small ratio related to the whole power installed in the world. The first offshore wind farm, Vindeby, was built in the Baltic Sea in 1991 and it was composed of 11 turbines of 450 kW. A decade later, in 2002, after the start-up of many wind farms of different power, was built in Denmark one of the greatest offshore wind farms of today. The Horns Rev counts 80 wind turbines developing together a power of 160 MW.

The main disadvantage of this technology is the high investment. Although nowadays its value is declining, it is still much higher than the onshore facilities, due to logistic and installation issues. We refer the reader to Chap. 6 for current development of wind as a source of energy.

Literature and Sources

Books:

- **BURTON, T.; SHARPE, D.; JENKINGS, N.; BOSSANYI, E.: Wind Energy Handbook. Chichester, UK: John Wiley & Sons, Ltd. 2002; ISBN: 9780471489979.**

Web pages:

- Renewable energies companies association (In Spanish). http://www.appa.es/

- Ministry of Industry, Energy and Tourism. Govern of Spain (In Spanish). http://www.minetur.gob.es/
- Wind Energy Association. http://www.aeeolica.org/
- Eurelectric. Electricity for Europe. http://www.eurelectric.org/
- Renewal Energy National Center (In Spanish). http://www.cener.com/

Part II
Infrastructure Design for Various Energy Sources

Development Planning of Offshore Oilfield Infrastructure

Vijay Gupta and Ignacio E. Grossmann

Abstract The development planning of an offshore oilfield infrastructure involves critical investment and operating decisions at an early stage of the project that impact the overall project profitability. These strategic/tactical decisions need to be made considering sufficient reservoir details, fiscal contracts with the government and uncertainty in the field parameters to be useful in practice. However, it makes the optimization problem difficult to model and solve. With this motivation, the objective of this chapter is to present a comprehensive review of a unified modeling framework and solution strategies to address the issues of complex fiscal rules and endogenous uncertainties in the development planning of offshore oil and gas field infrastructure. In particular, the chapter emphasizes the need to have as a basis an efficient deterministic model that can account for various alternatives in the decision-making process for a multi-field site incorporating sufficient level of details, while being computationally tractable for large instances. Consequently, such a model can effectively be extended to include other complexities, for instance a production sharing agreement and endogenous uncertainties. Computational results on the deterministic as well as multistage stochastic instances of the problem are discussed, with extension to incorporate Lagrangian decomposition solution algorithms.

Keywords Oil and gas exploration · Production sharing agreements · Ringfencing · Multistage stochastic programming · Endogenous uncertainties · Lagrangian decomposition

V. Gupta · I.E. Grossmann (✉)
Department of Chemical Engineering, Carnegie Mellon University,
Pittsburgh, PA 15213, USA
e-mail: grossmann@cmu.edu

V. Gupta
e-mail: v.gupta.cmu@gmail.com

M. Martín (ed.), *Alternative Energy Sources and Technologies*,
DOI 10.1007/978-3-319-28752-2_3

1 Introduction

The development planning of offshore oil and gas fields has received significant attention in recent years given the new discoveries of large oil and gas reserves in the last decade around the world. These have been facilitated by the new technologies available for exploration and production of oilfields in remote locations that are often hundreds of miles offshore. Surprisingly, there has been a net increase in the total oil reserves in the last decade because of these discoveries despite increase in the total demand (BP, Statistical review Report 2011). Therefore, there is currently a strong focus on exploration and development activities for new oil fields all around the world, specifically at offshore locations. However, installation and operating decisions in these projects involve very large investments that potentially can lead to large profits, but also to losses if these decisions are not made carefully.

In particular, the planning of offshore oil and gas field development represents a very complex problem and involves multi-billion dollar investments (Babusiaux et al. 2007). The major decisions involved in the oilfield development planning phase are the following:

(a) Selecting platforms to install and their sizes
(b) Deciding which fields to develop and what should be the order to develop them
(c) Deciding which wells and how many are to be drilled in the fields and in what sequence
(d) Deciding which fields are to be connected to which facility
(e) Determining how much oil and gas to produce from each field

Therefore, there are a very large number of alternatives that are available to develop a particular field or group of fields. However, these decisions should account for the physical and practical considerations, such as the following: a field can only be developed if a corresponding facility is present; nonlinear profiles of the reservoir to predict the actual flowrates of oil, water, and gas from each field; limitation on the number of wells that can be drilled each year due to availability of the drilling rigs; and long-term planning horizon that is the characteristics of the these projects. Therefore, optimal investment and operating decisions are essential for this problem to ensure the highest return on the investments over the time horizon considered. By including all the considerations described here in an optimization model, this leads to a large-scale multiperiod MINLP problem that is difficult to solve to global optimality. The extension of this model to the cases where we explicitly consider the fiscal rules and the uncertainties can further lead to a very complex problem to model and solve.

With the motivation described above, the chapter presents a comprehensive review of the optimal development planning of offshore oil and gas fields based on our recent work in this area, highlighting the models and solution approaches that we proposed, and the key issues that may be still be present. In particular, a unified

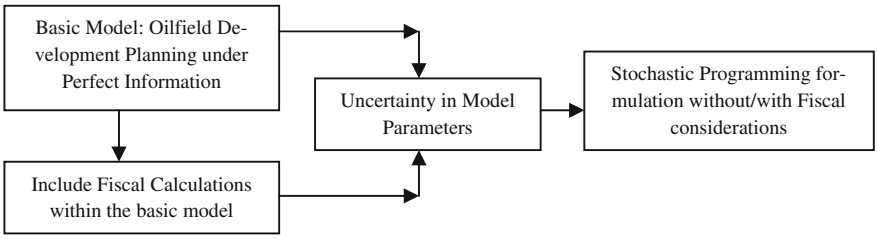


Fig. 1 A unified framework for oilfield development planning under complex fiscal rules and uncertainties

modeling framework (Fig. 1) is presented for this problem starting with a basic deterministic model that includes sufficient level of detail to be realistic as well as computationally efficient as proposed in Gupta and Grossmann (2012a). Moreover, we discuss the extension of the model for incorporating fiscal rules defined by the terms of the contract between oil companies and governments (Gupta and Grossmann 2012b), and uncertainty based on a multistage stochastic programming framework (Gupta and Grossmann 2014a).

The chapter starts with a detailed literature review of the various models and solution approaches that are proposed for oil/gas field development and operations emphasizing the advantage and the corresponding limitations. Then, a brief background on the oilfield development projects, the basic structure of an offshore oilfield site, fiscal contracts that are used in the industry and uncertain parameters are described. The specific problem under consideration involving nonlinear reservoir profiles, production sharing agreements and endogenous uncertainties, and the assumptions involved are presented. A basic MINLP model from Gupta and Grossman (2012a) relying on this description is reviewed and its extension to include fiscal rules (Gupta and Grossmann 2012b) and/or uncertain field parameters using multistage stochastic programming approach (Gupta and Grossmann 2014a), and the corresponding solution strategies are highlighted next. Numerical results on several examples ranging from deterministic to stochastic cases for oilfield planning problem are considered to emphasize the efficiency of the models and the solution methods.

2 Literature Review

2.1 Deterministic Approaches for Oil/Gas Field Development Planning

The oilfield development planning has traditionally been modeled as LP (Lee and Aranofsky 1958; Aronofsky and Williams 1962) or MILP (Frair 1973) models under certain assumptions to make them computationally tractable. Simultaneous

optimization of the investment and operating decisions was addressed in Bohannon (1970), Sullivan (1982) and Haugland et al. (1988) using MILP formulations with different levels of details in these models. Behrenbruch (1993) emphasized the need to consider a correct geological model and to incorporate flexibility into the decision process for an oilfield development project.

Iyer et al. (1998) proposed a multiperiod MILP model for optimal planning and scheduling of offshore oilfield infrastructure investment and operations. The model considers the facility allocation, production planning, and scheduling within a single model and incorporates the reservoir performance, surface pressure constraints, and oil rig resource constraints. To solve the resulting large-scale problem, the nonlinear reservoir performance equations are approximated through piecewise linear approximations. As the model considers the performance of each individual well, it becomes expensive to solve for realistic multi-field sites. Moreover, the flow rate of water was not considered explicitly for facility capacity calculations.

Van den Heever and Grossmann (2000) extended the work of Iyer et al. (1998) and proposed a multiperiod generalized disjunctive programming model for oil field infrastructure planning for which they developed a bilevel decomposition method. As opposed to Iyer and Grossmann (1998), they explicitly incorporated a nonlinear reservoir model into the formulation but did not consider the drill-rig limitations.

Grothey and McKinnon (2000) addressed an operational planning problem using an MINLP formulation where gas has to be injected into a network of low pressure oil wells to induce flow from these wells. Lagrangian decomposition and Benders decomposition algorithms were proposed for the efficient solution of the model. Kosmidis et al. (2002) considered a production system for oil and gas consisting of a reservoir with several wells, headers, and separators. The authors presented a mixed-integer dynamic optimization model and an efficient approximation solution strategy for this system.

Barnes et al. (2002) optimized the production capacity of a platform and the drilling decisions for wells associated with this platform. The authors addressed this problem by solving a sequence of MILPs. Ortiz-Gomez et al. (2002) presented three mixed-integer multiperiod optimization models of varying complexity for the oil production planning. The problem considers fixed topology and is concerned with the decisions involving the oil production profiles and operation/shut-in times of the wells in each time period assuming nonlinear reservoir behavior.

Lin and Floudas (2003) considered the long-term investment and operations planning of the integrated gas field site. A continuous-time modeling and optimization approach was proposed introducing the concept of event points and allowing the well platforms to come online at potentially any time within planning horizon. Two-level solution framework was developed to solve the resulting MINLP problems which showed that the continuous-time approach can reduce the computational efforts substantially and solve problems that were intractable for the discrete-time model.

Kosmidis et al. (2005) presented a mixed-integer nonlinear (MINLP) model for the daily well scheduling in petroleum fields, where the nonlinear reservoir behavior, the multiphase flow in wells, and constraints from the surface facilities

were simultaneously considered. The authors also proposed a solution strategy involving logic constraints, piecewise linear approximations of each well model, and an outer approximation based algorithm. Results showed an increase in oil production up to 10 % compared to a typical heuristic rules widely applied in practice.

Carvalho and Pinto (2006a) considered an MILP formulation for oilfield planning based on the model developed by Tsarbopoulou (2000), and proposed a bilevel decomposition algorithm for solving large-scale problems where the master problem determines the assignment of platforms to wells and a planning subproblem calculates the timing for the fixed assignments. The work was further extended by Carvalho and Pinto (2006b) to consider multiple reservoirs within the model.

Barnes et al. (2007) addressed the optimal design and operational management of offshore oil fields where at the design stage the optimal production capacity of a main field was determined with an adjacent satellite field and a well drilling schedule. The problem was formulated as an MILP model. Continuous variables involved individual well, jacket, and topsides costs, whereas binary variables were used to select individual wells within a defined field grid. An MINLP model was proposed for the operational management to model the pressure drops in pipes and wells for multiphase flow. Nonlinear cost equations were derived for the production costs of each well accounting for the length, the production rate and their maintenance. Operational decisions included the oil flow rates, the operation/shut-in for each well and the pressures for each point in the piping network.

Gunnerud and Foss (2010) considered the real-time optimization of oil production systems with a decentralized structure and modeled nonlinearities by piecewise linear approximations, resulting in a MILP model. The Lagrange relaxation and Dantzig–Wolfe decomposition methods were studied on a semi-realistic model of the Troll west oil rim which showed that both the approaches offers an interesting option to solve the complex oil production systems as compared to the fullspace method.

2.2 *Incorporating Complex Fiscal Rules*

The major limitation with the above approaches is that they do not consider the fiscal rules explicitly in the optimization model that are associated to these fields, and mostly rely on the simple net present value (NPV) as an objective function. Therefore, the models with these objectives may yield the solutions that are very optimistic, which can in fact be suboptimal after considering the impact of fiscal terms. Bagajewicz (2008) discussed the merits and limitations of using NPV in the investment planning problems and pointed out that additional consideration and procedures are needed for these problems, e.g., return on investments, to make the better decisions. Laínez et al. (2009) emphasizes that enterprise-wide decision problems must be formulated with realistic detail, not just in the technical aspects, but also in the financial components in order to generate solutions that are of value

to an enterprise. This requires systematically incorporating supplier/buyer options contracts within the framework of supply-chain problems.

In the context of oilfield planning, fiscal rules of the agreements between the oil company (contractor) and the host government, e.g., production sharing contracts, usually determine the share of each of these entities in the total oil production or gross revenues and the timing of these payments. Hence, including fiscal considerations as part of the oilfield development problem can significantly impact the optimal decisions and revenue flows over the planning horizon, as a large fraction of the total oil produced is paid as royalties, profit share, etc. The models and solutions approaches in the literature that consider the fiscal rules within oilfield infrastructure planning are either very specific or simplified. Van den Heever et al. (2000) and Van den Heever and Grossmann (2001) considered optimizing the complex economic objectives including royalties, tariffs, and taxes for the multiple gas field site where the schedule for the drilling of wells was predetermined as a function of the timing of the installation of the well platform. Moreover, the fiscal rules presented were specific to the gas field site considered, but not in generic form. Based on a continuous-time formulation for gas field development with complex economics of similar nature as Van den Heever and Grossmann (2001), Lin and Floudas (2003) proposed an MINLP model and solved it with a two-stage algorithm. Approaches based on simulation (Blake and Roberts 2006) and meta-modeling (Kaiser and Pulsipher 2004) have also been considered for the analysis of the different fiscal terms. However, the papers that address the mathematical programming models and solution approaches for the oilfield investments and operations with fiscal considerations are still very limited.

2.3 *Incorporating Uncertainties in the Development Planning*

In the literature work described above, one of the major assumptions is that there is no uncertainty in the model parameters, which in practice is generally not true. Although limited, there has been some work that accounts for uncertainty in the problem of optimal development of oil and/or gas fields. Haugen (1996) proposed a single parameter representation for uncertainty in the size of reserves and incorporates it into a stochastic dynamic programming model for scheduling of oil fields. However, only decisions related to the scheduling of fields were considered. Meister et al. (1996) presented a model to derive exploration and production strategies for one field under uncertainty in reserves and future oil prices. The model was analyzed using stochastic control techniques.

Jonsbraten (1998a) addressed the oilfield development planning problem under oil price uncertainty using an MILP formulation that was solved with a progressive hedging algorithm. Aseeri et al. (2004) introduced uncertainty in the oil prices and well productivity indexes, financial risk management, and budgeting constraints

into the model proposed by Iyer and Grossmann (1998), and solved the resulting stochastic model using a sampling average approximation algorithm.

Jonsbraten (1998b) presented an implicit enumeration algorithm for the sequencing of oil wells under uncertainty in the size and quality of oil reserves. The author uses a Bayesian approach to represent the resolution of uncertainty with investments. Both these papers consider investment and operation decisions for one field only. Lund (2000) addressed a stochastic dynamic programming model for evaluating the value of flexibility in offshore development projects under uncertainty in future oil prices and in the reserves of one field using simplified descriptions of the main variables.

Cullick et al. (2003) proposed a model based on the integration of a global optimization search algorithm, a finite-difference reservoir simulation, and economics. In the solution algorithm, new decision variables were generated using meta-heuristics, and uncertainties were handled through simulations for fixed design variables. They presented examples having multiple oil fields with uncertainties in the reservoir volume, fluid quality, deliverability, and costs. Few other papers (Begg et al. 2001; Zabalza-Mezghani et al. 2004; Bailey et al. 2005; Cullick et al. 2007), have also used a combination of reservoir modeling, economics, and decision-making under uncertainty through simulation-optimization frameworks.

Ulstein et al. (2007) addressed the tactical planning of petroleum production that involves regulation of production levels from wells, splitting of production flows into oil and gas products, further processing of gas and transportation in a pipeline network. The model was solved for different cases with demand variations, quality constraints, and system breakdowns.

Elgsæter et al. (2010) proposed a structured approach to optimize offshore oil and gas production with uncertain models that iteratively updates setpoints, while documenting the benefits of each proposed setpoint change through excitation planning and result analysis. The approach is able to realize a significant portion of the available profit potential, while ensuring feasibility despite large initial model uncertainty.

However, most of these works either consider the very limited flexibility in the investment and operating decisions or handle the uncertainty in an ad hoc manner. Stochastic programming provides a systematic framework to model problems that require decision-making in the presence of uncertainty by taking uncertainty into account of one or more parameters in terms of probability distribution functions, (Birge and Louveaux 1997). The concept of recourse action in the future, and availability of probability distribution in the context of oilfield development planning problems, makes it one of the most suitable candidates to address uncertainty. Moreover, extremely conservative decisions are usually ignored in the solution utilizing the probability information given the potential of high expected profits in the case of favorable outcomes.

In the context of stochastic programming, Goel and Grossmann (2004) considered a gas field development problem under uncertainty in the size and quality of reserves where decisions on the timing of field drilling were assumed to yield an immediate resolution of the uncertainty, i.e., the problem involves decision-dependent

uncertainty as discussed in Jonsbraten et al. (1998), Goel and Grossmann (2006), and Gupta and Grossmann (2011). Linear reservoir models, which can provide a reasonable approximation for gas fields, were used. In their solution strategy, the authors used a relaxation problem to predict upper bounds, and solved multistage stochastic programs for a fixed scenario tree for finding lower bounds. Goel et al. (2006) later proposed the theoretical conditions to reduce the number of non-anticipativity constraints in the model. The authors also developed a branch and bound algorithm for solving the corresponding disjunctive/mixed-integer programming model where lower bounds are generated by Lagrangian duality. The proposed decomposition strategy relies on relaxing the disjunctions and logic constraints for the non-anticipativity constraints. Ettehad et al. (2011) presented a case study for the development planning of an offshore gas field under uncertainty optimizing facility size, well counts, compression power, and production policy. A two-stage stochastic programming model was developed to investigate the impact of uncertainties in original gas in place and inter-compartment transmissibility. Results of two solution methods, optimization with Monte Carlo sampling and stochastic programming, were compared which showed that the stochastic programming approach is more efficient. The models were also used in a value of information (VOI) analysis.

Moreover, the gradual uncertainty reduction has also been addressed for problems in this class. Stensland and Tjøstheim (1991) have worked on a discrete-time problem for finding optimal decisions with uncertainty reduction over time and applied their approach to oil production. These authors expressed the uncertainty in terms of a number of production scenarios. Their main contribution was combining production scenarios and uncertainty reduction effectively for making optimal decisions. Dias (2002) presented four propositions to characterize technical uncertainty and the concept of revelation towards the true value of the variable. These four propositions, based on the theory of conditional expectations, are employed to model technical uncertainty.

Tarhan et al. (2009) addressed the planning of offshore oil field infrastructure involving endogenous uncertainty in the initial maximum oil flowrate, recoverable oil volume, and water breakthrough time of the reservoir, where decisions affect the resolution of these uncertainties. The authors extend the work of Goel and Grossmann (2004) and Goel et al. (2006) but with three major differences: (a) The model focuses on a single field consisting of several reservoirs rather than multiple fields but more detailed decisions are considered. (b) Nonlinear, rather than linear, reservoir models are considered. (c) The resolution of uncertainty is gradual over time instead of being resolved immediately. The authors also developed a multistage stochastic programming framework that was modeled as a disjunctive/mixed-integer nonlinear programming model consisting of individual non-convex MINLP subproblems connected to each other through initial and conditional non-anticipativity constraints. A duality-based branch and bound algorithm was proposed taking advantage of the problem structure and globally optimizing each scenario problem independently. An improved solution approach was also proposed that combines global optimization and outer approximation to optimize the investment and operations decisions (Tarhan et al. 2011). However, it considers

either gas/water or oil/water components for single field and single reservoir at a detailed level. Hence, realistic multi-field site instances can be expensive to solve with this model.

Most of the works described here either consider simple reservoir profiles, only fiscal rules or uncertainty. However, we are not aware of any paper in the literature that addresses nonlinear reservoir profiles, fiscal rules, and endogenous uncertainties for the multiple oilfields planning in a simultaneous manner. To overcome this limitation, we first consider a deterministic model and ways to improve its computational efficiency as described in Gupta and Grossmann (2012a). The model is then extended to incorporate complex fiscal rules (Gupta and Grossmann 2012b) and uncertainties using multistage stochastic programming (Gupta and Grossmann 2014a). Notice that the motivation for this chapter is not to claim a new contribution; rather we intend to present a comprehensive review of our recent work to address the above-mentioned complexities for the oilfield development problem.

3 Background

In this section, we present a general background on the offshore oil/gas field infrastructure, fiscal rules that are usually associated, and uncertainties in the various parameters. These elements will be used as a basis to describe the problem under considerations in the next section.

The life cycle of a typical offshore oilfield project consists of following five steps:

(1) *Exploration* This activity involves geological and seismic surveys followed by exploration wells to determine the presence of oil or gas.
(2) *Appraisal* It involves drilling of delineation wells to establish the size and quality of the potential field. Preliminary development planning and feasibility studies are also performed.
(3) *Development* Following a positive appraisal phase, this phase aims at selecting the most appropriate development plan among many alternatives. This step involves capital-intensive investment and operating decisions that include facility installations, drilling, subsea structures, etc.
(4) *Production* After facilities are built and wells are drilled, production starts where gas or water is usually injected in the field at a later time to enhance productivity.
(5) *Abandonment* This is the last phase of an oilfield development project and involves the decommissioning of facility installations and subsea structures associated with the field.

Given that most of the critical investments are usually associated with the development planning phase of the project, this chapter focuses on the key strategic/tactical decisions during this phase of the project.

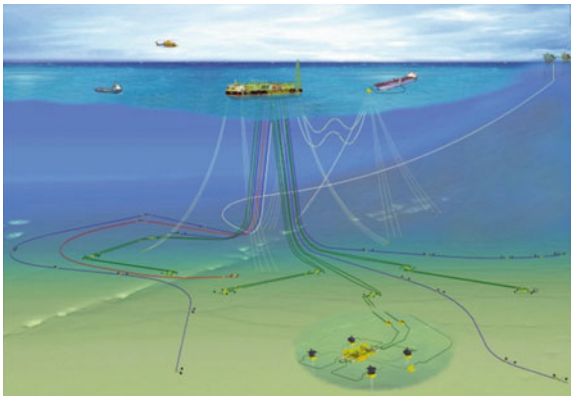


Fig. 2 A complex offshore oilfield infrastructure

An offshore oilfield infrastructure (Fig. 2) is usually very complex and involves various production facilities such as floating production, storage and offloading (FPSO), tension-leg platform (TLP), fields, wells, and connecting pipelines to produce oil and gas from the reserves. Each oilfield consists of a number of potential wells to be drilled using drilling rigs, which are then connected to the facilities through pipelines to produce oil. The field to facility connection involves trade-offs associated to the production rate of oil and gas, piping costs, and possibility of other fields to connect to that same facility. The number of wells that can be drilled in a field depends on the availability of the drilling rig that can drill a certain number of wells each year.

There is a multiphase flow in the connecting pipelines due to the presence of oil, water, and gas. Therefore, there may be several components present in the produced oil, and their relative amounts depend on certain parameters like cumulative oil produced. The facilities and piping connections in the offshore infrastructure are often in operation over many years. It is therefore important to anticipate future conditions when designing an initial infrastructure or any expansions. This can be accomplished by dividing the planning horizon, for example, 20 years, into a number of time periods with a length of 1 year, and allowing investment and operating decisions in each period, which leads to a multiperiod planning problem.

There are a variety of contracts that are used in the offshore oil and gas industry (Babusiaux et al. 2007; Johnston 1994; Sunley et al. 2002 and Tordo 2007). Although the terms of a particular agreement are usually negotiated between both the entities in practice, these contracts can broadly be classified into two main categories:

(1) *Concessionary System* A concessionary (or tax and royalty) system usually involves royalty, cost deduction, and tax. Royalty is paid to the government at a certain percentage of the gross revenues. The net revenue after deducting costs becomes taxable income on which a predefined percentage is paid as tax which may include both corporate income tax and a specific profit tax. The total contractor's share involves gross revenues minus royalty and taxes in

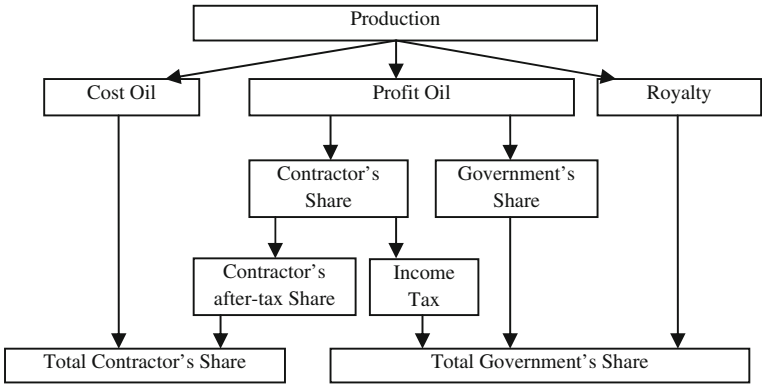


Fig. 3 Revenue flow for a typical production sharing agreement

each year. The basic difference as compared to the production sharing agreement is that the oil company keeps the right to all of the oil and gas produced at the wellhead and pays royalties, bonuses, and other taxes to the government. These contracts are used in countries such as Canada, USA, and the UK.

(2) *Production Sharing Agreements* (*PSAs*) The revenue flow in a typical Production Sharing Agreement can be seen as in Fig. 3 (World Bank 2007). First, in most cases, the company pays royalty to the government at a certain percentage of the total oil produced. After paying the royalties, some portion of the remaining oil is treated as cost oil by the oil company to recover its costs. There is a ceiling on the cost oil recovery to ensure revenues to the government as soon as production starts. The remaining part of the oil, called profit oil, is divided between the oil company and the host government at a certain percentage. The oil company needs to further pay income tax on its share of profit oil. Hence, the total contractor's (oil company) share in the gross revenue comprises of cost oil and contractor's profit oil share after tax. The other important feature of a PSA is that the government keeps rights to the oil produced at wellhead, and transfers title to a portion of the extracted oil and gas to oil company that works as a contractor at an agreed delivery point. Notice that the cost oil limit is one of the key differences with a concessionary system (World Bank 2007). These contracts are used in countries such as Cambodia, China, Egypt, India, Angola, and Nigeria.

The specific rules defined in such a contract (either concessionary or PSA, hybrid) between oil company and host government determine the profit that the oil company can keep, as well as the royalties and profit oil share that are paid to the government. These profit oil fractions, royalty rates define the fiscal terms of a particular contract and can be either "regressive or progressive". Regressive fiscal terms are not directly linked to the profitability of the project, e.g., fixed percentage of royalty or profit oil share for the entire planning horizon. However, the

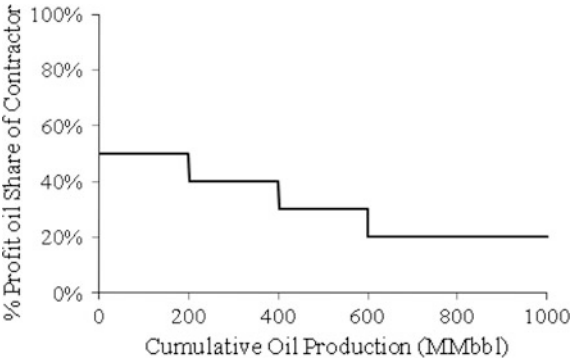


Fig. 4 Progressive profit oil share of the contractor

progressive fiscal terms (e.g., profit oil shares, royalty rates) are based on the profitability of the project, i.e., these terms penalize higher production rates, where cumulative oil produced, daily production, rate of return, R-factor, are the typical profitability measures that determine the tier structure (levels) for these contract terms. For instance, if the cumulative production is in the range of first tier, $0 \leq xc_t \leq 200$, the contractor receives 50 % of the profit oil, while if the cumulative production reaches in tier $2{,}200 \leq xc_t \leq 400$, the contractor receives 40 % of the profit oil, and so on (see Fig. 4). In practice, as we move to the higher tier, the percentage share of contractor in the total production decreases.

Ringfencing is an important concept that is usually part of the fiscal contracts and imposed by the government, which affects the cash flows over the planning horizon. In a typical ringfencing provision, investment and operational costs for a specified group of fields or block can only be recovered from the revenue generated from those fields or block (see Fig. 5). It means that the set of particular fields are "ring-fenced". Therefore, income derived from one contract area or project cannot be offset against losses from another contract area or project. In financial terms, a ringfencing provision basically defines the level at which all fiscal calculations need to be done, and restricts the oil companies to balance the costs and revenues across

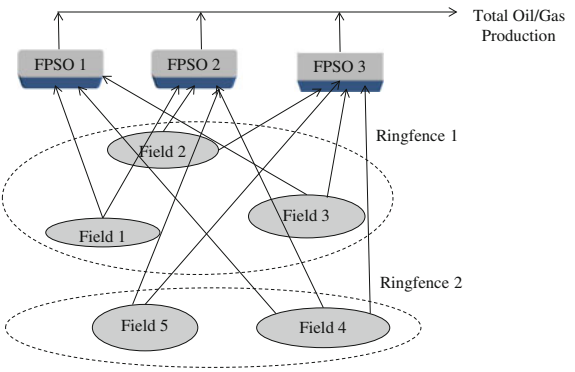


Fig. 5 A typical offshore oilfield infrastructure representation

various projects/blocks for minimizing the tax burden. For example, fiscal calculations for Fields 1–3 (Ringfence 1) and Field 4–5 (Ringfence 2) in Fig. 5 cannot be consolidated at one place. Ringfencing provisions are more popular in production sharing contracts.

These fiscal contracts, terms and ringfencing provisions are the backbone of most of the contracts that are currently used, and can have significant impact on the revenues. In addition, there can be some other fiscal considerations for a particular contract of interest, but for simplicity we only consider the important financial elements as described above. Notice that the royalties and/or government profit oil share that result from a particular contract can represent a significant amount of the gross revenues. Therefore, it is critical to consider these contract terms explicitly during the oilfield planning phase to assess the actual economic potential of such a project. For instance, a very promising oilfield or block can turn out to be a big loss or less profitable than projected in the long-term if significant royalties are attached to that field, which was not considered during the development planning phase involving large investments. On the contrary, there could be the possibility of missing an opportunity to invest in a field that has very difficult conditions for production and looks unattractive, but can have favorable fiscal terms resulting in large profits in the long term.

In addition, there are multiple sources of uncertainty in these oil/gas field development projects. The market price of oil/gas, quantity and quality of reserves at a field are the most important sources of the uncertainty in this context. The uncertainty in oil prices is influenced by the political, economic, or other market factors, and it belongs to the exogenous uncertainty problems. The uncertainty in the reserves on the other hand, is linked to the accuracy of the reservoir data (technical uncertainty). While the existence of oil and gas at a field is indicated by seismic surveys and preliminary exploratory tests, the actual amount of oil in a field, and the efficacy of extracting the oil will only be known after capital investment have been made at the field (Goel and Grossmann 2004), i.e., endogenous uncertainty. Both, the price of oil and the quality of reserves directly affect the overall profitability of a project, and hence it is important to consider the impact of these uncertainties when formulating the decision policy. However, in this chapter we only address the uncertainty in the field parameters that is already a very challenging problem to model and solve. In the next section, we outline the specific problem under consideration that incorporates nonlinear reservoir behavior, fiscal rules, and endogenous uncertainties.

4 Problem Description

In this chapter, we summarize our recent work on the development planning of an offshore oil and gas field infrastructure under complex fiscal rules and endogenous uncertainties. In particular, a multi-field site, $F = \{1, 2, \ldots f\}$, with potential investments in the floating production storage and offloading (FPSO) facilities,

Fig. 6 FPSO (Floating Production Storage and Offloading) facility

FPSO = {1, 2, ... fpso} with continuous capacities and ability to expand them in the future is considered (Fig. 5). The extension for including tension-leg platform (TLP) is straightforward but for simplicity we only consider FPSOs (Fig. 6). These FPSO facilities cost multi-billion dollars depending on their sizes, and have the capability of operating in remote locations for very deep offshore oilfields (200–2000 m) where seabed pipelines are not cost effective. FPSOs are large ships that can process the produced oil and store it until it is shipped to the onshore site or sales terminal. Processing includes the separation of oil, water, and gas into individual streams using separators located at these facilities. Each FPSO facility has a lead time between the construction or expansion decision, and its actual availability. The wells are subsea wells in each field that are drilled using drilling ships. Therefore, there is no need to have a facility present to drill a subsea well. The only requirement to recover oil from it is that the well must be connected to a FPSO facility.

The connection of a field to an installed FPSO facility and a number of wells need to be drilled to produce oil from these fields for the given planning horizon. The planning horizon is discretized into a number of time periods t, each with 1 year of duration. The location of each FPSO facility and its possible connections to the given fields are assumed to be known. Notice that each FPSO facility can be connected to more than one field to produce oil, while a field can only be connected to a single FPSO facility due to engineering requirements and economic viability of the project. The water produced with the oil is usually re-injected after separation, while the gas can be sold in the market. In this case, we consider natural depletion of the reserves i.e., no water or gas re-injection. The location of production facilities and possible field and facility allocation itself is a very complex problem. In this work, we assume that the potential location of facilities and field–facility connections are given. In addition, the potential number of wells in each field is also given. There are following three main complexities in the problem considered here:

(i) **Nonlinear Reservoir Profiles** We consider three components (oil, water, and gas) explicitly during production from a field. Field deliverability, i.e., maximum oil flow rate from a field, water-oil-ratio (WOR) and gas-oil-ratio

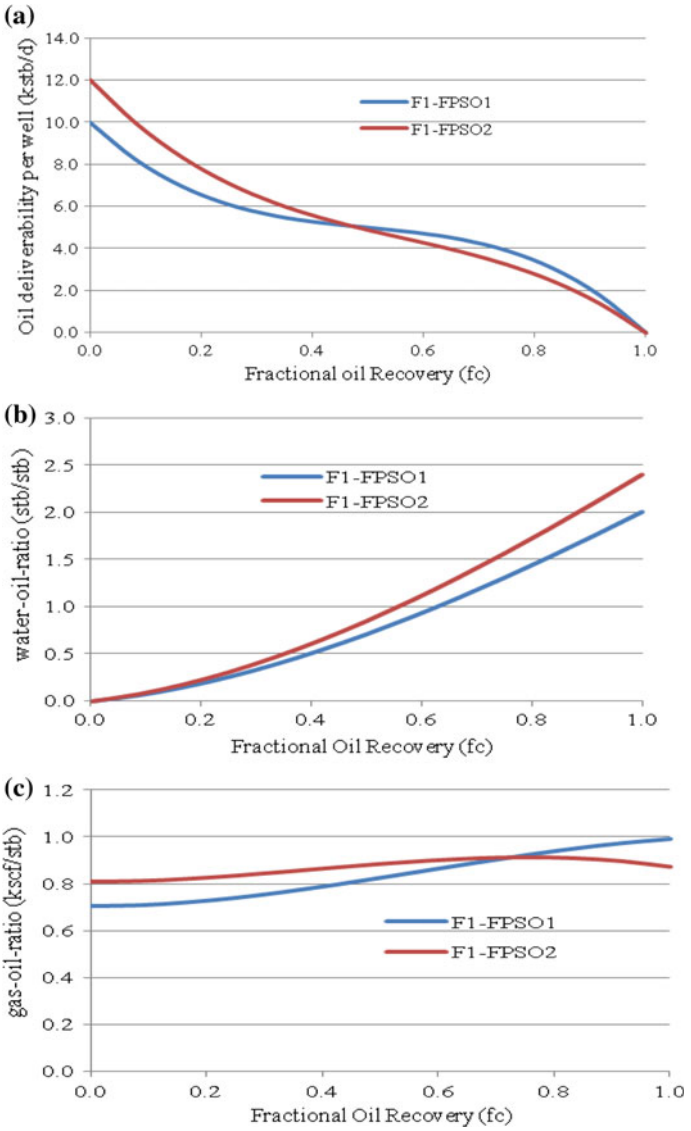


Fig. 7 Nonlinear reservoir profiles for field (F1) and 2 FPSOs (FPSO1–2). **a** Oil deliverability per well for field (F1). **b** Water to oil ratio for field (F1). **c** Gas-to-oil ratio for field (F1)

(GOR) are approximated by a cubic equations (a)–(c) (see Fig. 7), while cumulative water produced and cumulative gas produced from a field are represented by fourth order separable polynomials, Eqs. (d)–(e), that are derived in Gupta and Grossmann (2012a). The motivation for using polynomials for cumulative water produced and cumulative gas produced, Eqs. (d)–(e),

as compared to WOR and GOR, Eqs. (b)–(c), is to avoid bilinear terms, Eqs. (f)–(g), in the formulation and allow converting the resulting model into an MILP formulation. All the wells in a particular field f are assumed to be identical for the sake of simplicity leading to the same reservoir profiles, Eqs. (a)–(g), for each of these wells.

$$\hat{Q}_f^d = a_{1,f}(fc_{f,t})^3 + b_{1,f}(fc_f)^2 + c_{1,f}fc_f + d_1 \quad \forall f \tag{a}$$

$$\hat{\text{wor}}_f = a_{2,f}(fc_f)^3 + b_{2,f}(fc_f)^2 + c_{2,f}fc_f + d_{2,f} \quad \forall f \tag{b}$$

$$\hat{\text{gor}}_f = a_{3,f}(fc_f)^3 + b_{3,f}(fc_f)^2 + c_{3,f}fc_f + d_{3,f} \quad \forall f \tag{c}$$

$$\hat{\text{wc}}_f = a_{4,f}(fc_f)^4 + b_{4,f}(fc_f)^3 + c_{4,f}fc_f^2 + d_{4,f}fc_f \quad \forall f \tag{d}$$

$$\hat{\text{gc}}_f = a_{5,f}(fc_f)^4 + b_{5,f}(fc_f)^3 + c_{5,f}fc_f^2 + d_{5,f}fc_f \quad \forall f \tag{e}$$

$$w_f = \hat{\text{wor}}_f \cdot x_f \quad \forall f \tag{f}$$

$$g_f = \hat{\text{gor}}_f \cdot x_f \quad \forall f \tag{g}$$

(ii) **Production Sharing Agreements** We consider progressive (sliding scale) production sharing agreements with ringfencing provisions which are widely used in several countries (Fig. 3). In particular, we assume a sliding scale profit oil share of the contractor linked to the cumulative oil produced (see Fig. 4). Notice that this tier structure is a step function, which requires additional binary variables to model and makes the problem harder to solve. Moreover, the cost recovery ceiling is considered to be a fraction of the gross revenues in each time period t. For simplicity, the cost recovery ceiling fraction and income tax rates are assumed to be a fixed percentages (no sliding scale), and there are no explicit royalty provisions which is a straightforward extension. A set of ringfences RF = {1, 2, …} among the given fields is specified (see Fig. 5) to ensure that fiscal calculations are to be done for each ringfence separately (see Gupta and Grossmann 2012b for details). Qualitatively, a typical ringfencing provision states that the investment and operational costs for a specified group of fields or block can only be recovered from the revenue generated from those fields or block. Notice that in general a field is associated to a single ringfence, while a ringfence can include more than one field. In contrast, a facility can be connected to multiple fields from different ringfences for producing oil and gas. These ringfences may or may not have the same fiscal rules.

(iii) **Endogenous Uncertainties** We consider here the uncertainty in the field parameters, i.e., field size, oil deliverability per well, water-oil ratio, and gas-oil ratio. These are the endogenous uncertain parameters since our

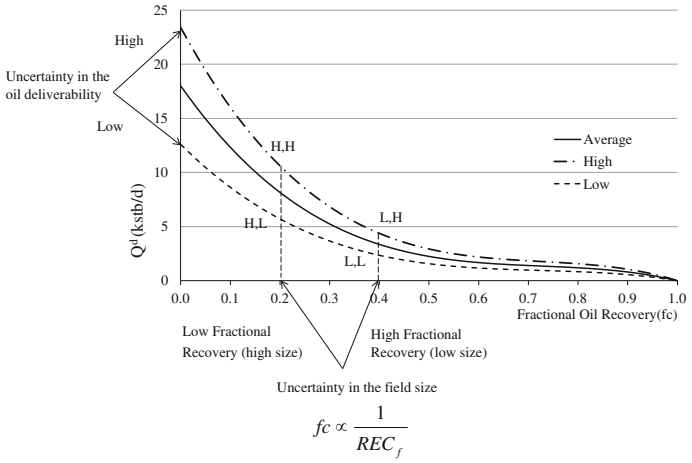


Fig. 8 Oil deliverability per well for a field under uncertainty

investment and operating decisions affect the stochastic process (Jonsbraten et al. 1998; Goel and Grossmann 2006; Tarhan et al. 2008; and Gupta and Grossmann 2011). In particular, the uncertainty in the field parameters can only resolved when an investment is made in that field and we start producing from it. Therefore, our decisions determine the timing of uncertainty realization i.e., decision-dependent uncertainty.

The average profile in Fig. 8 represents the oil deliverability per well for a field as a nonlinear polynomial in terms of the fractional oil recovery (Eq. a) under the perfect information. However, due to the uncertainty in the oil deliverability, the actual profile is assumed to be either lower or upper side of the average profile with a given probability. In particular, Eq. (h) represents the oil deliverability per well for a field under uncertainty where parameter α_{oil} is used to characterize this uncertainty. For instance, if $\alpha_{\text{oil}} > 1$, then we have a higher oil deliverability than expected $(\alpha_{\text{oil}} = 1)$, whereas for $\alpha_{\text{oil}} < 1$ a lower than expected oil deliverability is observed. Since, the uncertain field size is an inverse function of the fraction oil recovery, a higher field size will correspond to the low fractional oil recovery whereas a small field size will correspond to the higher factional oil recovery for a given amount of the cumulative oil production.

$$Q_f^d = \alpha_{\text{oil}} \cdot \hat{Q}_f^d \quad \forall f \tag{h}$$

Similarly, Eqs. (i) and (j) correspond to the uncertain field profiles for water-oil-ratio and gas-oil-ratio that are characterized by the uncertain parameters α_{wor} and α_{gor}, respectively. Notice that since cumulative water produced (Eq. d) and cumulative gas produced (Eq. e) profiles are used in the model, instead of water-oil-ratio (Eq. b) and gas-oil-ratio (Eq. c), the uncertainty in the parameters

α_{wor} and α_{gor} is transformed into the corresponding uncertainty in the parameters α_{wc} and α_{gc} as in Eqs. (k) and (l), respectively (Gupta and Grossmann 2014a).

$$\text{wor}_f = \alpha_{\text{wor}} \cdot \hat{\text{wor}}_f \quad \forall f \tag{i}$$

$$\text{gor}_f = \alpha_{\text{wor}} \cdot \hat{\text{gor}}_f \quad \forall f \tag{j}$$

$$\text{wc}_f = \alpha_{\text{wc}} \cdot \hat{\text{wc}}_f \quad \forall f \tag{k}$$

$$\text{gc}_f = \alpha_{\text{wc}} \cdot \hat{\text{gc}}_f \quad \forall f \tag{l}$$

Moreover, the uncertain parameters for every field, i.e., $\theta_f = \left\{\text{REC}_f, \alpha_{\text{oil}}, \alpha_{\text{wor}}, \alpha_{\text{gor}}\right\}$ are considered to have a number of possible discrete realizations $\tilde{\theta}_f^k$ with a given probability. Therefore, all the possible combinations of these realizations yield a set of scenarios $s \in S^{\text{sup}}$ where each scenario has the corresponding probability p^s. However, the total number of scenarios in set S^{sup} grows exponentially with the number of uncertain parameters and their possible realizations, which makes the problem intractable. To overcome this limitation, we assume that the uncertain parameters for a field $\theta_f = \left\{\text{REC}_f, \alpha_{\text{oil}}, \alpha_{\text{wor}}, \alpha_{\text{gor}}\right\}$ are correlated using practical considerations. Therefore, only a subset of the possible scenarios $S \subset S^{\text{sup}}$ is sufficient to represent the uncertainty (Gupta and Grossmann 2014a). We also assume that the uncertainty in the parameters of a field is resolved if we drill N_1 number of wells in that field and produce from it for a duration of N_2 years. These assumptions on uncertainty resolution rules are flexible and can changed depending on the available field information.

The discrete planning horizon and a discrete set of the selected scenarios for each field with given probabilities can be represented by the scenario trees. However, since the timing of the uncertainty realization for a field (or its corresponding scenarios) depends on the drilling and operating decisions, the resulting scenario tree is decision-dependent. For instance, if we consider a set of two uncertain fields $F = \{1, 2\}$ where each field has four uncertain parameters $\theta_f = \left\{\text{REC}_f, \alpha_{\text{oil}}, \alpha_{\text{wor}}, \alpha_{\text{gor}}\right\}$. The sampled scenario set based on the parameter correlations for each field has 2 elements, i.e., $\{\tilde{\theta}_f^1, \tilde{\theta}_f^2\}$ with equal probability. Therefore, the problem involves only 4 scenarios $S = \{1 : (\tilde{\theta}_1^1, \tilde{\theta}_2^1);\ 2 : (\tilde{\theta}_1^1, \tilde{\theta}_2^2);\ 3 : (\tilde{\theta}_1^2, \tilde{\theta}_2^1);\ 4 : (\tilde{\theta}_1^2, \tilde{\theta}_2^2)\}$ each with a probability of 0.25. Figure 9 represents the scenario tree for this problem, where the uncertainty in the first field is resolved at the end of first year, since we drill N_1 wells in the field at the beginning of year 1 and produce from this field during that year ($N_2 = 1$). The system can be in two states in year 2 depending on the value of the $\tilde{\theta}_1^k$. Similarly, uncertainty in the field 2 is resolved in year 4 under the scenarios 3 and 4 due to drilling and operating decisions, whereas it remains uncertain in the scenarios 1 and 2. Therefore, the resulting scenario tree depends on the optimization decisions, which are not known a priori, requiring to model the superstructure of the all possible

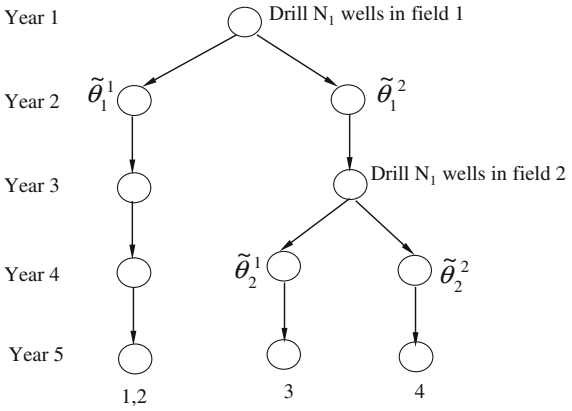


Fig. 9 Decision-dependent scenario tree for two fields

scenario trees that can occur based on our decisions, and therefore, increasing the complexity of the model.

The problem is to determine the optimal investment and operating decisions to maximize the contractor's expected NPV for a given planning horizon considering the above nonlinear profiles, production sharing agreements and endogenous uncertainties. In particular, investment decisions in each time period t and scenario s include FPSO facilities installation or expansion, and their respective installation or expansion capacities for oil, liquid, and gas; fields-FPSO connections; and the number of wells that need to be drilled in each field f given the restrictions on the total number of wells that can be drilled in each time period t over all the given fields. Operating decisions include the oil/gas production rates from each field f in each time period t under every scenario s. It is assumed that the installation and expansion decisions occur at the beginning of each time period t, while operations take place throughout the time period. There is a lead time of l_1 years for each FPSO facility initial installation, and a lead time of l_2 years for the expansion of an earlier installed FPSO facility. Once installed, we assume that the oil, liquid (oil and water), and gas capacities of a FPSO facility can only be expanded once. These assumptions are made for the sake of simplicity, and both the model and the solution approaches are flexible enough to incorporate more complexities. In the next section, we outline the basic oilfield development planning model proposed in Gupta and Grossmann (2012a) that relies on this description, which is further extended to include fiscal rules and/or uncertainties in the subsequent sections.

5 Basic Deterministic Model

We outline in this section the basic MINLP model presented by Gupta and Grossmann (2012a) that takes the infrastructure (Fig. 5) and reservoir characteristics (Eqs. a–f) as reference. Notice that for simplicity we only outline the constraints

that correspond to Model 1 (MINLP) of the paper. With the goal to determine the optimal investment and operating decisions and maximizing the NPV considering a long-term planning horizon, the model involves following set of constraints:

(a) ***Economic Constraints***

The objective function (1) is to maximize the total net present value (NPV) of the project where constraint (2) represents the overall NPV as a function of the difference between total revenue and total cost in each time period t taking the discount factors d_t into account. The total revenues (3) in each time period t are calculated based on the total amount of oil and gas produced in that time period and respective selling prices. The total cost incurred in (4) is the sum of capital and operating expenses in each time period t. The overall capital expenses (5) consist of the fixed installation costs for FPSO facilities, variable installation, and expansion costs corresponding to the FPSOs liquid and gas capacities, connection costs between a field and a FPSO facility and cost of drilling the wells for each field in each time period t. The total operating expenses (6) are the operation cost occurred corresponding to the total amount of liquid and gas produced in each time period t.

$$\text{Max} \quad \text{NPV} \tag{1}$$

$$\text{NPV} = \sum_t d_t(\text{REV}_t - \text{COST}_t) \tag{2}$$

$$\text{REV}_t = \delta_t(\alpha_t x_t^{\text{tot}} + \beta_t g_t^{\text{tot}}) \quad \forall t \tag{3}$$

$$\text{COST}_t = \text{CAP}_t + \text{OPER}_t \quad \forall t \tag{4}$$

$$\text{CAP}_t = \sum_{\text{fpso}} \left[\text{FC}_{\text{fpso},t} b_{\text{fpso},t} + \text{VC}^{\text{liq}}_{\text{fpso},t}(\text{QI}^{\text{liq}}_{\text{fpso},t} + \text{QE}^{\text{liq}}_{\text{fpso},t}) + \text{VC}^{\text{gas}}_{\text{fpso},t}(\text{QI}^{\text{gas}}_{\text{fpso},t} + \text{QE}^{\text{gas}}_{\text{fpso},t})\right] + \sum_f \sum_{\text{fpso}} \text{FC}^{C}_{f,\text{fpso},t} b^{c}_{f,\text{fpso},t} + \sum_f \text{FC}^{\text{well}}_{f,t} I^{\text{well}}_{f,t} \quad \forall t \tag{5}$$

$$\text{OPER}_t = \delta_t\left[\text{OC}^{\text{liq}}_t(x^{\text{tot}}_t + w^{\text{tot}}_t) + \text{OC}^{\text{gas}}_t g^{\text{tot}}_t\right] \quad \forall t \tag{6}$$

(b) ***Reservoir Constraints***

Constraints (7)–(11) predict the reservoir behavior for each field f in each time period t. In particular, constraint (7) restricts the oil flow rate from each well for a particular FPSO-field connection in time period t to be less than the deliverability (maximum oil flow rate) of that field per well. Equations (8)–(9) calculate the field deliverability per well at the beginning of each time period for a particular FPSO-field connection as the cubic equation in terms of the fractional oil recovered. Constraints (10) and (11) represent the value of water-to-oil and gas-to-oil ratios in time period t for a specific field-FPSO connection as cubic equations in terms of the

fractional oil recovery by the end of previous time period, respectively. The predicted WOR and GOR values in Eqs. (10) and (11) are further used in Eqs. (12) and (13) to calculate the respective water and gas flow rates from field to FPSO in time period t by multiplying it with the corresponding oil flow rate. Notice that these equations give rise to the bilinear terms in the model. Equation (14) computes the cumulative amount of oil produced from field f by the end of time period t, while (15) represents the fractional oil recovery by the end of time period t. The cumulative oil produced is also restricted in (16) by the recoverable amount of oil from the field.

$$x^{\text{well}}_{f,\text{fpso},t} \le Q^{d,\text{well}}_{f,\text{fpso},t} \quad \forall f, \text{fpso}, t \tag{7}$$

$$Q^{d,\text{well}}_{f,\text{fpso},1} = d_{1,f,\text{fpso}} \quad \forall f, \text{fpso} \tag{8}$$

$$\begin{aligned} Q^{d,\text{well}}_{f,\text{fpso},t+1} = {} & a_{1,f,\text{fpso}}(fc_{f,t})^3 + b_{1,f,\text{fpso}}(fc_{f,t})^2 \\ & + c_{1,f,\text{fpso}} fc_{f,t} + d_{1,f,\text{fpso}} \quad \forall f, \text{fpso}, t < |T| \end{aligned} \tag{9}$$

$$\text{wor}_{f,\text{fpso},t} = a_{2,f,\text{fpso}}(fc_{f,t-1})^3 + b_{2,f,\text{fpso}}(fc_{f,t-1})^2 + c_{2,f,\text{fpso}} fc_{f,t-1} + d_{2,f,\text{fpso}} \quad \forall f, \text{fpso}, t \tag{10}$$

$$\text{gor}_{f,\text{fpso},t} = a_{3,f,\text{fpso}}(fc_{f,t-1})^3 + b_{3,f,\text{fpso}}(fc_{f,t-1})^2 + c_{3,f,\text{fpso}} fc_{f,t-1} + d_{3,f,\text{fpso}} \quad \forall f, \text{fpso}, t \tag{11}$$

$$w_{f,\text{fpso},t} = \text{wor}_{f,\text{fpso},t} x_{f,\text{fpso},t} \quad \forall f, \text{fpso}, t \tag{12}$$

$$g_{f,\text{fpso},t} = \text{gor}_{f,\text{fpso},t} x_{f,\text{fpso},t} \quad \forall f, \text{fpso}, t \tag{13}$$

$$xc_{f,t} = \sum_{\tau=1}^{t} (x_{f,\tau}\delta_{\tau}) \quad \forall f, t \tag{14}$$

$$fc_{f,t} = \frac{xc_{f,t}}{\text{REC}_f} \quad \forall f, t \tag{15}$$

$$xc_{f,t} \le \text{REC}_f \quad \forall f, t \tag{16}$$

(c) ***Field-FPSO Flow constraints***

The total oil flow rate in (17) from each field f in time period t is the sum of the oil flow rates that are directed to FPSO facilities in that time period t, whereas oil that is directed to a particular FPSO facility from a field f is calculated as the multiplication of the oil flow rate per well and number of wells available for production in that field, Eq. (18). Equations (19)–(21) compute total oil, water, and gas flow rates into each FPSO facility, respectively, in time period t from all the given fields. The total oil, water, and gas flow rates in each time period t are calculated as the sum of the

production rate of these components over all the FPSO facilities in Eqs. (22)–(24), respectively.

$$x_{f,t} = \sum_{\text{fpso}} x_{f,\text{fpso},t} \quad \forall f, t \tag{17}$$

$$x_{f,\text{fpso},t} = N^{\text{well}}_{f,t} \cdot x^{\text{well}}_{f,\text{fpso},t} \quad \forall f, \text{fpso}, t \tag{18}$$

$$x_{\text{fpso},t} = \sum_{f} x_{f,\text{fpso},t} \quad \forall \text{fpso}, t \tag{19}$$

$$w_{\text{fpso},t} = \sum_{f} w_{f,\text{fpso},t} \quad \forall \text{fpso}, t \tag{20}$$

$$g_{\text{fpso},t} = \sum_{f} g_{f,\text{fpso},t} \quad \forall \text{fpso}, t \tag{21}$$

$$x^{\text{tot}}_{t} = \sum_{\text{fpso}} x_{\text{fpso},t} \quad \forall t \tag{22}$$

$$w^{\text{tot}}_{t} = \sum_{\text{fpso}} w_{\text{fpso},t} \quad \forall t \tag{23}$$

$$g^{\text{tot}}_{t} = \sum_{\text{fpso}} g_{\text{fpso},t} \quad \forall t \tag{24}$$

(d) ***FPSO Capacity Constraints***

There are three types of capacities, i.e., for oil, liquid (oil and water), and gas, that are used for modeling the capacity constraints for FPSO facilities. Specifically, Eqs. (25)–(27) restrict the total oil, liquid, and gas flow rates into each FPSO facility to be less than its corresponding capacity in each time period t respectively. These three different kinds of capacities of a FPSO facility in time period t are computed by equalities (28)–(30) as the sum of the corresponding capacity at the end of previous time period $t-1$, installation capacity at the beginning of time period $t-l_1$ and expansion capacity at the beginning of time period $t-l_2$, where l_1 and l_2 are the lead times for FPSO capacity installations and expansions, respectively.

$$x_{\text{fpso},t} \leq Q^{\text{oil}}_{\text{fpso},t} \quad \forall \text{fpso}, t \tag{25}$$

$$x_{\text{fpso},t} + w_{\text{fpso},t} \leq Q^{\text{liq}}_{\text{fpso},t} \quad \forall \text{fpso}, t \tag{26}$$

$$g_{\text{fpso},t} \leq Q^{\text{gas}}_{\text{fpso},t} \quad \forall \text{fpso}, t \tag{27}$$

$$Q^{\text{oil}}_{\text{fpso},t} = Q^{\text{oil}}_{\text{fpso},t-1} + QI^{\text{oil}}_{\text{fpso},t-l_1} + QE^{\text{oil}}_{\text{fpso},t-l_2} \quad \forall \text{fpso}, t \tag{28}$$

$$Q^{\text{liq}}_{\text{fpso},t} = Q^{\text{liq}}_{\text{fpso},t-1} + QI^{\text{liq}}_{\text{fpso},t-l_1} + QE^{\text{liq}}_{\text{fpso},t-l_2} \quad \forall \text{fpso}, t \tag{29}$$

$$Q^{\text{gas}}_{\text{fpso},t} = Q^{\text{gas}}_{\text{fpso},t-1} + QI^{\text{gas}}_{\text{fpso},t-l_1} + QE^{\text{gas}}_{\text{fpso},t-l_2} \quad \forall \text{fpso}, t \tag{30}$$

(e) ***Logic Constraints***

Inequalities (31) and (32) restrict the installation and expansion of a FPSO facility to take place only once, respectively, while inequality (33) states that the connection between a FPSO facility and a field can be installed only once during the whole planning horizon. Inequality (34) ensures that a field can be connected to at most one FPSO facility in each time period t, while (35) states that at most one FPSO-field connection is possible for a field f during the entire planning horizon T due to engineering considerations. Constraints (36) and (37) state that the expansion in the capacity of a FPSO facility and the connection between a field and a FPSO facility, respectively, in time period t can occur only if that FPSO facility has already been installed by that time period. Inequality (38) states that the oil flow rate per well from a field f to a FPSO facility in time period t will be zero if that FPSO-field connection is not available in that time period. Constraints (39)–(44) are the upper-bounding constraints on the installation and expansion capacities for FPSO facilities in time period t corresponding to the three different kinds of capacities mentioned earlier. The additional restrictions on the oil, liquid, and gas expansion capacities of FPSO facilities, (45)–(47), come from the fact that these expansion capacities should be less than a certain fraction (μ) of the initial built capacities, respectively.

$$\sum_{t \in T} b_{\text{fpso},t} \leq 1 \quad \forall \text{fpso} \tag{31}$$

$$\sum_{t \in T} b^{ex}_{\text{fpso},t} \leq 1 \quad \forall \text{fpso} \tag{32}$$

$$\sum_{t \in T} b^{c}_{f,\text{fpso},t} \leq 1 \quad \forall f, \text{fpso} \tag{33}$$

$$\sum_{\text{fpso}} b^{c}_{f,\text{fpso},t} \leq 1 \quad \forall f, t \tag{34}$$

$$\sum_{t \in T} \sum_{\text{fpso}} b^{c}_{f,\text{fpso},t} \leq 1 \quad \forall f \tag{35}$$

$$b^{ex}_{\text{fpso},t} \leq \sum_{\tau=1}^{t} b_{\text{fpso},\tau} \quad \forall \text{fpso}, t \tag{36}$$

$$b^{c}_{f,\text{fpso},t} \leq \sum_{\tau=1}^{t} b_{\text{fpso},\tau} \quad \forall f, \text{fpso}, t \tag{37}$$

$$x^{\text{well}}_{f,\text{fpso},t} \leq U^{\text{well,oil}}_{f,\text{fpso}} \sum_{\tau=1}^{t} b^{c}_{f,\text{fpso},\tau} \quad \forall f, \text{fpso}, t \tag{38}$$

$$QI^{\text{oil}}_{\text{fpso},t} \leq U^{\text{oil}}_{\text{fpso}} b_{\text{fpso},t} \quad \forall \text{fpso}, t \tag{39}$$

$$QI^{\text{liq}}_{\text{fpso},t} \leq U^{\text{liq}}_{\text{fpso}} b_{\text{fpso},t} \quad \forall \text{fpso}, t \tag{40}$$

$$QI^{\text{gas}}_{\text{fpso},t} \leq U^{\text{gas}}_{\text{fpso}} b_{\text{fpso},t} \quad \forall \text{fpso}, t \tag{41}$$

$$QE^{\text{oil}}_{\text{fpso},t} \leq U^{\text{oil}}_{\text{fpso}} b^{ex}_{\text{fpso},t} \quad \forall \text{fpso}, t \tag{42}$$

$$QE^{\text{liq}}_{\text{fpso},t} \leq U^{\text{liq}}_{\text{fpso}} b^{ex}_{\text{fpso},t} \quad \forall \text{fpso}, t \tag{43}$$

$$QE^{\text{gas}}_{\text{fpso},t} \leq U^{\text{gas}}_{\text{fpso}} b^{\text{exp}}_{\text{fpso},t} \quad \forall \text{fpso}, t \tag{44}$$

$$QE^{\text{oil}}_{\text{fpso},t} \leq \mu Q^{\text{oil}}_{\text{fpso},t-1} \quad \forall \text{fpso}, t \tag{45}$$

$$QE^{\text{liq}}_{\text{fpso},t} \leq \mu Q^{\text{liq}}_{\text{fpso},t-1} \quad \forall \text{fpso}, t \tag{46}$$

$$QE^{\text{gas}}_{\text{fpso},t} \leq \mu Q^{\text{gas}}_{\text{fpso},t-1} \quad \forall \text{fpso}, t \tag{47}$$

(f) ***Well drilling limitations***

The number of wells available for the production from a field is calculated from (48) as the sum of the wells available at the end of previous time period and the number of wells drilled at the beginning of time period t. The maximum number of wells that can be drilled over all the fields during each time period t and in each field f during complete planning horizon T are restricted by respective upper bounds in (49) and (50).

$$N^{\text{well}}_{f,t} = N^{\text{well}}_{f,t-1} + I^{\text{well}}_{f,t} \quad \forall f, t \tag{48}$$

$$\sum_{f} I^{\text{well}}_{f,t} \leq \text{UI}^{\text{well}}_{t} \quad \forall t \tag{49}$$

$$N^{\text{well}}_{f,t} \leq \text{UN}^{\text{well}}_{f} \quad \forall f, t \tag{50}$$

The non-convex MINLP model **(Model 1)** for offshore oilfield investment and operations planning involves constraint (1)–(50). In particular, constraints (9)–(13)

and (18) are nonlinear and non-convex constraints in the model that can lead to suboptimal solutions when solved with a method that assumes convexity (e.g., branch and bound, outer approximation).

Based on the following two theoretical properties, Gupta and Grossmann (2012a) reformulated this MINLP model to one that includes univariate polynomials representing reservoir profiles in terms of cumulative water and cumulative gas produced (Model 2):

Property 1. The area under the curve water-oil-ratio versus cumulative oil produced for a field (Fig. 7b) yields the cumulative amount of water produced.

Property 2. The area under the curve gas-oil-ratio versus cumulative oil produced for a field (Fig. 7c) yields the cumulative amount of gas produced.

In particular, the profiles for water-oil-ratio (Eq. 10) and gas-oil-ratio (Eq. 11) are integrated over the factional oil recovery to obtain the corresponding polynomials for cumulative water and cumulative gas produced to be used in Model 2. Furthermore, the improved MINLP formulation (Model 2) is converted to an MILP (Model 3) using piecewise linearization of these univariate polynomials for cumulative water and cumulative gas produced, and exact linearization of the bilinear term with integer variables (Eq. 18). The resulting MILP formulation allows solving the problem to global optimality efficiently; see Gupta and Grossmann (2012a) for details.

Table 1 summarizes the main features of these MINLP and reformulated MILP models. In particular, the reservoir profiles and respective nonlinearities involved in the models are compared in the table. Realistic instances involving 10 fields, 3 FPSO's and 20 years planning horizon have been solved and comparisons of the computational performance of the proposed MINLP and MILP formulations are presented in the paper. The computational efficiency of the proposed MINLP and MILP models have been further improved by a binary reduction scheme that yields

Table 1 Comparison of the nonlinearities involved in 3 model types

	Model 1	Model 2	Model 3
Model type	MINLP	MINLP	MILP
Oil deliverability	3rd order polynomial	3rd order polynomial	Piecewise linear
WOR	3rd order polynomial	–	–
GOR	3rd order polynomial	–	–
wc	–	4th order polynomial	Piecewise linear
gc	–	4th order polynomial	Piecewise linear
Bilinear terms	N^*x x^*WOR x^*GOR	N^*x	None
MILP reformulation	Not possible	Possible	Reformulated MILP

an order of magnitude reduction in the solution time. A large-scale example form Gupta and Grossmann (2012a) is explained in the results section of this chapter.

The more complex models with fiscal contracts and endogenous uncertainties can be derived from these deterministic formulations, as explained in Gupta and Grossmann (2012b, 2014a), respectively.

6 Incorporating Fiscal Contracts in Oilfield Planning

In this section, we incorporate the complex fiscal rules in the above MINLP/MILP models as described in Gupta and Grossmann (2012b) emphasizing the unified framework that we presented earlier (Fig. 1). Particularly, we consider the progressive production sharing agreements (PSAs) with ringfencing provisions (Fig. 3) that is one of the most generic forms of fiscal rules. Notice that the motivation for including progressive PSAs with ringfencing provisions is to consider the key elements of the most of the available contracts. Particular fiscal rules of interest can be modeled as the specific case of this representation as derived in Gupta and Grossmann (2012b).

The objective function now becomes to maximize the contractor's NPV which is the difference between total contractor's revenue share and total cost occurred over the planning horizon, taking discounting into consideration. The idea of cost recovery ceiling is included in terms of min function (m) to limit the amount of total oil produced each year that can be used to recover the capital and operational expenses. In particular, it represents the minimum of the cost oil recovery and a given fraction $(f^{\mathrm{CR}}_{rf,t})$ of the revenues. Constraint (m) is further converted to mixed-integer linear constraint using big-M formulation (see Gupta and Grossmann 2012b for details). This ceiling on the cost oil recovery is usually enforced to ensure early revenues to the government as soon as production starts.

$$\mathrm{CO}_{rf,t} = \min\left(\mathrm{CR}_{rf,t}, f^{\mathrm{CR}}_{rf,t} \cdot \mathrm{REV}_{rf,t}\right) \quad \forall rf, t \tag{m}$$

Moreover, a sliding scale based profit oil share of contractor linked to the cumulative oil production is also included in the model (Fig. 4). In particular, disjunction (**n**) is used to model this tier structure for profit oil split which states that the variable $Z_{rf,i,t}$ will be true if cumulative oil production by the end of time period t for ringfence rf lies between given tier thresholds $L^{\mathrm{oil}}_{rf,i} \leq xc_{rf,t} \leq U^{\mathrm{oil}}_{rf,i}$, i.e., tier i is active and split fraction $f^{\mathrm{PO}}_{rf,i}$ is used to determine the contractor share in that time period for ringfence rf. The disjunction (**n**) in the model is further reformulated into linear and mixed-integer linear constraints using the convex-hull formulation. In addition, the total cost in each year need to be disaggregated for each ringfence separately for the fiscal calculations that further increase the complexity of the model (see Gupta and Grossmann 2012b for details). The other constraints and

features remain the same as the MINLP and MILP models described in the earlier section.

$$\bigvee_{i}\begin{bmatrix} Z_{rf,i,t} \\ \mathrm{ConSh}_{rf,t}^{\mathrm{beforetax}} = f_{rf,i}^{\mathrm{PO}} \cdot \mathrm{PO}_{rf,t} \\ L_{rf,i}^{\mathrm{oil}} \le xc_{rf,t} \le U_{rf,i}^{\mathrm{oil}} \end{bmatrix} \quad \forall rf, t \tag{n}$$

The resulting MINLP/MILP models consider the trade-offs involved between investment and operating decisions and resulting royalties, profit shares that are paid to the government, and yields the maximum overall NPV for the contractor due to the improved decisions, as can be seen from the computational results in Sect. 8.

To improve the computational efficiency of these models, the logic constraints (o) and (p) that defines the tier sequencing are included in the model to tighten its relaxation as proposed in Gupta and Grossmann (2012b). These constraints can be expressed as integer linear inequalities. In addition, the valid inequalities (q), are also included that bounds the cumulative contractor's share in the cumulative profit oil by the end of time period t based on the sliding scale profit oil share and cost oil that has been recovered (see Gupta and Grossmann 2012a, b for derivation).

$$Z_{rf,i,t}^{s} \Rightarrow \bigwedge_{\tau=t}^{T} \neg Z_{rf,i',\tau}^{s} \quad \forall rf, i, i' < i, t, s \tag{o}$$

$$Z_{rf,i,t}^{s} \Rightarrow \bigwedge_{\tau=1}^{t} \neg Z_{rf,i',\tau}^{s} \quad \forall rf, i, i' > i, t, s \tag{p}$$

$$\begin{aligned} \sum_{\tau \le t} (\mathrm{Contsh}_{rf,\tau}^{\mathrm{beforetax},s}/\alpha_{\tau}) \le \sum_{i'=1}^{i' \le i} (f_{rf,i'}^{\mathrm{PO}} - f_{rf,i'-1}^{\mathrm{PO}}) \cdot (xc_{rf,t}^{s} - L_{rf,i'}) \\ - f_{rf,i^{\mathrm{end}}}^{\mathrm{PO}} \quad \cdot \sum_{\tau \le t} (\mathrm{CO}_{rf,\tau}^{s}/\alpha_{\tau}) \quad \forall rf, i, t, s \end{aligned} \tag{q}$$

However, due to the presence of additional binary variables in the model for the disjunction (**n**) and a weak relaxation of the problem, the large-scale instances still become intractable. To solve these large instances involving fiscal contracts, Gupta and Grossmann (2012b) also proposed reformation technique for this fiscal model that does not rely on disjunction (**n**) and yields a valid upper bound and decisions to be fixed in the fullspace model to obtain a feasible solution. Moreover, a heuristic approach is also proposed that yields near optimal solution in reasonable time for the intractable instances of the problem.

Notice that the deterministic model with fiscal considerations presented here can also be used as the basis for the stochastic programming approaches explained in the next section to incorporate uncertainty in the model parameters under the unified framework (Fig. 1). Optimal investment and operations decisions, the computational impact of adding a typical progressive production sharing agreement

(PSA) terms and effectiveness of the proposed approaches is demonstrated in the results section with a small example.

7 Incorporating Endogenous Uncertainty in Oilfield Development Planning

In this section, we consider the oilfield development planning problem under endogenous uncertainties in the field parameters with/without fiscal terms as described in Sect. 4. In particular, we use a multistage stochastic programming (MSSP) approach to model this uncertainty taking the above deterministic models as a basis.

Given that the uncertainty can only be revealed once we invest in a field, the resulting scenario tree is decisions dependent (Fig. 10a). Therefore, an alternative representation of the scenario tree is used (Ruszczynski 1997), where each scenario is represented by a set of unique nodes (Fig. 10b). The horizontal lines connecting nodes in time period *t*, mean that nodes are identical as they have the same information, and those scenarios are said to be indistinguishable in that time period. These horizontal lines correspond to the non-anticipativity (NA) constraints in the model that link different scenarios and prevent the problem from being decomposable. The alternative scenario tree representation allows to model the uncertainty in the problem more effectively (Goel and Grossmann 2006; and Gupta and Grossmann 2011).

7.1 Multistage Stochastic Formulation

The mixed-integer linear disjunctive multistage stochastic program with endogenous uncertainties for oilfield development problem using the alternative scenario-tree can be represented in the compact form **(MD)**. The detailed model and solution approaches are proposed in our recent paper, Gupta and Grossmann (2014a).

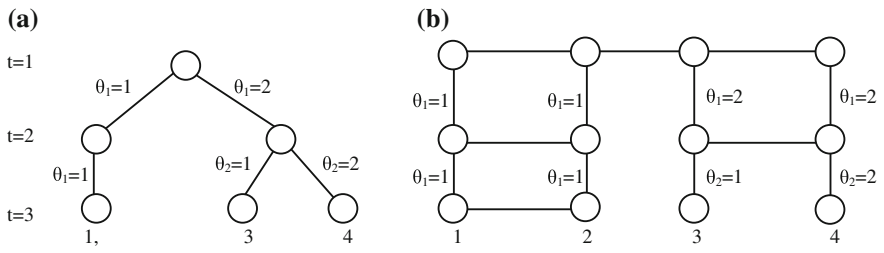


Fig. 10 Tree representations for discrete uncertainties over three stages

$$(\mathbf{MD}) \quad \max \quad z = \sum_{s \in S} p^s \sum_{t \in T} c_t x_t^s \tag{51}$$

$$\text{s.t.} \quad \sum_{\tau \le t} A_\tau^s x_\tau^s \le a_t^s \quad \forall t, s \tag{52}$$

$$x_t^s = x_t^{s'} \quad \forall t \in T_I, \forall s, s' \in S \tag{53}$$

$$Z_t^{s,s'} \Leftrightarrow F(x_1^s, x_2^s \ldots x_{t-1}^s) \quad \forall t \in T_C, \forall s, s' \in S \tag{54}$$

$$\begin{bmatrix} Z_t^{s,s'} \\ x_t^s = x_t^{s'} \end{bmatrix} \vee \left[\neg Z_t^{s,s'}\right] \quad \forall t \in T_C, \forall s, s' \in S \tag{55}$$

$$x_{jt}^s \in \{0, 1\} \quad \forall t, s, \forall j \in J' \tag{56}$$

$$x_{jt}^s \in R \quad \forall t, s, \forall j \in J \backslash J' \tag{57}$$

The objective function (51) in the above model **(MD)** maximizes the expectation of an economic criterion such as NPV or total contractor share over the set of scenarios $s \in S$, and over a set of time periods $t \in T$. For a particular scenario s, inequality (52) represents constraints that govern decisions x_t^s in time period t and link decisions across time periods. Notice that these constraints correspond to the constraints (2)–(50) in the basic deterministic model we outlined in Sect. 5. Moreover, in the case of the model with fiscal contracts, additional constraints as in Sect. 6 need to be included (see Gupta and Grossmann 2014a for details). Non-anticipativity (NA) constraints for initial time periods $T_I \subset T$ are given by Eq. (53) for each scenario pair (s, s') to ensure the same decisions in all the scenarios. The conditional NA constraints are written for the later time periods $T_C \subset T$ in terms of logic propositions (54) and disjunctions (55). Notice that the set of initial time periods T_I may include first few years of the planning horizon until uncertainty cannot be revealed, while T_C represents the rest of the time periods in the planning horizon. The function $F(x_1^s, x_2^s \ldots x_{t-1}^s)$ in Eq. (54) is an uncertainty resolution rule for a given pair of scenarios s and s' that determines the value of the corresponding boolean variable $Z_t^{s,s'}$ based on the decisions that have been implemented so far. The variable $Z_t^{s,s'}$ is further used in disjunction (55) to ensure the same decisions in scenarios s and s' if these are still indistinguishable in time period t. Equations (56)–(57) define the domain of the discrete and continuous variables in the model.

Notice that the model with reduced number of scenario pairs (s, s') that are sufficient to represent the non-anticipativity constraints can be obtained from model (MD) after applying the three properties presented in the paper by Gupta and Grossmann (2011). These properties are defined on the basis of symmetry, adjacency, and transitivity relationship among the scenarios. The reduced model

(MDR) can be formulated as follows, where P_3 is the set of minimum number of scenario pairs that are required to represent non-anticipativity in each time period t,

$$(\textbf{MDR}) \quad \min \quad z = \sum_{s \in S} p^s \sum_{t \in T} c_t x_t^s \tag{51}$$

$$\text{s.t.} \quad \sum_{\tau \le t} A_\tau^s x_\tau^s \le a_t^s \quad \forall t, s \tag{52}$$

$$x_t^s = x_t^{s'} \quad \forall t \in T_I, \forall (s, s') \in P_3 \tag{53a}$$

$$Z_t^{s,s'} \Leftrightarrow F(x_1^s, x_2^s \ldots x_{t-1}^s) \quad \forall t \in T_C, \forall (s, s') \in P_3 \tag{54a}$$

$$\begin{bmatrix} Z_t^{s,s'} \\ x_t^s = x_t^{s'} \end{bmatrix} \vee \left[\neg Z_t^{s,s'} \right] \quad \forall t \in T_C, \forall (s, s') \in P_3 \tag{55a}$$

$$x_{jt}^s \in \{0, 1\} \quad \forall t, s, \forall j \in J' \tag{56}$$

$$x_{jt}^s \in R \quad \forall t, s, \forall j \in J \backslash J' \tag{57}$$

The mixed-integer linear disjunctive model (MDR) can further be converted to a mixed-integer linear programming model **(MLR)**. First, the logic constraints (54a) are re-written as the mixed-integer linear constraints Eq. (54b) based on the uncertainty resolution rule, where $z_t^{s,s'}$ is a binary variable that takes a value of 1 if scenario pair (s, s') is indistinguishable in time period t, and zero otherwise. The disjunction (55a) can then be converted to mixed-integer linear constraints (55b) and (55c) using the big-M formulation. The resulting mixed-integer linear model (MLR) includes constraints (51), (52), (53a), (54b), (55b), (55c), (56), and (57).

$$B_t^s x_t^s + C_t^s z_t^{s,s'} \le d_t^s \quad \forall t \in T_C, \forall (s, s') \in P_3 \tag{54b}$$

$$-M(1 - z_t^{s,s'}) \quad \le x_t^s - x_t^{s'} \quad \forall t \in T_C, \forall (s, s') \in P_3 \tag{55b}$$

$$M(1 - z_t^{s,s'}) \quad \ge x_t^s - x_t^{s'} \quad \forall t \in T_C, \forall (s, s') \in P_3 \tag{55c}$$

Figure 11 represents the block angular structure of model (MLR), where we can observe that the initial (Eq. 53a) and conditional (Eqs. 54b, 55b, and 55c) non-anticipativity constraints link the scenario subproblems (Eq. 52), i.e., these are the complicating constraints in the model. However, this structure allows decomposing the fullspace problem into smaller subproblems by relaxing the linking constraints. It should be noted that the NACs (especially conditional NACs) represent a large fraction of the total constraints in the model.

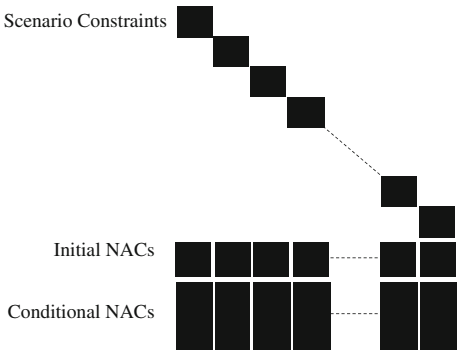


Fig. 11 Structure of a typical multistage stochastic program with endogenous uncertainties

7.2 *Standard Lagrangian Decomposition Approach*

The reduced model (MLR) is composed of scenario subproblems connected through initial (Eq. 53a) and conditional (Eqs. 54b, 55b, and 55c) NA constraints. If these NA constraints are either relaxed or dualized using Lagrangian decomposition, then the problem decomposes into smaller subproblems that can be solved independently for each scenario within an iterative scheme for the multipliers as described in Caroe and Schultz (1999) and in Gupta and Grossmann (2011). In this way, we can effectively decompose the large-scale problems in this class. The Lagrangian decomposition algorithm of Fig. 12 for MSSP with endogenous uncertainties as proposed in Gupta and Grossmann (2011) involves obtaining the upper bound (UB) by solving the Lagrangian problem **(L1-MLR)** with fixed multipliers $\lambda_t^{s,s'}$. Whereas, the Lagrangian problem (L1-MLR) is formulated from the mixed-integer linear reduced model (MLR) by relaxing all the conditional NA constraints (54b), (55b), and (55c) and dualizing all the initial NA constraints (53a) as penalty terms in the objective function. It gives rise to the subproblems for each scenario $s \in S$ **(L1-MLR$^{\mathbf{s}}$)** that can be solved in parallel.

$$(L1-MLR) \quad \max \quad \sum_{s\in S} p^s \sum_{t\in T} c_t x_t^s + \sum_{t\in T_1} \sum_{(s,s')\in P_3} \lambda_t^{s,s'}(x_t^s - x_t^{s'}) \tag{51a}$$

$$\text{s.t.} \quad \sum_{\tau\le t} A_\tau^s x_\tau^s \le a_t^s \quad \forall t, s \tag{52}$$

$$x_{jt}^s \in \{0,1\} \qquad \forall t, s, \forall j \in J' \tag{56}$$

$$x_{jt}^s \in R \quad \forall t, s, \forall j \in J \backslash J' \tag{57}$$

$$(L1-MLR^{\mathbf{s}}) \quad \max \quad \sum_{t\in T} p^s c_t x_t^s + \sum_{t\in T_1} x_t^s \Big(\sum_{\substack{(s,s')\in P_3 \\ s<s'}} \lambda_t^{s,s'} - \sum_{\substack{(s',s)\in P_3 \\ s>s'}} \lambda_t^{s',s} \Big) \tag{51b}$$

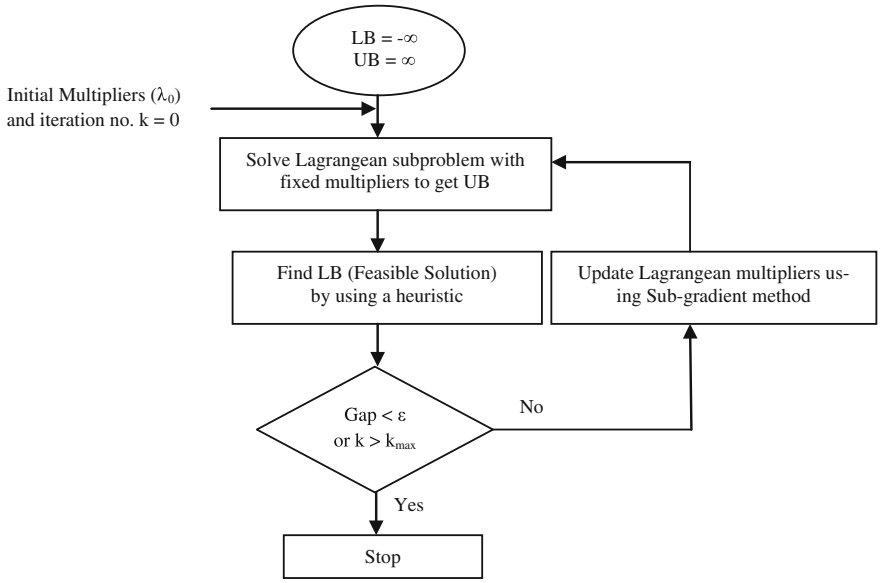


Fig. 12 Lagrangian decomposition algorithm

$$\text{s.t.} \quad \sum_{\tau \le t} A^s_\tau x^s_\tau \le a^s_t \quad \forall t \tag{52a}$$

$$x^s_{jt} \in \{0,1\} \quad \forall t, \forall j \in J' \tag{56a}$$

$$x^s_{jt} \in R \quad \forall t, \forall j \in J \backslash J' \tag{57a}$$

The lower bound (LB) or feasible solution is generated by using a heuristic based on the solution of the Lagrangian problem (L1-MLR). In this heuristic, we fix the decisions obtained from the above problem (L1-MLR) in the reduced problem (MLR) such that there is no violation of NA constraints and solve it to obtain the lower bound. The subgradient method by Fisher (1985) or an alternative update scheme (see Mouret et al. 2011; Oliveira et al. 2013; and Tarhan et al. 2013) is used during each iteration to update the Lagrangian multipliers. The algorithm stops when either a maximum iteration/time limit is reached, or the difference between the lower and upper bounds, LB and UB, is less than a prespecified tolerance.

Notice that the extended form of this method relying on duality-based branch and bound search has also been proposed in Goel and Grossmann (2006), Tarhan et al. (2009, 2011) to close the gap between the upper and the lower bounds.

7.2.1 Limitations

The major limitation of this Lagrangian Decomposition algorithm for endogenous uncertainty problems (Gupta and Grossmann 2011; Goel and Grossmann 2006; Tarhan et al. 2009, 2011) is that all the conditional NA constraints (54b), (55b), and (55c) are removed while formulating the scenario subproblems at the root node. These constraints represent a large fraction of the total constraints in the model, and can therefore have significant impact on the decisions. For instance, in Fig. 13, the scenario tree for the later time periods T_C (conditional NACs) can be constructed in several ways even though the initial NACs (for time periods T_I) are satisfied. Therefore, there can be several undesired consequences that can occur with this relaxation approach:

1. The dual bound at root node can be significantly weaker since a large amount of information from the conditional NACs is ignored. In particular, the initial NACs that represent only a first few time periods are considered (dualized) while formulating the subproblems at the root node. This means that the dynamics of the problem corresponding to the later periods is completely relaxed.
2. It is theoretically impossible to obtain a dual bound that is stronger than the optimal solution of the model without all conditional NACs at the root node.
3. The total number of nodes in the branch and bound search tree and the number of iterations required at each node can be quite large.
4. Since many constraints are relaxed from the model, a good heuristic is needed to generate a feasible solution based on the solution of the dual problem.
5. The number of subproblems grows with the number of uncertain parameters and their realizations in an exponential manner.
6. It is problem specific and non-intuitive to define the branching rules/variables in the tree search since there are several alternatives.

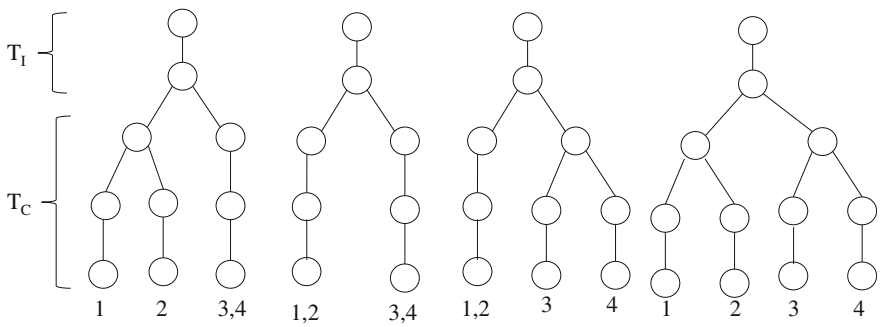


Fig. 13 Impact of conditional NACs on the scenario tree structure

7.3 Proposed Lagrangian Decomposition Approach

To overcome these limitations, we propose a new decomposition scheme (Gupta and Grossmann 2014b) that neither relaxes nor dualizes all the conditional NACs. The basic idea relies on decomposing the fullspace model into scenario group subproblems instead of individual scenarios. This allows keeping a subset of the NACs in the subproblems as constraints, while dualizing and relaxing the rest of the NACs. Therefore, it can be considered as a partial decomposition approach. A systematic methodology to construct these scenario groups for partial decomposition approach is described in our recent paper Gupta and Grossmann (2014b). In particular, the rules to formulate the scenario groups for the proposed algorithm can be summarized as follows:

1. Each scenario s occurs in only one of the scenario group S_g and every scenario is included in at least one of the groups. All the scenario groups $S_g \in G$ have equal number of scenarios. Therefore, the total number of scenarios equal to the number of scenario groups times the number of scenarios in each group, i.e., $|S| = |G| \cdot |S_g|$. Notice that here we assume the symmetry of the scenario groups to formulate the subproblems that have almost similar complexity. However, we can always consider an asymmetric approach as shown in Fig. 14 for the 4 scenario instance. Specifically, Fig. 14a, b decompose the problem into two scenario groups $\{g_1$: (1,2,3), g_2: (4)$\}$ and $\{g_1$: (1,3,4), g_2: (2)$\}$, respectively, where the subproblems with 3 scenarios should be more expensive to solve than the one with a single scenario.
2. Scenario groups S_g are formulated by first selecting an endogenous uncertain parameter and then taking those scenarios in a group which differ in the realization of only that particular uncertain parameter.
3. Since there can be many uncertain parameters $\{\theta_p\}$ each with its own scenario sets $(s_1, s_2, \ldots, s_k) \in L_p$, the selection of a particular scenario group is not unique.

   (i) Ideally, one may consider selecting a scenario group that provides the tightest initial bound compared to the others. However, unless all the combinations are tested, it is not obvious to select such a scenario group.

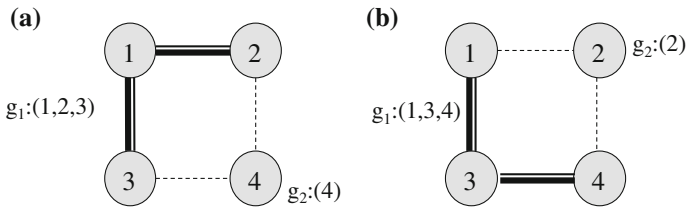


Fig. 14 Asymmetric scenario group decomposition

4. A relatively simpler approach can be to first solve each scenario independently, and selecting the scenario group corresponding to the uncertain parameter that has the largest total difference in the objective function values of the corresponding scenarios. This is due to the fact that most likely the corresponding NACs for those scenarios will be active at the optimal solution. Therefore, keeping these NACs in the subproblem as constraints should yield a tighter bound.
5. Even after selecting the scenario group correspond to an uncertain parameter $\{\theta_p\}$, it may still be difficult to solve the resulting scenario group subproblems. For instance if a parameter has many realizations, then each scenario group subproblem will have that many scenarios which may increase the computational expense. Therefore, one may further divide the scenario groups into subgroups and solve the resulting smaller problems.
   As an example, say if we have 2 uncertain parameters and 4 realizations of each parameter, there are a total of 16 scenarios. There are two possibilities of the scenario groups {(1,2,3,4), (5,6,7,8), (9,10,11,12), (13,14,15,16)} (Fig. 15c) and {(1,5,9,13), (2,6,10,14), (3,7,11,15), (4,8,12,16)} (Fig. 15d) according to the rules 1–3. Based on the problem characteristics, it may be difficult to solve each scenario group subproblem with four scenarios. Therefore, these groups can be

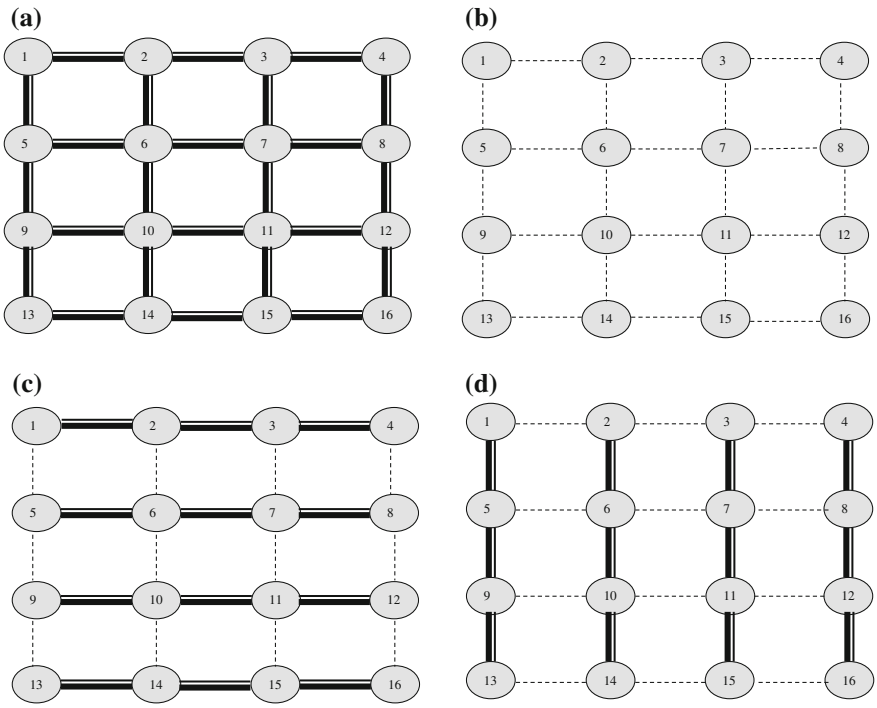


Fig. 15 2 parameters, 16 scenarios and its scenario/scenario group decomposition

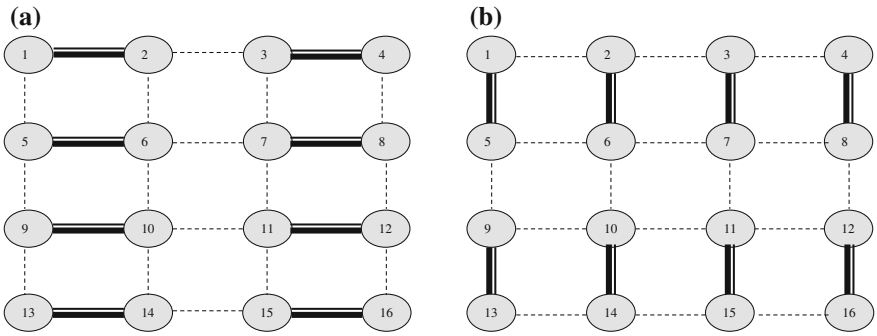


Fig. 16 Decomposition of the scenario groups into subgroups

further decomposed into a total of 8 scenario groups each with two scenarios, respectively (Fig. 16a, b). However, the quality of the bound may deteriorate since the corresponding conditional NACs need to be relaxed. Therefore, there is a trade-off between the quality of the bound and the complexity of solving a scenario group problem.

6. In general, if the problem is expensive to solve for each scenario, it is better to use scenario groups each with only few scenarios. On the other hand, if individual scenarios are not expensive to solve, then one may consider more scenarios in each group to improve the quality of the bound.

Based on the scenario groups that are constructed, we first present the corresponding reformulated Reduced (MILP) model. Notice that these scenario group partitions will be used to decompose the resulting reduced model into scenario group subproblems during the proposed Lagrangian decomposition algorithm. Let us consider that G is the set of scenario groups $S_g \in G$ that are selected based on the above rules, where each of these scenario groups S_g may have one or more scenarios. The reduced model (MLR) can now be represented as an equivalent model (**MLR$^{\mathbf{G}}$**) in terms of the scenario groups $S_g \in G$ where we disaggregate the total NACs for the scenario pairs that corresponds to the same scenario group $(s, s') \in S_g$ (i.e., Eqs. 53i, 54i, 55i, and 55j) with those which belong to the different scenario groups $(s \in S_g) \wedge (s' \notin S_g)$(i.e., Eqs. 53j, 54j, 55k, and 55l).

$$(\mathbf{MLR^G}) \quad \min \quad \sum_{S_g \in G} \left\{ \sum_{s \in S_g} p^s \sum_{t \in T} c_t x_t^s \right\} \tag{51i}$$

$$\text{s.t.} \quad \sum_{\tau \le t} A_\tau^s x_\tau^s \le a_t^s \quad \forall t, s \in S_g \in G \tag{52i}$$

$$x_t^s = x_t^{s'} \quad \forall t \in T_I, \forall (s, s') \in P_3, s, s' \in S_g \in G \tag{53i}$$

$$B_t^s x_t^s + C_t^s z_t^{s,s'} \le d_t^s \quad \forall t \in T_C, \forall (s,s') \in P_3,\ s,s' \in S_g \in G \tag{54i}$$

$$-M(1 - z_t^{s,s'}) \le x_t^s - x_t^{s'} \quad \forall t \in T_C, \forall (s,s') \in P_3,\ s,s' \in S_g \in G \tag{55i}$$

$$M(1 - z_t^{s,s'}) \ge x_t^s - x_t^{s'} \quad \forall t \in T_C, \forall (s,s') \in P_3, s,s' \in S_g \in G \tag{55j}$$

$$x_t^s = x_t^{s'} \quad \forall t \in T_I, \forall (s,s') \in P_3,\ (s \in S_g) \wedge (s' \notin S_g), S_g \in G \tag{53j}$$

$$B_t^s x_t^s + C_t^s z_t^{s,s'} \le d_t^s \quad \forall t \in T_C, \forall (s,s') \in P_3,\ (s \in S_g) \wedge (s' \notin S_g), S_g \in G \tag{54j}$$

$$-M(1 - z_t^{s,s'}) \le x_t^s - x_t^{s'} \quad \forall t \in T_C, \forall (s,s') \in P_3,\ (s \in S_g) \wedge (s' \notin S_g), S_g \in G \tag{55k}$$

$$M(1 - z_t^{s,s'}) \ge x_t^s - x_t^{s'} \quad \forall t \in T_C, \forall (s,s') \in P_3,\ (s \in S_g) \wedge (s' \notin S_g), S_g \in G \tag{55l}$$

$$x_{jt}^s \in \{0,1\} \quad \forall t, \forall s \in S_g \in G, \forall j \in J' \tag{56i}$$

$$x_{jt}^s \in R \quad \forall t, \forall s \in S_g \in G, \forall j \in J \backslash J' \tag{57i}$$

The Lagrangian problem (**L4-MLR$^{\mathbf{G}}$**) corresponding to the model (MLR$^{\mathrm{G}}$) can be formulated by dualizing only those initial NAC constraints for the pairs of scenarios (s, s') that link the two scenario groups, i.e., Eq. (53j), and removing the corresponding conditional NACs (Eqs. 54j, 55k, and 55l). Therefore, the initial and conditional NACs (Eqs. 53i, 54i, 55i, and 55j) among the scenario pairs (s, s') that belong to the same scenario group remain in the Lagrangian problem as explicit constraints.

$$(\mathbf{L4-MLR^G}) \quad \min \sum_{S_g \in G} \left\{ \sum_{s \in S_g} p^s \sum_{t \in T} c_t x_t^s + \sum_{t \in T_1} \sum_{\substack{(s,s') \in P_3 \\ s.t.(s \in S_g) \wedge (s' \notin S_g)}} \lambda_t^{s,s'} (x_t^s - x_t^{s'}) \right\} \tag{51j}$$

$$\text{s.t.} \quad \sum_{\tau \le t} A_\tau^s x_\tau^s \le a_t^s \quad \forall t, s \in S_g \in G \tag{52i}$$

$$x_t^s = x_t^{s'} \quad \forall t \in T_I, \forall (s,s') \in P_3,\ s,s' \in S_g \in G \tag{53i}$$

$$B_t^s x_t^s + C_t^s z_t^{s,s'} \le d_t^s \quad \forall t \in T_C, \forall (s,s') \in P_3, s,s' \in S_g \in G \tag{54i}$$

$$-M(1-z_t^{s,s'}) \le x_t^s - x_t^{s'} \quad \forall t \in T_C, \forall (s,s') \in P_3, s, s' \in S_g \in G \tag{55i}$$

$$M(1-z_t^{s,s'}) \ge x_t^s - x_t^{s'} \quad \forall t \in T_C, \forall (s,s') \in P_3, s, s' \in S_g \in G \tag{55j}$$

$$x_{jt}^s \in \{0,1\} \quad \forall t, \forall s \in S_g \in G, \forall j \in J' \tag{56i}$$

$$x_{jt}^s \in R \quad \forall t, \forall s \in S_g \in G, \forall j \in J \backslash J' \tag{57i}$$

In contrast to the previous approach, we can observe that the main idea in the proposed decomposition approach is that instead of removing all the conditional NACs from the model, we only remove a subset of conditional NACs from the model and dualize a subset of the initial NACs in the objective function instead of dualizing all the initial NACs while formulating the Lagrangian problem (L4-MLR^{G}). This results in the decomposition of the reduced model (MLR) into scenario group subproblems (**L4-MLR^{Gs})** rather than individual scenarios in the previous case. Therefore, we also refer to it as a partial decomposition approach.

$$(\mathbf{L4-MLR^{Gs}}) \quad \min \sum_{s \in S_g} p^s \sum_{t \in T} c_t x_t^s + \sum_{t \in T_1} x_t^s \left(\sum_{\substack{(s,s') \in P_3, s < s' \\ s.t.(s \in S_g) \wedge (s' \notin S_g)}} \lambda_t^{s,s'} - \sum_{\substack{(s',s) \in P_3, s > s' \\ s.t.(s \in S_g) \wedge (s' \notin S_g)}} \lambda_t^{s',s} \right) \tag{51k}$$

$$\text{s.t.} \quad \sum_{\tau \le t} A_\tau^s x_\tau^s \le a_t^s \quad \forall t, s \in S_g \tag{52k}$$

$$x_t^s = x_t^{s'} \quad \forall t \in T_I, \forall (s,s') \in P_3, \; s, s' \in S_g \tag{53k}$$

$$B_t^s x_t^s + C_t^s z_t^{s,s'} \le d_t^s \quad \forall t \in T_C, \forall (s,s') \in P_3, s, s' \in S_g \tag{54k}$$

$$-M(1-z_t^{s,s'}) \le x_t^s - x_t^{s'} \quad \forall t \in T_C, \forall (s,s') \in P_3, s, s' \in S_g \tag{55m}$$

$$M(1-z_t^{s,s'}) \ge x_t^s - x_t^{s'} \quad \forall t \in T_C, \forall (s,s') \in P_3, s, s' \in S_g \tag{55n}$$

$$x_{jt}^s \in \{0,1\} \quad \forall t, \forall s \in S_g, \forall j \in J' \tag{56k}$$

$$x_{jt}^s \in R \quad \forall t, \forall s \in S_g, \forall j \in J \backslash J' \tag{57k}$$

The structure of model (L4-MLR^{G}) can be seen in Fig. 17, where each scenario group subproblem that contains its corresponding initial and conditional NACs can be solved independently, and where only a small fraction of the total initial and conditional NACs are dualized and removed, respectively. Since, the resulting subproblems capture the more relevant information, i.e., the one corresponding to

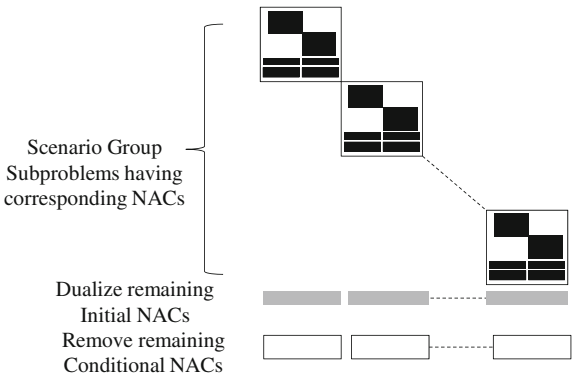


Fig. 17 Scenario decomposition approach in the proposed Lagrangian decomposition

the later time periods, the dual bound should be tighter. We can then state the following proposition (see Gupta and Grossmann 2014b for proof):

Proposition 1: *The dual bound obtained from the proposed Lagrangian problem (L4-MLR^G) at the root node is at least as tight as the dual bound obtained from the standard Lagrangian decomposition approach (L1-MLR), i.e., the model (L1-MLR) is a relaxation of the model (L4-MLR^G).*

The rest of the steps of the algorithm are similar to the standard Lagrangian decomposition (Fig. 12) where scenario group subproblems L4-MLR^{Gs} are solved during each iteration, and multipliers are updated using either subgradient method (Fisher 1985) or an alternative scheme as in Mouret et al. (2011), Oliveira et al. (2013), and Tarhan et al. (2013). Moreover, the algorithm can be further extended within a duality-based branch and bound search (as proposed in Goel and Grossmann 2006; Tarhan et al. 2009, 2011) if the gap between the LB and UB is still large. As will be shown in the results, the main advantage with the proposed approach is that the resulting dual bound is significantly strengthened at the root node itself since a large fraction of the NACs are included as explicit constraints in the subproblems. This will eventually reduce the number of iterations required to converge at each node and the total number of nodes in the branch and bound search.

8 Examples

In this section, we consider a variety of the examples for the oilfield development planning problem that covers deterministic, complex fiscal rules and stochastic features as discussed in the earlier sections.

8.1 Instance 1: Deterministic Case

We first present an example of the oilfield planning problem assuming that there are no fiscal contracts and no uncertainty in the model parameters. In particular, we compare the computational results of the various deterministic MINLP and MILP models proposed in Gupta and Grossmann (2012a).

An oilfield infrastructure is considered with 10 oil fields (Fig. 18) that can be connected to 3 FPSOs with 23 possible connections. There are a total of 84 wells that can be drilled in all of these 10 fields, and the planning horizon considered is 20 years, which is discretized into 20 periods of each 1 year of duration. We need to determine which of the FPSO facilities is to be installed or expanded, in what time period, and what should be its capacity of oil, liquid, and gas, to which fields it should be connected and at what time, and the number of wells to be drilled in each field during each time period. Other than these installation decisions, there are operating decisions involving the flow rate of oil, water, and gas from each field in each time period. The objective function is to maximize the total NPV over the given planning horizon.

The problem is solved using the DICOPT 2x-C solver for MINLP Models 1 and 2. These models were implemented in GAMS 23.6.3 and run on Intel Core i7 machine. The optimal solution of this problem that corresponds to the reduced MINLP Model 2-R solved with DICOPT 2x-C, suggests to install all the 3 FPSO facilities in the first time period with their respective liquid (Fig. 19a) and gas (Fig. 19b) capacities. These FPSO facilities are further expanded in the future when more fields come online or liquid/gas flow rates increases as can be seen from these figures.

After initial installation of the FPSO facilities by the end of time period 3, these are connected to the various fields to produce oil in their respective time periods for

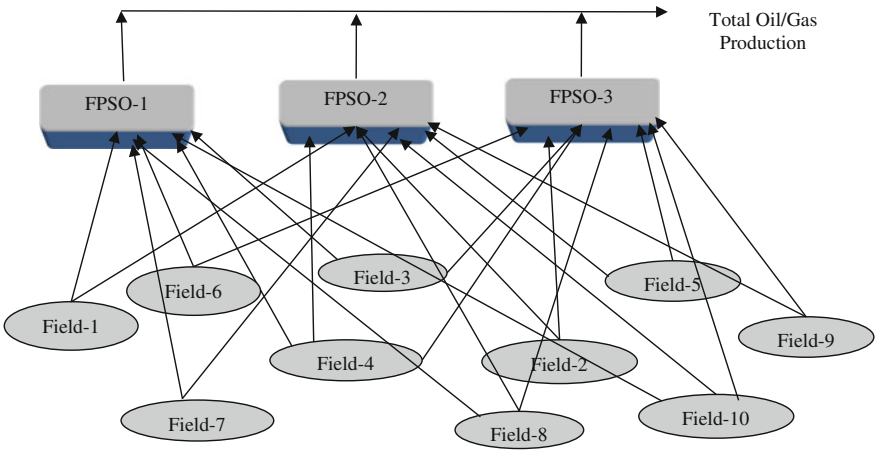


Fig. 18 Deterministic Instance 1 (10 Fields, 3 FPSOs, 20 years)

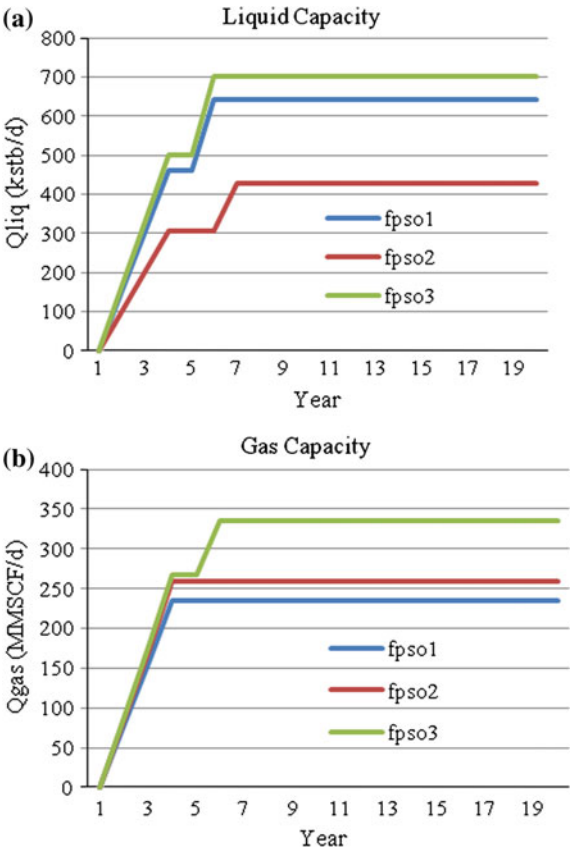


Fig. 19 FPSO installation and expansion schedule. **a** Liquid capacities of FPSO facilities. **b** Gas capacities of FPSO facilities

coming online as indicated in Fig. 20. The well drilling schedule for these fields in Fig. 21 ensures that the maximum number of wells drilling limit and maximum potential wells in a field are not violated in each time period t. We can observe from these results that most of the installation and expansions are in the first few time periods of the planning horizon.

Other than these investment decisions, the operating decisions are the production rates of oil and gas from each of the fields, and hence, the total flow rates for the installed FPSO facilities that are connected to these fields as can be seen from Fig. 22a, b. Notice that the oil flow rates increases initially until all the fields come online and then they start to decrease as the oil deliverability decreases when time progresses. Gas flow rate, which depends on the amount of oil produced, also follows a similar trend. The total NPV of the project is \$30946.39 M.

Tables 2 and 3 represent the results for the various model types considered for this instance. DICOPT performs best in terms of solution time and quality, even for the largest instance compared to other solvers as can be seen from Table 2. There are significant computational savings with the reduced models as compared to the

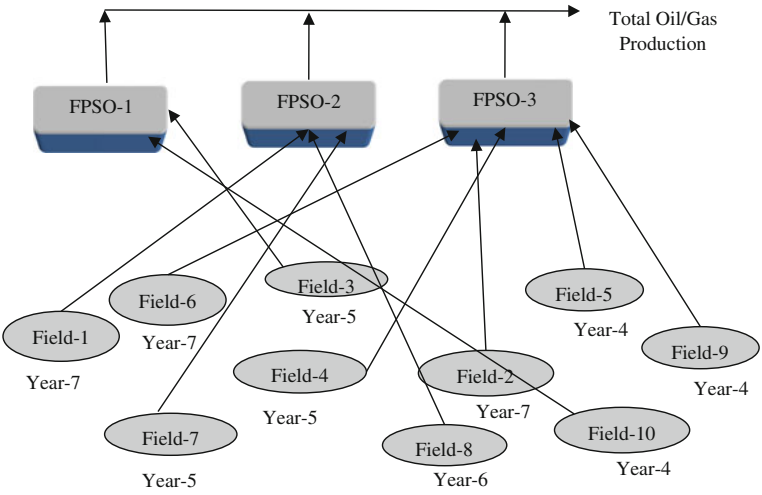


Fig. 20 FPSO-field connection schedule

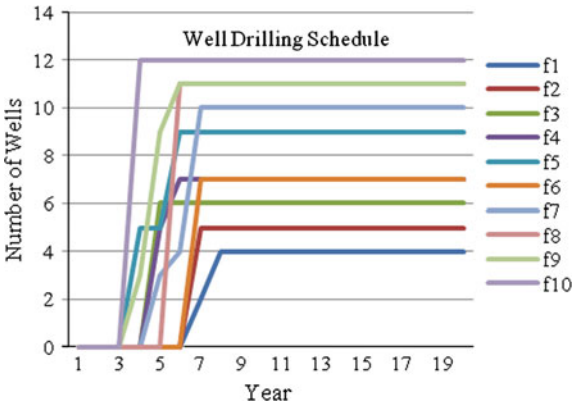


Fig. 21 Well drilling schedule for fields

original ones for all the model types in Table 3. Even after binary reduction of the reformulated MILP, Model 3-R becomes expensive to solve, but yields global solutions, and provides a good discrete solution to be fixed/initialized in the MINLPs for finding better solutions.

The optimal NPV that come from the Models 1 and 2 after fixing discrete variables based on MILP solution (even though it was solved within 10 % of optimality tolerance) are \$31329.81 and \$31022.48 M, respectively. These are the best solutions among all other solutions obtained in Table 2 for the respective MINLPs. Therefore, while the MINLP model may sometimes lead to near optimal solutions, the MILP approximation is an effective way to consistently obtain these solutions. Notice also that the optimal discrete decisions for Models 1 and 2 are

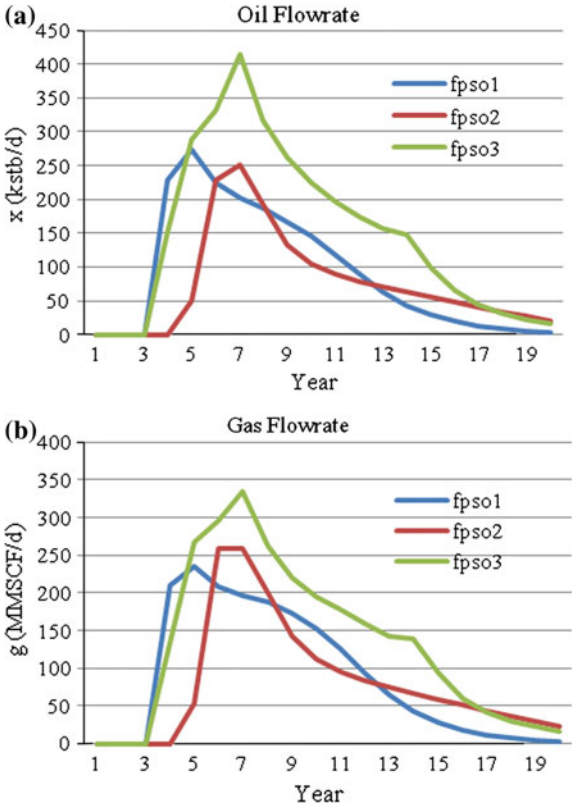


Fig. 22 Total flow rates from each FPSO facility. **a** Total oil flow rates from FPSO's. **b** Total gas flow rates from FPSO's

Table 2 Comparison of various models and solvers for Instance 1

		Model 1		Model 2
Constraints		5,900		10,100
Continuous Var.		4,681		6,121
Discrete Var.		851		851
Solver	Optimal NPV (Model 1) (million$)	Time (Model 1) (s)	Optimal NPV (Model 2) (million$)	Time (Model 2) (s)
DICOPT	31297.94	132.34	30562.95	114.51
SBB	30466.36	4973.94	30005.33	18152.03
BARON	31297.94	>72,000	30562.95	>72,000

very similar even though they are formulated in a different way. However, only Model 2 can be reformulated into an MILP problem that gives a good estimate of the near optimal decisions to be used for these MINLPs.

Notice that although the advantage from using the MILP formulation in terms of the NPV value is not very large for this instance since the required solution time is

Table 3 Comparison of models 1, 2, and 3 with and without binary reduction

	Model 1	Model 1-R	Model 2	Model 2-R	Model 3-R
Constraints	5,900	5,677	10,100	9,877	17,140
Continuous Var.	4,681	4,244	6,121	5,684	12,007
Discrete Var.	851	483	851	483	863
SOS1 Var.	0	0	0	0	800
NPV (million$)	31297.94	30982.42	30562.95	30946.39	30986.22
Time (s)	132.34	53.08	114.51	67.66	16295.26

*Model 1 and 2 solved with DICOPT 2x-C, Model 3 with CPLEX 12.2

large, but it does yield a global solution that is difficult to obtain for the MINLPs. In addition, when we increase the complexity of the basic deterministic model to fiscal contracts and/or stochastic model, the advantage of MILP formulation becomes more apparent due to the availability of the robust MILP solvers compared to MINLP. In that case, the solutions from the local MINLP solvers are significantly suboptimal and they cannot provide any guarantees of the valid upper bound in the Lagrangian decomposition unless each subproblem is solved to global optimality. Therefore, we use the MILP model for instances 2 and 3 that includes complex fiscal rules and/or uncertainties.

8.2 Instance 2: Development Planning with Complex Fiscal Rules

We demonstrate here the impact of adding a typical progressive production sharing agreement (PSA) terms for a deterministic case of the oilfield planning problem in terms of the investment and operating decisions, and the required computational effort.

In this instance, we consider 5 oilfields (Fig. 5) that can be connected to 3 FPSOs with 11 possible connections, see Gupta and Grossmann (2012b) for details. There are a total of 31 wells that can be drilled in these 5 fields, and the planning horizon considered is 20 years. There is a cost recovery ceiling of 50 % and 4 tiers (see Fig. 4) that are defined for profit oil fraction between the contractor and host government based on the cumulative oil production but there are no ringfences present. The problem is solved to maximize the NPV of the contractor's share after paying taxes and the corresponding optimal investment/operating decisions.

Table 4 compares the performance of the MILP (Model 3F) involving detailed connections and reduced MILP model (Model 3RF) that are the extension of the deterministic Models 3 and 3R, respectively, with progressive PSAs. The models are implemented in GAMS 23.6.3 and run on Intel Core i7, 4 GB RAM machine using CPLEX 12.2. We can observe that there is significant increase in the computational time with fiscal consideration for the MILP formulation Model 3F with

Table 4 Computational results for Instance 2 (Model 3F vs. Model 3RF)

Model	# of constraints	# of continuous variables	# of discrete variables	NPV ($Million)	Time (s)	Optimality gap (%)
Model 3F	9,474	6,432	727	2183.63	>36,000	<14
Model 3RF	9,363	6,223	551	2228.94	1,164	<2

this larger instance, which takes more than 10 h with a 14 % of optimality gap as compared to the reduced MILP model (Model 3RF), which terminates the search with a 2 % gap in reasonable time.

The optimal solution from Model 3RF suggests installing 1 FPSO facility (FPSO3) with expansions in the future (see Fig. 23), while Fig. 24 represents the well drilling schedule for this example. The tiers 2, 3, and 4 for profit oil split become active in years 6, 8, and 12, respectively, based on the cumulative oil production profile during the given planning horizon. Notice that the optimal solution of this problem fails to develop field 1, which is not intuitive. The reason for not developing field 1 is that the size of the field 1 is quite small as compared to the other fields and the superstructure we consider does not allow connecting field 1 to FPSO 3, which is the only FPSO that is installed. Therefore, based on the superstructure and field size, it is not worth to install an additional FPSO to produce from this field after paying government share. In contrast, the solution from the

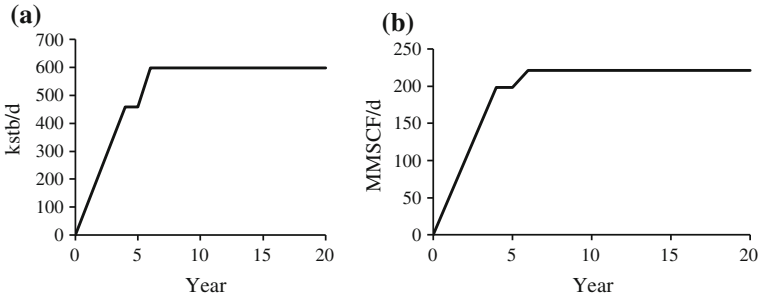


Fig. 23 Optimal liquid and gas capacities of FPSO 3 facility for Instance 2. **a** Liquid capacity. **b** Gas capacity

Fig. 24 Optimal well drilling schedule for Instance 2

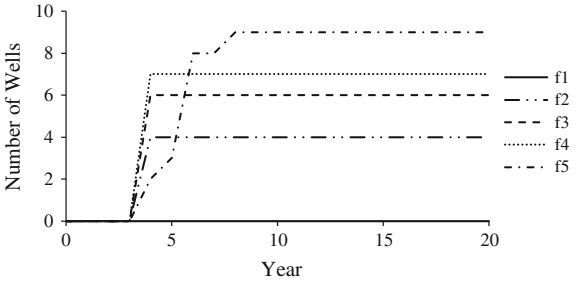


Table 5 Comparison of number of tiers versus solution time for Model 3RF

# of tiers	Time (s)
4	1,164
2	694
1	73
No fiscal rules	190

simple NPV based optimization suggests exploring field 1 as well since it is worth in that case to install 2 small FPSO facilities and also produce from field 1 given that the trade-offs due to fiscal rules are neglected. Whereas, the total NPV of the contractor's share in this case is lower than the optimal solution of Model 3RF ($1914.71 M vs. $2228.94 M). Therefore, we can observe that incorporating fiscal terms within development planning can yield significantly different investment and operations decisions, and higher profit compared to a simple NPV based optimization.

Note that fiscal terms without tier structure, for instance fixed percentage of profit share, royalty rates, often reduces the computational expense of solving the deterministic model directly without any fiscal terms instead. Surprisingly, the problem with flat 35 % of the profit share of contractor is solved in 73 s which is even smaller than the solution time for deterministic case without any fiscal terms (190 s). On the other hand, the problem with 2 tiers instead of 4 as considered above is solved in 694 s which is more than the model without fiscal terms and less than the model with 4 tiers as can be seen in Table 5. Therefore, the increase in computational time while including fiscal rules within development planning, is directly related to the number of tiers (levels) that are present in the model to determine the profit oil shares or royalties.

Table 6 compares the further improvements in the solution time for Model 3RF (1,164 s) after using the reformulation/approximation techniques and strategies that are proposed in Gupta and Grossmann (2012b), and reviewed in Sect. 6. In particular, the tighter formulation Model 3RF-L that is obtained after including logic constraint (o)–(p) and valid inequalities (q) is solved in one fourth of the time than Model 3RF. Notice that these MILP models are solved with a 2 % optimality tolerance yielding a slightly different objective values for Model 3RF and Model 3RF-L. Model 3RI, which relaxes the disjunction (n), can be solved more than 20 times faster than the original Model 3RF. Although the solution obtained is a relaxed one (upper bound of 2591.10), it gives the optimal investment decisions

Table 6 Results for Instance 2 after using various solution strategies

Model	# of constraints	# of continuous variables	# of discrete variables	NPV ($Million)	NPV after fixing decisions in Model 3RF ($Million)	Time (s)
Model 3RF	9,363	6,223	551	2228.94	–	1,164
Model 3RF-L	11,963	6,223	551	2222.40	–	275
Model 3RI-A	8,803	5,903	471	2197.63	2228.94	82
Model 3RI	8,803	5,903	471	2591.10	2228.94	48

Table 7 Results for Instance 2 with ringfencing provisions

Model	# of constraints	# of continuous variables	# of discrete variables	NPV ($Million)	NPV after fixing decisions in Model 3RF ($Million)	Time (s)	Optimality gap (%)
Model 3RF	14,634	9,674	651	2149.39	–	>36,000	<15.4
Model 3RF-L	19,834	9,674	651	2161.27	–	3,334	<2
Model 3RI-A	13,514	9,034	491	2148.90	2142.75	134	<2
Model 3RI	13,514	9,034	491	2533.06	2151.75	112	<2

that result in the same solution as we obtained from solving Model 3RF directly. The approximate version of this Model 3RI-A, takes only 82 s as compared to Model 3RF (1164 s) and yields the optimal solution after we fix the decisions from this model in the original one. Notice that the quality of the approximate solution itself is very good ($\sim$1.5 % accurate) and both relaxed/approximate models are even $\sim$3 times faster than the model without any fiscal terms (Model 3R) that takes 190 s. Therefore, the approximate model can be use as a good heuristic to solve the large instances of the problem.

In addition, when we consider two ringfences for this Instance 2 (see Fig. 5) where progressive PSA terms are defined for each of these ringfences separately. Based on the computational performance of the Model 3RF as compared to Model 3F, we only show the results for Model 3RF, which is more efficient.

Table 7 compares the results for various models for this oilfield development instance involving ringfencing provisions. We can observe that including ringfencing provisions makes Model 3RF expensive to solve (>10 h), compared to the previous instance without any ringfences that required only 1,164 s. This is due to the additional binary variables that are required in the model for each of the two ringfences, their trade-offs and FPSO cost disaggregation. In contrast, since Models 3RI and 3RI-A do not need binary variable for the sliding scale in disjunction, they solve much faster than Model 3RF (>300 times faster) and Model 3RF-L ($\sim$30 times faster). Notice that even after including ringfencing provisions, these two models are faster than the simple NPV based Model 3R. This is due to the trade-off from the fiscal part in the simple NPV based model without binary variables for the sliding scale.

Notice that Model 3RI and 3RI-A are solved here in one of the most generic forms of the fiscal terms where the solutions may not be the global optimal, but the relaxed Model 3RI, which provides a valid upper bound, also allows to compare the solution quality. The optimal NPV after ringfencing provisions is lower as compared to the one without ringfencing provisions due to the additional restrictions it imposes on the revenue and cash flows.

8.3 *Instance 3: Stochastic Case*

In this example, we consider the planning of offshore oilfield under decision-dependent uncertainty in the field parameters, which resolves as a function of investment and operating decisions as described in Sect. 7. Moreover, we also extend this MSSP instance to include the complex fiscal rules.

We consider the multistage stochastic MILP model by Gupta and Grossmann (2014a) as outlined in Sect. 7 for maximizing the expected NPV in the development planning of an offshore oilfield, which is an extension of the deterministic MILP model (Model 3) presented in Gupta and Grossmann (2012a). The model for this oilfield planning instance is implemented in GAMS 23.6.3 and run on Intel Core i7, 4 GB RAM machine using CPLEX 12.2 solver. In particular, we consider three oilfields, three potential FPSOs, and nine possible connections among field-FPSO (Fig. 25). A total of 30 wells can be drilled in the fields and the planning horizon is 10 years. Field 3 has a recoverable oil volume (field size) of 500 MMbbls. However, there is uncertainty in the size of fields 1 and 2 where each one has two possible realizations (low, high) with equal probability. Therefore, there are a total of 4 scenarios each with a probability of 0.25 (see Table 8).

The problem is to determine the investment (FPSO installations and expansions, field-FPSO connections and well drilling) and operating decisions (oil production rate) for this infrastructure with an objective to maximize the total expected NPV (ENPV) over the planning horizon. The optimal ENPV for this problem is \$11.50 × 10^9 when the reduced model (MLR) is solved in fullspace, and requires 1184 s. Table 9 represents the model statistics for this instance. The solution suggests

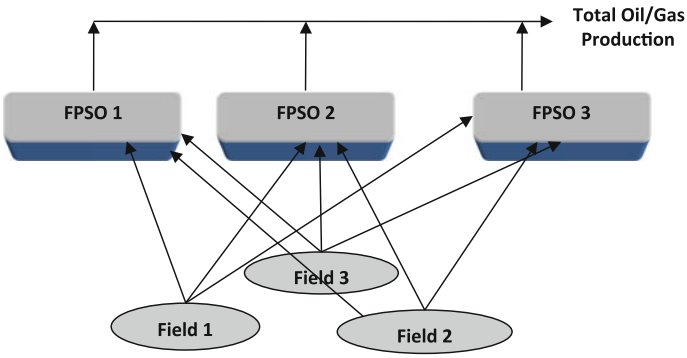


Fig. 25 Stochastic Instance 3 (3 oilfields, 4 scenarios)

Table 8 Stochastic Instance 3 (3 oilfields, 4 scenarios)

Scenarios	s1	s2	s3	s4
Field 1 size (MMbbls)	57	403	57	403
Field 2 size (MMbbls)	80	80	560	560
Scenario probability	0.25	0.25	0.25	0.25

Table 9 Model statistics for the 3 oilfield example (Instance 3)

Problem type	Number of constraints	Continuous variables	Discrete variables	SOS1 variables
Reduced model (MLR)	16,473	9,717	876	240
Individual scenario	3,580	2,390	179	60

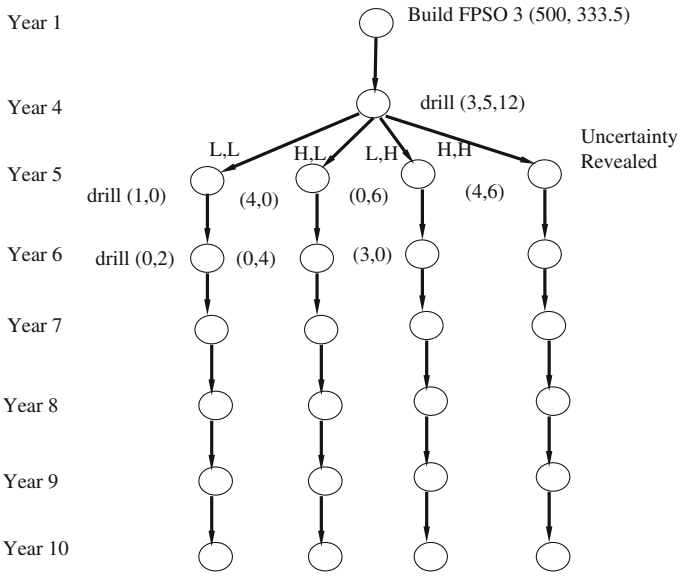


Fig. 26 Optimal solution for 3 oilfield example (Instance 3)

installing only FPSO 3 in the first year (see Fig. 26) and when the facility is available to produce at the beginning of year 4 due to lead time, 3, 5, and 12 wells are drilled in field 1, 2, and 3, respectively. Since, field 1 and 2 are uncertain, therefore based on the realization of the uncertainty in their field sizes, more wells are drilled in these fields in the future for the favorable scenarios compared to the unfavorable outcomes, whereas no more wells are drilled in field 3. We can observe that the optimal scenario tree is decision-dependent which is not known a priori (Fig. 26).

Figure 27 demonstrates the progress of the bounds obtained at the root node using standard Lagrangian decomposition described in Sect. 7.2 that relies on dualizing the initial NACs and removing the conditional NACs. A termination criterion of either 1 % gap or 20 iterations is used. We can observe that the problem can be solved in $\sim$1 % optimality tolerance in only 466 s compared to the fullspace model that takes 1184 s. However, the solution of the expected value problem considering the mean value of the field sizes is $\$11.28 \times 10^9$. Therefore, the value

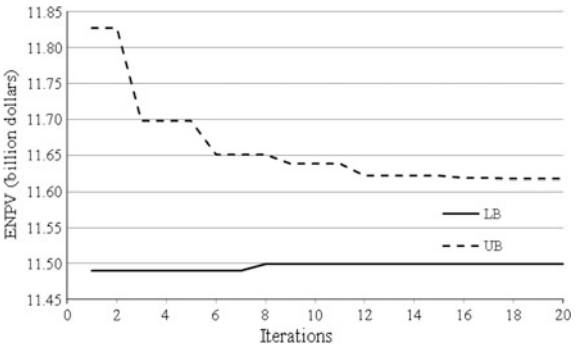


Fig. 27 Standard Lagrangian decomposition results for 3 oilfield example (Instance 3)

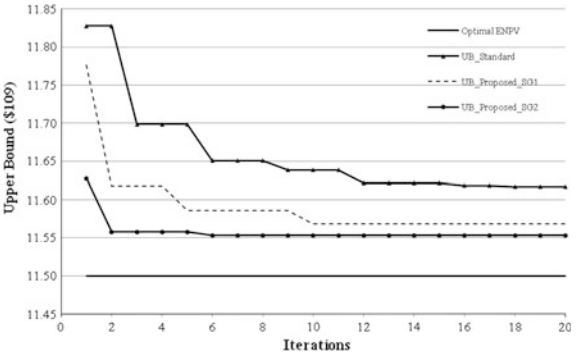


Fig. 28 Comparison of the various decomposition schemes for oilfield example (Instance 3)

of stochastic solution for this case is $\sim$2 % and the solution is more robust as compared to the deterministic approach (Gupta and Grossmann 2014a).

Figure 28 compares the performance of the UBs obtained at the root node using standard Lagrangian decomposition (Sect. 7.2) based on dualizing the initial NACs and removing the conditional NACs with the proposed partial decomposition approach (Sect. 7.3). A termination criterion of either 1 % gap or 20 iterations is used. The proposed decomposition algorithm based on scenario groups SG1: {(1,2), (3,4)}, and SG2: {(1,3), (2,4)} yield stronger upper bounds, $\$11.59 \times 10^9$ and $\$11.56 \times 10^9$ respectively, than the standard Lagrangian decomposition ($\$11.62 \times 10^9$). Additionally, the total computational effort is less with the proposed approach since only two subproblems need to be solved at each iteration, and only few iterations are needed to satisfy a 1 % of optimality tolerance (Table 10). SG2 performs better than SG1 as can be observed from the total variations in the scenario NPVs with respect to the change in the field sizes as calculated in Table 11 ($\$6.63 \times 10^9$ vs. $\$4.77 \times 10^9$). It is important to see that the quality of the UB from SG2 is similar in the first iteration with the quality of the UB obtained from the scenario subproblems after 20 iterations (see Fig. 28). Moreover, for clarity we only plotted the progress of the UBs with iterations and the optimal NPV in Fig. 28.

Table 10 Comparison of the various decomposition schemes for oilfield example (Instance 3)

	Standard	Proposed SG1	Proposed SG2
UB ($\$10^9$)	11.62	11.59	11.56
LB ($\$10^9$)	11.50	11.50	11.50
Solution time (s)	466	382	172
% Gap	1.02 %	<1 %	<1 %
# iterations	20	5	2

Table 11 Variations in the objective function value with uncertain parameters (Instance 3)

(a) Individual Scenario NPV		(b) Scenario groups NPV variations		
	NPV ($\$10^9$)		SG1	SG2
s1	8.95	s1–s2	2.44	–
s2	11.39	s3–s4	2.33	–
s3	12.32	s1–s3	–	3.37
s4	14.65	s2–s4	–	3.26
		Total NPV variations ($\$10^9$)	4.77	6.63

Table 12 Multistage stochastic instance with fiscal rules for 3 oilfield example (Instance 3)

Fullspace model					Lagrangian decomposition		
# Constraints	# Dis. Var.	# Cont. Var.	ENPV ($\$10^9$)	Time (s)	ENPV ($\$10^9$)	Sequential time (s)	Parallel time (s)
27,113	1,536	15,857	\$2.97 (>21 %)	>36,000	\$3.04 (0.7 %)	8,990	4,002

In addition, we also consider the extension of this 3 oilfield instance to the case where we include the progressive production sharing agreements with 15 years of planning horizon, and uncertain field sizes (field 1 and 2) having a total of 4 scenarios as described above. The resulting fullspace model becomes very difficult to solve (see Table 12). In particular, the best solution obtained after 10 h is \$2.97 $\times$ 10^9 with more than 21 % of optimality gap. On the other hand, Lagrangian decomposition can solve this problem in approximately 2 h for sequential implementation of the scenario subproblem solutions and in about 1 h for parallel implementation, and yields a higher ENPV \$3.04 $\times$ 10^9 within 0.7 % optimality tolerance.

Therefore, the importance of the decomposition algorithm, especially parallel solution of the scenario subproblems increases as more complexities are added to the deterministic problem such as fiscal contracts. Several multistage stochastic instances of the problems in this class with/without fiscal contracts involving many scenarios are presented in Gupta and Grossmann (2014a, b).

9 Conclusions

In this chapter, we have first reviewed the models and solution approaches that have been used in the literature for oil and gas field development problems with a particular focus on the type of infrastructure, fiscal rules, and uncertainties that have been considered. Given the limitations of the previous works, we first outlined the recently proposed deterministic multiperiod MINLP model by Gupta and Grossmann (2012a) for offshore oil and gas field infrastructure investment and operational planning. The model considers realistic details to be useful in practice such as multiple fields, three components (oil, water, and gas), facility expansions decisions, well drilling schedules, and nonlinear reservoir profiles. The reformulation of this model without bilinear terms and its further conversion into an MILP formulation that allow solving the problem to global optimality has also been summarized. Furthermore, a discussion has been presented on the extensions of this basic deterministic model to incorporate fiscal rules (Gupta and Grossmann 2012b) and/or endogenous uncertainties (Gupta and Grossmann 2014a) in a unified modeling framework, addressing the complications and ways to overcome them based on our recent work in this area. Numerical results on development planning problems with and without fiscal contracts and uncertainties are included to emphasize the efficiency of the models and solution approaches such as Lagrangian decomposition that are proposed in aforementioned papers. It is hoped that this chapter has shown that there has been very significant progress in the mathematical programming models and solution algorithms for offshore oilfield development planning. In the future, these approaches can serve as a basis to incorporate additional complexities such as uncertain oil/gas prices.

Acknowledgments The author would like to acknowledge Center for Advanced Process Decision-making, Carnegie Mellon University, Pittsburgh for the financial support of this work.

References

Aronofsky JS, Williams AC. The use of linear programming and mathematical models in underground oil production. Manage Sci. 1962; 8:394–407.

Aseeri, A., Gorman, P., Bagajewicz, M. J. Financial risk management in offshore oil infrastructure planning and scheduling. Ind. Eng. Chem. Res. 2004, 43, 3063–3072.

Babusiaux, D., Favennec, J. P., Bauquis, P. R., Bret-Rouzaut, N., and Guirauden, D. (2007). Oil and Gas Exploration and Production: Reserves, costs, contracts. Technip edition, 2007, chap. 4.

Bagajewicz, M. J. On the use of net present value in investment capacity planning models. Ind. Eng. Chem. Res., 2008, 47, 9413–9416.

Bailey, W. J.; Couet, B.; Wilkinson, D. Field Optimization Tool for Maximizing Asset Value. SPE ReserVoir EVal. Eng. 2005, 8 (1), 7–21.

Barnes, R., Linke, P., Kokossis, A. Optimization of oil-field development production capacity. European symposium on computer aided process engineering, 2002, vol. 12 (p. 631).

Barnes, R.J., Kokossis, A. and Shang, Z. An integrated mathematical programming approach for the design and optimisation of offshore fields. Computers and Chemical Engineering, 31, Issues 5–6, 2007, 612–629.

Begg, S. H., Bratvold, R., Campbell, J. C. Improving Investment Decisions Using A Stochastic Integrated Asset Model Presented at the SPE Annual Technical Conference and Exhibition, New Orleans, LA, Sep 30–Oct 3, 2001; Paper SPE 71414.

Behrenbruch, P. Offshore oilfield development planning. J. Pet. Technol. 1993, 45 (8), 735–743.

Birge, J. R., Louveaux, F., 1997. Introduction to stochastic programming. New York, NY: Springer.

Blake, A. J.; Roberts, M. C. Comparing petroleum fiscal regimes under oil price uncertainty. Resources Policy, 2006, 31(2), 95–105.

Bohannon J. A linear programming model for optimum development of multi-reservoir pipeline systems. J Petrol Tech. 1970;22:1429–1436.

BP Statistical Review of World Energy, 2011.

Caroe, C.C., Schultz, R., 1999: Dual decomposition in stochastic integer programming. Oper. Res. Lett. 24, 37–45.

Carvalho, M., Pinto, J. M. A bilevel decomposition technique for the optimal planning of offshore platforms. Brazilian J. Chem. Eng. 2006a, 23, 67–82.

Carvalho, M., Pinto, J. M. An MILP model and solution technique for the planning of infrastructure in offshore oilfields. J. Pet. Sci. Eng. 2006b, 51, 97–110.

Cullick, A. S., Heath D., Narayanan K., April J., Kelly J. Optimizing Multiple-Field Scheduling and Production Strategy with Reduced Risk. Presented at the SPE Annual Technical Conference and Exhibition, Denver, CO, Oct 5–8, 2003; Paper SPE 84239.

Cullick, A. S., Cude, R., and Tarman, M. 2007. Optimizing Field Development Concepts for Complex Offshore Production Systems, Paper SPE 108562-MS presented at Offshore Europe, Aberdeen, Scotland, U.K., 4–7 September.

Dias, M., 2002. Investment in information in petroleum, real options and revelation. In: Proceedings of the 6th Annual International Conference on Real Options. Real Options Group at Cyprus, Cyprus.

Elgsæter, S. M., O. Slupphaug, Johansen T. A. A structured approach to optimizing offshore oil and gas production with uncertain models. Comput. Chem. Eng. 2010, 34(2), 163–176.

Ettehad, A., Jablonowski, C. J., Lake, L. W. Stochastic Optimization and Uncertainty Analysis for E&P Projects: A Case in Offshore Gas Field Development, Offshore Technology Conference, 2011.

Fisher, M. L., 1985. An Applications Oriented Guide to Lagrangian Relaxation. Interfaces 15, 10–21.

Frair, L. C. Economic Optimization of Offshore Oilfield Development. PhD Dissertation, University of Oklahoma, Tulsa, OK, 1973.

Goel, V., Grossmann, I. E. A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves. Comput. Chem. Eng. 2004, 28 (8), 1409–1429.

Goel, V., Grossmann, I. E. A class of stochastic programs with decision dependent uncertainty. Mathematical Programming, 2006, 108(2–3, Ser. B), 355–394.

Goel, V., Grossmann, I. E., El-Bakry, A. S., Mulkay, E. L. A novel branch and bound algorithm for optimal development of gas fields under uncertainty in reserves. Comput. Chem. Eng. 2006, 30, 1076–1092.

Grothey, A., McKinnon, K., 2000. Decomposing the optimization of a gas lifted oil well network. Tech Rept. MS 00-005, Department of Mathematics and Statistics, University of Edinburgh.

Gunnerud, V., Foss, B. Oil production optimization – A piecewise linear model, solved with two decomposition strategies. Computers and Chemical Engineering, 2010, 34, 1803–1812.

Gupta, V., Grossmann, I. E., 2011. Solution Strategies for Multistage Stochastic Programming with Endogenous Uncertainties. Computers and Chemical Engineering 35, 2235–2247.

Gupta, V., Grossmann, I. E., 2012a. An Efficient Multiperiod MINLP Model for Optimal planning of Offshore Oil and Gas Field Infrastructure. Industrial and Engineering Chemistry Research 51 (19), 6823–6840.

Gupta, V., Grossmann, I. E., 2012b. Modeling and Computational Strategies for Offshore Oilfield Development Planning under Complex Fiscal Rules. Industrial and Engineering Chemistry Research 51, 14438–14460.

Gupta, V., Grossmann, I. E., 2014a. Multistage Stochastic Programming Approach for Offshore Oilfield Infrastructure Planning under Production Sharing Agreements and Endogenous Uncertainties, Journal of Petroleum Science and Engineering 124, 180–197.

Gupta, V., Grossmann, I. E., 2014b. A new Decomposition Algorithm for Multistage Stochastic Programs with Endogenous Uncertainties, Computers and Chemical Engineering 62, 62–79.

Haugen, K. K., 1996. A stochastic dynamic programming model for scheduling of offshore petroleum fields with resource uncertainty. European Journal of Operational Research 88 (1), 88–100.

Haugland, D., Hallefjord, Å.; Asheim, H. Models for Petroleum Field Exploitation. Eur. J. Oper. Res. 1988, 37, 58.

Iyer, R. R., Grossmann, I. E., Vasantharajan, S., Cullick, A. S. Optimal planning and scheduling offshore oilfield infrastructure investment and operations. Ind. Eng. Chem. Res. 1998, 37, 1380–1397.

Johnston, D. Petroleum Fiscal Systems and Production Sharing Contracts; PennWell Publishing Co.: Tulsa, Oklahoma, 1994.

Jonsbraten, T. W. Oil-field optimization under price uncertainty. Journal of the Operational Research Society, 1998a, 49, 811.

Jonsbraten, T., 1998b. Optimization models for petroleum field exploitation. PhD thesis, Norwegian School of Economics and Business Administration.

Jonsbraten, T. W.; Wets, R. J. B.; Woodruff, D. L. A class of stochastic programs with decision dependent random elements. Ann. Oper. Res. 1998, 82, 83–106.

Kaiser, M.J. and A.G. Pulsipher. Fiscal System Analysis: Concessionary and Contractual Systems Used in Offshore Petroleum Arrangements. US Department of the Interior Mineral Management Services MMS 2004-016, 1–78.

Kosmidis, V., Perkins, J., Pistikopoulos, E., 2002. A mixed integer optimization strategy for integrated gas/oil production. In: European Symposium on Computer Aided Process Engineering. Vol. 12.

Kosmidis, V. D., Perkins, J. D., Pistikopoulos, E. N. A mixed integer optimization for the well scheduling problem on petroleum fields. Computers & Chemical Engineering Journal, 2005, 29, 1523–1541.

Lainez, J. M.; Puigjaner, L.; Reklaitis, G. V. Financial and financial engineering considerations in supply chain and product development pipeline management. Comput. Chem. Eng., 2009, 33, 1999–2011.

Lee AS, Aronofsky JS. A linear programming model for scheduling crude oil production. J Petrol Tech. 1958;10:51–54.

Lin, X., Floudas, C. A. A Novel Continuous-Time Modeling and Optimization Framework for Well Platform Planning Problems. Optim. Eng. 2003, 4 (1–2), 65–95.

Lund, M. W. Valuing Flexibility in Offshore Petroleum Projects. Ann. Oper. Res. 2000, 99 (1–4), 325–349.

Meister, B., Clark, J. M. C., Shah, N., 1996. Optimisation of oilfield exploitation under uncertainty. Computers and Chemical Engineering 20, S1251–S1256, suppl. B.

Mouret, S., Grossmann, I. E., Pestiaux, P., 2011. A new Lagrangean decomposition approach applied to the integration of refinery planning and crude-oil scheduling. Computers and Chemical Engineering 35, 2750–2766.

Oliveira, F., Gupta, V., Hamacher, S., Grossmann, I. E., 2013. A Lagrangean Decomposition Approach for Oil Supply Chain Investment Planning under Uncertainty with Risk Considerations. Computers and Chemical Engineering 50, 184–195.

Ortiz-Gomez, A., Rico-Ramirez, V., & Hernandez-Castro, S. Mixed-integer multiperiod model for the planning of oil-field production. Comput. Chem. Eng., 2002, 26(4–5), 703.

Ruszczynski, A., 1997. Decomposition methods in stochastic programming. Math. Programming (Ser. B) 79, 333–353.

Stensland, G., Tjøstheim, D., 1991. Optimal decisions with reduction of uncertainty over time - an application to oil production. In: Lund, D., øksendal, B. (Eds.), Stochastic Models and Option Values. pp. 267–291.

Sullivan J. A computer model for planning the development of an offshore gas field. J Petrol Tech. 1982;34:1555–1564.

Sunley, E. M., T. Baunsgaard, and D. Simard. 2002. "Revenue from the Oil and Gas Sector: Issues and Country Experience." IMF Conference Paper.

Tarhan, B., Grossmann, I. E., 2008. A multistage stochastic programming approach with strategies for uncertainty reduction in the synthesis of process networks with uncertain yields. Computers and Chemical Engineering 32, 766–788.

Tarhan, B., Grossmann, I.E., Goel, V. Stochastic programming approach for the planning of offshore oil or gas field infrastructure under decision- dependent uncertainty. Ind. Eng. Chem. Res. 2009, 48(6), 3078–3097.

Tarhan, B., Grossmann, I.E., Goel, V. Computational strategies for non-convex multistage MINLP models with decision-dependent uncertainty and gradual uncertainty resolution. Ann. Oper. Res., 2011, DOI: 10.1007/s10479-011-0855-x.

Tarhan, B., Gupta V., Grossmann, I. E., 2013. Improving Dual Bound for Stochastic MILP Models using Sensitivity Analysis, manuscript in preparation.

Tordo, S. 2007. Fiscal Systems of Hydrocarbons: Design Issues. Working Paper 123/07, World Bank, Washington, D.C.

Tsarbopoulou, C. Optimization of oil facilities and oil production. M.Sc. Dissertation, University College London, London, U.K., 2000.

Ulstein, N. L., Nygreen, B.; Sagli, J. R. Tactical planning of offshore petroleum production. Eur. J. Oper. Res. 2007, 176, 550–564.

van den Heever, S. A.; Grossmann, I. E. An iterative aggregation/disaggregation approach for the solution of a mixed integer nonlinear oilfield infrastructure planning model. Ind. Eng. Chem. Res. 2000, 39, 1955–1971.

van den Heever, S. A., Grossmann, I. E., Vasantharajan, S.; Edwards, K. Integrating complex economic objectives with the design and planning of offshore oilfield infrastructures. Comput. Chem. Eng. 2000, 24, 1049–1055.

van den Heever, S. A., Grossmann, I. E. A Lagrangean Decomposition Heuristic for the Design and Planning of Offshore Hydrocarbon Field Infrastructures with Complex Economic Objectives. Ind. Eng. Chem. Res. 2001, 40, 2857–2875.

World Bank. 2007. "Contracts for Petroleum Development – Part 1–3." Petroleum Sector Briefing Note No. 7.

Zabalza-Mezghani, I.; Manceau, E.; Feraille, M.; Jourdan, A. Uncertainty management: From geological scenarios to production scheme optimization. J. Pet. Sci. Eng. 2004, 44, 11–25.

Emerging Optimal Control Models and Solvers for Interconnected Natural Gas and Electricity Networks

Nai-Yuan Chiang and Victor M. Zavala

Abstract This chapter reviews emerging optimal control models for interconnected natural gas and electricity networks and discusses economic drivers motivating the development of such models. We also review computational patterns and structures arising in these models and assess the potential and limitations of state-of-the-art optimization solvers.

1 Motivation

Natural gas and power grid infrastructures are becoming increasingly interdependent. A major factor driving this situation is the increasing deployment of gas-fired power plants. These plants are modular and less capital intensive compared with large, centralized generation facilities running on nuclear and coal fuel sources. In addition, gas-fired plants are more flexible and can quickly ramp up and down their power output. This flexibility becomes an asset as the share of intermittent solar and wind power increases. Moreover, the high availability of gas resulting from new fracking technologies that extract gas from shales has led to lower prices, making gas-fired plants economically more attractive (Cafaro and Grossmann 2014).

N.-Y. Chiang (✉)
Mathematics and Computer Science Division, Argonne National Laboratory,
9700 South Cass Avenue, Argonne IL 60439, USA
e-mail: nychiang@mcs.anl.gov

V.M. Zavala
Department of Chemical and Biological Engineering,
University of Wisconsin-Madison, 1415 Engineering Drive,
Madison, WI 53706, USA
e-mail: victor.zavala@wisc.edu

M. Martín (ed.), *Alternative Energy Sources and Technologies*,
DOI 10.1007/978-3-319-28752-2_4

An important feature of gas-fired generation (compared with other generation technologies) is that large amounts of fuel must be transported to power generation facilities in gaseous form through a sophisticated network that spans thousands of miles. A key advantage of having such a network infrastructure is that significant amounts of gas can be stored inside the pipelines. The stored gas is distributed spatially in the network and is normally referred to as *line-pack* (Carter et al. 2004). Line-pack is used by pipeline operators to modulate variations of gas demands at multiple spatial points in intraday operations. Some of the strongest variations in gas demands are the result of on-demand start-up and shutdown of gas-fired power plants (Rachford et al. 2009). Modulating these variations is challenging because the fast release of line-pack at multiple simultaneous locations can trigger complex spatiotemporal responses that propagate hundreds to thousands of miles and that can take hours to stabilize. Therefore, line-pack management is performed by using sophisticated optimal control and pipeline simulation tools. These automation tools orchestrate the operation of a multitude of compressor stations distributed throughout the system with the objectives of satisfying demands, maintaining pressure levels, and minimizing compression costs (Rachford and Carter 2004).

An important issue faced by gas-fired power plants is that they compete for natural gas with industrial facilities and with local distribution companies (LDCs) that supply gas to urban areas. Therefore, natural gas cannot be guaranteed to be available at each power generation facility at all times. This limitation is particularly evident during the winter season when residential and office buildings require large amounts of gas for heating. An extreme manifestation of this issue was observed during the polar vortex of 2014 in which sustained low temperatures in the Midwest region of the U.S. led to high gas demands in urban areas and to equipment failures.[1,2] These factors resulted in widespread shortages of natural gas in places as remote as California, Massachusetts, and Texas. These gas shortages in turn resulted in lost electrical generation capacity totaling 35 GW. At a value of lost load of 5,000 $/MWh, shortages of this magnitude represent economic losses of 175 million $/hr. The New England area alone lost 1.5 GW of power generation capacity (FERC 2014). The polar vortex also exposed market inefficiencies resulting from the increasing interaction between grid and gas systems. In particular, gas-fired plants required significant uplift payments from the independent system operators (ISOs). These payments compensated the power plants for the lost revenue resulting from the inability of the gas infrastructure to deliver fuel. These operational and economic issues question the ability of the gas network infrastructure to sustain additional gas-fired generation. Consequently, it is important to investigate the economic and flexibility of natural gas infrastructures by developing detailed physical models and advanced optimal control tools.

[1] http://tinyurl.com/pvouhym.

[2] http://tinyurl.com/no5v2za.

Optimal control formulations of natural gas networks using high-resolution dynamic models have been reported by several researchers (Steinbach 2007; Carter et al. 2004; Zavala 2014). These studies, however, do not consider coordination with power grid networks and treat power plants as exogenous uncertain disturbances. By coupling gas and electric models one can better understand the enhanced flexibility and resiliency that can be gained by coordination. Researchers have also developed models of interconnected power grid and natural gas networks but the dynamics of natural gas systems are often neglected (Geidl and Andersson 2007; Arnold et al. 2010; An et al. 2003). Steady-state models cannot capture line-pack storage dynamics and thus significantly misrepresent the flexibility of the system in real-time operations. Consequently, steady-state models are more appropriate for long-term planning studies (Ríos-Mercado and Borraz-Sánchez 2015). In this respect, the models that we discuss in this chapter can better capture the flexibility provided by line-pack in real-time operations. Recent studies have also reported models and strategies for co-optimization of gas and power grid transmission systems using detailed dynamic gas models. The studies in (Liu et al. 2009, 2011) use full-resolution models but focus on small synthetic models to assess economic improvements resulting from coordination and to evaluate the impact of using dynamic over steady-state gas pipeline models. The studies in (Chaudry et al. 2008; Qadrdan et al. 2010) focus on the Great Britain network and provide more in-depth analyses. In particular, the study in (Chaudry et al. 2008) presents a multi-time period model to study the effects of gas terminal (supply) failures on the integrated gas–electric system. The model, however, uses simplifications to address computational complexity; in particular, an aggregated Great Britain model with 16 buses is used and the gas system dynamics are only captured at the daily timescale (dynamics in intraday operations are ignored). The model proposed in (Qadrdan et al. 2010) studies the effect of wind power adoption levels on gas generation and demonstrates that line-pack can limit system performance during periods with low wind generation. The simplified 16-bus Great Britain network is also used in this study, the gas network is simplified by aggregating parallel branches, and gas dynamics are ignored. None of the studies reported in the literature discuss computational efficiency and scalability issues. In this chapter we review modern optimal control formulations of natural gas networks and discuss the economic drivers motivating such applications. We also provide a review of existing computational capabilities that enable the solution of these complex optimization problems and provide pointers to open challenges. Our objective is to focus on real-sized networks and to capture full physical and spatiotemporal resolutions. This can enable us to discover nonintuitive physical behavior that can inform ISOs and gas pipeline operators. Focusing on real-sized systems will also enable us to assess the limits of state-of-the-art optimization algorithms and to identify scalability bottlenecks.

This chapter is structured as follows: Sect. 2 presents elements of a typical optimal control model for natural gas networks and discusses interconnection with

power grid models and derivation of stochastic formulations. Section 3 discusses economic and resiliency insights that can be gained with detailed optimal control models and Sect. 4 discusses computational solution strategies and open challenges. Concluding remarks are provided in Sect. 5.

2 Optimal Control Formulations

In this section we present typical components of an optimal control formulation for gas networks and provide pointers to suitable extensions.

2.1 Transport Equations

We use the following isothermal form of the continuity and momentum equations:

$$\frac{\partial p_\ell(x,\tau)}{\partial \tau} + \frac{\mathrm{ZRT}}{A_\ell}\frac{\partial f_\ell(x,\tau)}{\partial x} = 0 \tag{1a}$$

$$\frac{1}{A_\ell}\frac{\partial f_\ell(x,\tau)}{\partial \tau} + \frac{\partial p_\ell(x,\tau)}{\partial x} + \frac{8\lambda_\ell}{\pi^2 D_\ell^5}\frac{f_\ell(x,\tau)|f_\ell(x,\tau)|}{\rho_\ell(x,\tau)} = 0 \tag{1b}$$

Here, $\tau \in [0,N] := \mathscr{T}$ is the time dimension with final time N and $x \in \mathscr{X}_\ell := [0,L_\ell]$ is the axial dimension with length L_ℓ. The link diameters are denoted as D_ℓ and the friction coefficients as λ_ℓ. The states of the link are the gas density $\rho_\ell(x,\tau)$, speed $v_\ell(x,\tau)$, and pressure $p_\ell(x,\tau)$. The gas pressure and density are related as

$$\frac{p_\ell(x,\tau)}{\rho_\ell(x,\tau)} = \mathrm{ZRT}, \tag{2}$$

where R is the universal gas constant and T is temperature. For a detailed non-isothermal formulation the reader can refer to (Abbaspour and Chapman 2008). The pipeline links are connected through a network comprising a set $\mathscr{N}$ of nodes, a set $\mathscr{S}$ of supply flows, and a set $\mathscr{D}$ of demand flows. For each node $n \in \mathscr{N}$ we define the set of inlet and outlet links, $\mathscr{L}_n^{\mathrm{snd}} := \{\ell \,|\, \mathrm{snd}(\ell) = n\}$, $\mathscr{L}_n^{\mathrm{rec}} := \{\ell \,|\, \mathrm{rec}(\ell) = n\}$. Here, $\mathrm{snd}(\ell) \in \mathscr{N}$ is the start node of link ℓ and $\mathrm{rec}(\ell) \in \mathscr{N}$ is the end node. We assume that the direction of the flow is given. For formulations that allow for flow reversals the reader is referred to (Baumrucker and Biegler 2010). We define $\mathrm{dn}(j) \in \mathscr{N}$ as the node at which the demand flow $d_j(\tau)$ is located and $\mathrm{sn}(i) \in \mathscr{N}$ as the node at which the supply flow $s_i(\tau)$ is located. Accordingly, we define the sets $\mathscr{S}_n := \{j \in \mathscr{S} \,|\, \mathrm{sn}(j) = n\}$ and $\mathscr{D}_n := \{j \in \mathscr{D} \,|\, \mathrm{dn}(j) = n\}$ for each node $n \in \mathscr{N}$. For each node $n \in \mathscr{N}$ we also define pressures $\boldsymbol{\theta}_n(\cdot)$.

For modeling convenience we lift the network system by introducing *dummy* inlet flows for each link $f_\ell^{\text{in}}(\cdot)$, respectively; and outlets $f_\ell^{\text{out}}(\cdot)$. Using these definitions we can express mass balances at the nodes as

$$\sum_{\ell \in \mathscr{L}_n^{\text{rec}}} f_\ell^{\text{out}}(\tau) - \sum_{\ell \in \mathscr{L}_n^{\text{snd}}} f_\ell^{\text{in}}(\tau) + \sum_{i \in \mathscr{S}_n} s_i(\tau) - \sum_{j \in \mathscr{D}_n} d_j(\tau) = 0, \quad n \in \mathscr{N} \tag{3}$$

We split the set of links $\mathscr{L}$ into subsets of passive $\mathscr{L}_p$ links and active links $\mathscr{L}_a$. For the active links we define the *boost pressures* $\Delta\theta_\ell(\cdot)$ which are the additional (nonnegative) pressures introduced by the compressor located at the inlet (sending) node of the link. For the passive links, there is no compression. The boundary conditions for the link pressures are given by

$$p_\ell(L_\ell, \tau) = \theta_{\text{rec}(\ell)}(\tau), \quad \ell \in \mathscr{L} \tag{4a}$$

$$p_\ell(0, \tau) = \theta_\ell^{\text{dis}}(\tau), \quad \ell \in \mathscr{L}. \tag{4b}$$

The boundary conditions for the link flows are,

$$f_\ell(0, \tau) = f_\ell^{\text{in}}(\tau), \quad \ell \in \mathscr{L} \tag{5a}$$

$$f_\ell(L_\ell, \tau) = f_\ell^{\text{out}}(\tau), \quad \ell \in \mathscr{L}. \tag{5b}$$

The discharge pressures of the compressors (located at the inlet of the active links) are given by,

$$\theta_\ell^{\text{dis}}(\tau) = \theta_{\text{snd}(\ell)}(\tau) + \Delta\theta_\ell(\tau), \quad \ell \in \mathscr{L}_a. \tag{6a}$$

For the passive links we simply have that,

$$\theta_\ell^{\text{dis}}(\tau) = \theta_{\text{snd}(\ell)}(\tau), \quad \ell \in \mathscr{L}_p. \tag{7}$$

The total compression power consumed in the active links is given by

$$P_\ell(\tau) = c_p f_\ell^{\text{in}}(\tau) T \left(\left(\frac{\theta_\ell^{\text{dis}}(\tau)}{\theta_{\text{rec}(\ell)}(\tau)} \right)^\beta - 1 \right), \quad \ell \in \mathscr{L}_a. \tag{8}$$

We note that this constraint implies that the compressor ratios are variable and not fixed. As can be seen, the transport equations are highly nonlinear. Additional complexity can be added to the model by consider control valves and more complex compressor configurations (Steinbach 2007; Marqués and Morari 1988).

2.2 Constraints

We consider the following constraints on available compressor power, suction and discharge pressures, and demand delivery pressures,

$$P_\ell^L \le P_\ell(\tau) \le P_\ell^U, \quad \ell \in \mathscr{L}_a \tag{9a}$$

$$\theta_\ell^{\mathrm{suc},L} \le \theta_{\mathrm{snd}(\ell)}(\tau) \le \theta_\ell^{\mathrm{suc},U}, \quad \ell \in \mathscr{L}_a \tag{9b}$$

$$\theta_\ell^{\mathrm{dis},L} \le \theta_\ell^{\mathrm{dis}}(\tau) \le \theta_\ell^{\mathrm{dis},U}, \quad \ell \in \mathscr{L}_a \tag{9c}$$

$$\theta_j^{\mathrm{dem},L} \le \theta_{\mathrm{dn}(j)}(\tau) \le \theta_j^{\mathrm{dem},U}, \quad j \in \mathscr{D}. \tag{9d}$$

We also have the implicit physical bounds $f_\ell^{\mathrm{in}}(\cdot), f_\ell^{\mathrm{out}}(\cdot), P_\ell(\cdot) \ge 0$ and we assume that the pressures (or inlet flows but not both) at the supply points are given:

$$\theta_{\mathrm{sn}(j)}(\tau) = \text{given} \quad \text{or} \quad s_j(\tau) = \text{given}, \quad j \in \mathscr{S}. \tag{10}$$

The gas demands are bounded by the given targets $d_j^{\mathrm{target}}(\cdot)$,

$$0 \le d_j(\tau) \le d_j^{\mathrm{target}}(\tau), \quad j \in \mathscr{D}. \tag{11}$$

We enforce periodicity of the line-pack in each pipeline using the constraint:

$$\int_0^{L_\ell} f_\ell(x,T)\mathrm{d}x \ge \int_0^{L_\ell} f_\ell(x,0)\mathrm{d}x, \quad \ell \in \mathscr{L}_a. \tag{12}$$

Without this constraint, the system will tend to deplete line-pack in order to minimize compressor power and this can make operation in the following day infeasible. The periodicity constraint thus provides a mechanism to deal with the finite horizon of the optimal control problem (OCP) and ensure recursive feasibility (Rawlings et al. 2012).

2.3 Initial State

We assume that the system is at steady state at the initial time $\tau = 0$. Such steady-state satisfies the transport equations

$$\frac{\mathrm{ZRT}}{A_\ell}\frac{\partial f_\ell(x,\tau)}{\partial x} = 0 \tag{13a}$$

$$\frac{\partial p_\ell(x,\tau)}{\partial x} + \frac{8\lambda_\ell}{\pi^2 D_\ell^5}\frac{f_\ell(x,\tau)|f_\ell(x,\tau)|}{\rho_\ell(x,\tau)} = 0 \tag{13b}$$

One can use the boundary conditions defined at $\tau = 0$ and the assumption of given supply pressures (or flows) at $\tau = 0$ to prove that fixing boost pressures $\Delta\theta_\ell(0)$ and demand flows $d_j(0)$ fully defines the initial states (Zavala 2014).

2.4 Objective Function

The cost function can be set as a combination of gas supply, compression cost, and gas demand delivery,

$$\varphi_{\mathrm{gas}} := \int_0^N \left(\sum_{i\in\mathscr{S}} \alpha_i^s s_i(\tau) + \sum_{\ell\in\mathscr{L}_a} \alpha_\ell^P P_\ell(\tau) - \sum_{j\in\mathscr{D}} \alpha_j^d d_j(\tau) \right) \mathrm{d}\tau \tag{14}$$

The first term is the supply cost, the second term in the cost function is the compressor power with cost α_ℓ^P, and the third term is the total value of served demand. The demand term seeks to maximize the served demand (this explains the negative sign) and the parameter α_j^d can be interpreted as the value of the served demand. Consequently, the cost function is an analog of the negative social welfare used in electricity market clearing formulations (Pritchard et al. 2010). We explore this topic in Sect. 2.5. From a control standpoint, the demand term can also be interpreted as penalty for unserved demand. In particular, when the value α_j^d is high relative to compression power and supply, the control system will tend to push the demands $d_j(\cdot)$ to the targets $d_j^{\mathrm{target}}(\cdot)$ defined in the constraints (11). The parameter α_j^d can also be used to *prioritize* certain demand locations (e.g., residential over power plants) and can be used to model gas bids provided by gas-fired power plants. We highlight that the cost function (14) is linear and this tends to introduce ill-conditioning in the Hessian matrix and slowdown convergence. This is an issue arising in optimal control models with economic objectives and is exacerbated in distributed systems (Zavala and Biegler 2009). To ameliorate this issue, it is also possible to define a cost function of the form

$$\varphi_{\mathrm{gas}} := \int_0^N \left(\sum_{i\in\mathcal{S}} \alpha_i^s s_i(\tau) + \sum_{\ell\in\mathcal{L}_a} \alpha_\ell^P P_\ell(\tau) + \sum_{j\in\mathcal{D}} \alpha_j^d (d_j(\tau) - d_j^{\mathrm{target}}(\tau))^2 \right) \mathrm{d}\tau. \tag{15}$$

Here, the quadratic term in the objective is a tracking term that seeks to bring the demands to the targets. An advantage of this formulation is that it adds positive curvature to the cost function and this can aid computational performance. A disadvantage of this formulation is that the parameter α_j^d does not necessarily have an economic interpretation and, consequently, it might be difficult to tune.

The optimal control model for the gas side can be summarized as:

$$\begin{aligned} \min \quad & \varphi_{\text{gas}} \\ \text{s.t.} \quad & (1)-(13). \end{aligned}$$

One can use a structural model analysis to prove that the number of *degrees of freedom* of the problem corresponds to that of the control trajectories $\Delta\theta_\ell(\tau),\ \ell \in \mathscr{L}_a, \tau \in \mathscr{T}$ and $d_j(\tau),\ j \in \mathscr{D}, \tau \in \mathscr{T}$ (Zavala 2014).

2.5 Integrated Gas–Electric Formulations

A major source of uncertainty in natural gas systems are the large gas demands of power plants (Rachford et al. 2009). At the same time, gas-fired power plants cannot guarantee that the entire amount of gas requested can be delivered at all times and this uncertainty affects ISO operations. Consequently, it is natural to believe that better communication/coordination between natural gas and power grid infrastructures can significantly help mitigate uncertainty on both sides. Moreover, we can expect that by combining control flexibility through coordination we can increase overall resiliency. The question is, however, by how much? To answer this question it is necessary to develop integrated gas–electric optimal control models.

Economic dispatch is an OCP that is solved by ISOs to balance supply and demand and to price electricity in intraday operations. We formulate the dispatch problem as the continuous-time OCP shown in (16). For a more detailed explanation of the model the reader is referred to (Chiang and Zavala 2015). Here, we use a notation similar to that used in the description of the gas system to highlight similarities and differences between gas and electric models. We highlight differences in notation when appropriate. We define the sets of electricity suppliers (i.e., power plants or generators) as $\mathscr{S}$, the set of electrical loads as $\mathscr{D}$, the set of network nodes as $\mathscr{N}$, and the set of transmission lines as $\mathscr{L}$. For each link $\ell \in \mathscr{L}$ we denote $\text{snd}(\ell) \in \mathscr{N}$ as the sending node and $\text{rec}(\ell) \in \mathscr{N}$ as the receiving node. We define $\mathscr{S}_g \subseteq \mathscr{S}$ as the subset of power plants that are fired by natural gas. The rest of the suppliers are either thermal or renewable power suppliers. We define $\mathscr{S}_n \subseteq \mathscr{S}$ as the subset of suppliers connected to node n and $\mathscr{D}_n \subseteq D$ as the subset of loads connected to node n. We define $\mathscr{L}_n^{\text{snd}} \subseteq \mathscr{L}$ as the set of links originating from node n and $\mathscr{L}_n^{\text{rec}} \subseteq \mathscr{L}$ as the set of lines ending in node n.

$$\min \ \varphi_{\text{grid}} := \int_0^N \left(\sum_{i \in \mathscr{S}} \alpha_i^s s_i(\tau) - \sum_{j \in \mathscr{D}} \alpha_j^d d_j^{\text{grid}}(\tau) \right) \mathrm{d}\tau \tag{16a}$$

$$\text{s.t.} \ \frac{\mathrm{d}s_i(\tau)}{\mathrm{d}\tau} = r_i(\tau), \quad \in \mathscr{S} \tag{16b}$$

$$\sum_{\ell \in \mathscr{L}_n^{\text{rec}}} f_\ell(\tau) - \sum_{\ell \in \mathscr{L}_n^{\text{snd}}} f_\ell(\tau) + \sum_{i \in \mathscr{S}_n} s_i(\tau) - \sum_{j \in \mathscr{D}_n^{\text{grid}}} d_j^{\text{grid}}(\tau) = 0, \quad n \in \mathscr{N} \tag{16c}$$

$$f_\ell(\tau) = \beta_\ell \big(\theta_{\text{snd}(\ell)}(\tau) - \theta_{\text{rec}(\ell)}(\tau) \big), \quad \ell \in \mathscr{L} \tag{16d}$$

$$\underline{f}_\ell \leq f_\ell(\tau) \leq \overline{f}_\ell, \quad \ell \in \mathscr{L} \tag{16e}$$

$$\underline{\theta}_n \leq \theta_n(\tau) \leq \overline{\theta}_n, \quad n \in \mathscr{N} \tag{16f}$$

$$\underline{s}_i \leq s_i(\tau) \leq \overline{s}_i, \quad i \in \mathscr{S} \tag{16g}$$

$$\underline{r}_i \leq r_i(\tau) \leq \overline{r}_i, \quad i \in \mathscr{S} \tag{16h}$$

$$0 \leq d_j^{\text{grid}}(\tau) \leq d_j^{\text{grid,target}}(\tau), \quad i \in \mathscr{S} \tag{16i}$$

$$d_i^{\text{gas,grid}}(\tau) = \eta_i \cdot s_i(\tau), \quad i \in \mathscr{S}_g. \tag{16j}$$

The time profiles for power generation, delivered loads, flows, and voltage angles are denoted as $s_i(\cdot)$, $i \in \mathscr{S}$; $d_j^{\text{grid}}(\cdot)$, $j \in \mathscr{D}^{\text{grid}}$; $f_\ell(\cdot)$, $\ell \in \mathscr{L}$; and θ_n, $n \in \mathscr{N}$, respectively. The time profiles for target loads are denoted as $d_j^{\text{grid,target}}(\cdot)$. The power flows are bounded by the capacity limits $\underline{f}_\ell, \overline{f}_\ell$, the angles are bounded by $\underline{\theta}_n, \overline{\theta}_n$, and the supply flows are bounded by $\underline{s}_i, \overline{s}_i$. The supply change rates are given by $r_i(\cdot)$, $i \in \mathscr{S}$ and are bounded by the ramp limits $\underline{r}_i, \overline{r}_i$.

The cost function (16a) is the negative social welfare (Pritchard et al. 2010). Where the term "negative" indicates a sign reversal in the standard social welfare (i.e., we minimize suppliers costs minus consumer costs rather than maximize consumer costs minus supplier costs). The generation and demand costs are α_i^s, $i \in \mathscr{S}$ and α_j^d, $j \in \mathscr{D}^{\text{grid}}$, respectively. In the case of inelastic demands the demand costs are typically set to the value of lost load (VOLL). This value is typically in the range 1,000–10,000 $/MWh (Centolella and Ott 2009). This parameter has similar interpretation to the one used in the gas side.

The dual variables of the network balance Eq. (16c) are the locational marginal prices, which we denote as $\pi_n(\cdot)$, $n \in \mathscr{N}$. Equations (16d) are the DC power flow relationships, Eqs. (16f)–(16h) are bounds for voltage angles, generation, and ramps, respectively.

We define $d_i^{\text{gas,grid}}(\cdot)$, $i \in \mathscr{S}_g$ as the gas demands originating from the gas-fired plants and η_i, $i \in \mathscr{S}_g$ as the heat rates of the different power plants. The *heat rate* is a measure of the conversion efficiency of a power plant installation and is defined as the amount of fuel (in BTUs) needed to produce a KWh of electrical energy. The ideal heat rate is 3,412 BTU/KWh, which indicates a one-to-one conversion between fuel and electrical energy (efficiency of 100 %). Consequently, the smaller the heat rate, the more efficient the technology. Different gas-fired plants can have different heat rates (combined cycle plants have much smaller heat rates than do simple cycle plants). In fact, the energy information administration (EIA) reports that the average heat rate for gas-fired plants has been reducing steadily from 9,207 BTU/KWh in 2003 (efficiency of 37 %) to 7,948 BTU/KWh in 2013 (efficiency of 42.9 %) because of the adoption of new technologies. In comparison, the average heat rate for coal power plants was reported to be 10,459 BTU/KWh (efficiency of 32 %) in 2013 and this has remained at similar levels since 2003.[3]

The coupling between gas and power grid infrastructures is given by the following constraints:

$$d_j^{\text{target}}(\tau) = d_{\text{gn}(j)}^{\text{gas,grid}}(\tau) + d_j^{\text{base}}(\tau), \quad j \in \mathscr{D} \tag{17a}$$

$$d_{\text{gn}(j)}^{\text{gas,grid}}(\tau) \leq d_j(\tau) - d_j^{\text{base}}(\tau), \quad j \in \mathscr{D}. \tag{17b}$$

The first constraint states that the gas demand targets for the gas infrastructure are given by the gas demands of the gas-fired power plants plus an exogenous base gas demand that arises from industrial facilities and/or LDCs serving urban areas. Here, $\text{gn}(j) \in \mathscr{S}_g$ denotes the power plant corresponding to the gas demand $j \in \mathscr{D}$. The second constraint states that the gas used by the power plants cannot physically exceed the gas delivered by the gas infrastructure. We use this second constraint to model situations in which the gas infrastructure is physically constrained and thus the delivered demand $d_j^{\text{gas}}(\cdot)$ cannot match the target demand $d_{\text{gn}(j)}^{\text{gas,target}}(\cdot)$. The simultaneous solution of the grid and gas models together with the coupling constraints (17a) gives the *coordinated* dispatch model.

To compare the performance of a coordinated gas–electric system, we consider an *uncoordinated* setting in which the two infrastructures do not dispatch jointly (as is currently done). This is simulated by first solving the economic dispatch problem for the power grid (16) to set the *predicted* generation $s_i(\cdot)$, $i \in \mathscr{S}$, the natural gas demand targets $d_{\text{gn}(j)}^{\text{gas,grid}}(\cdot)$, $j \in \mathscr{D}_g$, and the predicted locational marginal prices $\pi_n(\cdot)$, $n \in \mathscr{N}$. Having the gas demand targets, we solve the gas dispatch problem to maximize the gas delivered and minimize compression costs. The solution of this problem sets the *realized* gas demands delivered to the gas-fired power plants $d_j^{\text{gas}}(\cdot) - d_j^{\text{gas,base}}(\cdot)$. Because the realized gas demands might not be able to match

[3] http://www.eia.gov/electricity/annual/html/epa_08_01.html.

the power grid targets, we solve the economic dispatch problem for the power grid (16a) again to determine the realized generation schedule and locational marginal prices corresponding given the realized delivered gas demands. We denote the realized power generation schedules as $s_i^{\text{real}}(\cdot)$ and the prices as $\pi_n^{\text{real}}(\cdot)$. Differences between the target and delivered gas demands will introduce a difference between the predicted and realized generation schedules and prices. When the gas-fired power plants cannot obtain the total gas requested, they will need to curtail power and must pay for the unserved electricity generation at the realized price (Ott 2003). In such a case, the revenue for the power plants is given by

$$\mathscr{R}_i := \int_0^N \Big(\pi_{\text{sn}(i)} s_i(\tau) + s_i^{\text{real}}(\tau)\big(\pi_{\text{sn}(i)}^{\text{real}}(\tau) - \pi_{\text{sn}(i)}(\tau)\big) - \alpha_i^s s_i^{\text{real}}(\tau)\Big)\mathrm{d}\tau, \quad i \in \mathscr{S}_g. \tag{18}$$

Here, $\text{sn}(i) \in \mathscr{N}$ denotes the node at which supplier $i \in \mathscr{S}_g$ is connected to. The total revenue for the gas-fired generators is denoted by $\mathscr{R} = \sum_{i \in \mathscr{S}_g} \mathscr{R}_i$.

When the requested and delivered gas demands coincide we have that the predicted and realized generation schedules coincide (i.e., the predicted generation schedule is feasible to the gas system). Consequently, we have that $\pi_{\text{sn}(i)}^{\text{real}}(\cdot) = \pi_{\text{sn}(i)}(\cdot)$ and the revenue reduces to

$$\mathscr{R}_i = \int_0^N \big(\pi_{\text{sn}(i)}(\tau) s_i(\tau) - \alpha_i^s s_i^{\text{real}}(\tau)\big)\mathrm{d}\tau, \quad i \in \mathscr{S}_g. \tag{19}$$

After space–time discretization (for a review on discretization techniques the reader is referred to (Abbaspour and Chapman 2008; Biegler 2010)), we can represent the coupled gas–electric problem as the nonlinear program (NLP):

$$\min\ \varphi_{\text{grid}}(w_{\text{grid}}) + \varphi_{\text{gas}}(w_{\text{gas}}) \tag{20a}$$

$$\text{s.t. } c_{\text{grid}}(w_{\text{grid}}) \geq 0, \quad (\lambda_{\text{grid}}) \tag{20b}$$

$$c_{\text{gas}}(w_{\text{gas}}) \geq 0, \quad (\lambda_{\text{gas}}) \tag{20c}$$

$$\Pi_{\text{gas}} w_{\text{gas}} + \Pi_{\text{grid}} w_{\text{grid}} = 0, \quad (\lambda). \tag{20d}$$

Here, w_{grid} are all the variables in the grid side and w_{gas} are all the variables in the gas side. This problem decomposes if the coupling constraints (20d) are removed. The coupling constraints correspond to (17) (after introducing slack variables) and we note that all that is needed to form the coupling constraints are the matrices Π_{gas} and Π_{grid}. These are trivial matrices (containing only zeros and ones).

2.6 Stochastic Formulations

In the presence of uncertainty, both gas and grid sides will seek to make decisions in anticipation of the future. We can formulate this decision process as a two-stage problem of the form in (21a). Here, $\omega \in \Omega$ is the scenario realization, $\mathbb{E}[\cdot]$ denotes the expectation, and constraints (21e) model nonanticipativity conditions for the grid and gas sides. The expected value form can also be used to model conditional value at risk functions (Zavala 2014). The nonanticipativity conditions separate the variables into here-and-now decisions (decisions made before uncertainty is revealed) and recourse variables (decisions made after uncertainty is revealed). Here-and-now decisions in the gas side can be the allocation of line-pack while in the grid side are forward generation schedules. Recourse decisions on the gas side can be corrections of compressor power to mitigate variability of gas demand while on the power grid side can be corrections on generation to mitigate wind power fluctuations. For more details on stochastic formulations for gas and power grid networks the reader is referred to (Zavala 2014; Pritchard et al. 2010).

$$\min \ \mathbb{E}[\varphi_{\text{grid}}(w_{\text{grid},\omega}) + \varphi_{\text{gas}}(w_{\text{gas},\omega})] \tag{21a}$$

$$\text{s.t. } c_{\text{grid},\omega}(w_{\text{grid},\omega}) \geq 0, \ \omega \in \Omega, \quad (\lambda_{\text{grid},\omega}) \tag{21b}$$

$$c_{\text{gas},\omega}(w_{\text{gas},\omega}) \geq 0, \ \omega \in \Omega, \quad (\lambda_{\text{gas},\omega}) \tag{21c}$$

$$\Pi_{\text{gas},\omega} w_{\text{gas},\omega} + \Pi_{\text{grid},\omega} w_{\text{grid},\omega} = 0, \ \omega \in \Omega \quad (\lambda_\omega) \tag{21d}$$

$$\sum_{\omega \in \Omega} \Pi_\omega(w_{\text{gas},\omega}, w_{\text{grid},\omega}) = 0. \quad (\lambda) \tag{21e}$$

The solution the coupled stochastic problem (21a) gives rise to here-and-now control policies. Uncertainty can model equipment failures or externalities such as wind and solar power. When the systems are decoupled, uncertainty can also model the interface variables between infrastructures (i.e., gas demands from power plants $d_j^{\text{target}}(\cdot)$). We also recall that a deterministic formulation computes the policies using average information of the uncertain data and we recall that a wait-and-see (perfect information) policy can be obtained by dropping the nonanticipativity constraints.

3 Economic and Resiliency Issues

In this section we highlight economic and resiliency gains that can be achieved by coordination. We first compare how a stochastic formulation that uses uncertainty information from the power grid can significantly aid resiliency of pipeline operations.

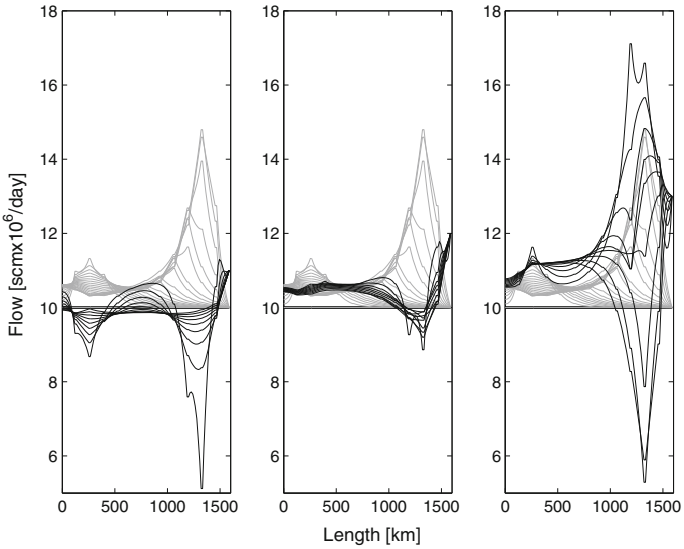


Fig. 1 Optimal flow profiles for low (*left*), medium (*middle*), and high (*right*) demand scenarios. Stochastic (here-and-now) policy

To demonstrate this, we use a long pipeline system with 13 nodes and 12 pipelines and 10 compressors that spans 1,600 km. This system is reported by (Zavala 2014).

In Fig. 1 we present the axial flow profiles for a pipeline system for three demand scenarios. The gray lines are the axial profiles during the charging phase of line-pack in the system. The axial profiles in this charging phase are the same for all scenarios because these are nonanticipative decisions. Note that demand is satisfied at each time step for all scenarios. From Fig. 1 we also see that the optimal here-and-now (HN) policy consists of progressively accumulate line-pack toward the end of the system reflected by a large increase of flow close to the demand node. Once uncertainty is revealed, the system takes three different paths. The total compression energy for the low, medium, and high demand scenarios is 92,742, 99,783 and 113,672 kWh, respectively.

We now compare the performance of the deterministic and wait-and-see (WS) policies under perfect information. The optimal flow profiles for the deterministic formulation are presented in Fig. 2. As can be seen, the policy does not build as much inventory as does the HN solution presented in Fig. 1. As a result, while the low and medium demand scenarios are feasible, the system struggles to satisfy the high demand scenario. In fact, we have found that the demand needs to be curtailed for the system to remain feasible. This high-stress behavior is also reflected in highly volatile flow profiles resulting from aggressive recourse actions.

The axial flow profiles for the WS solution are presented in Fig. 3. As can be seen, the profiles are similar to those of the HN solution presented in Fig. 1. The

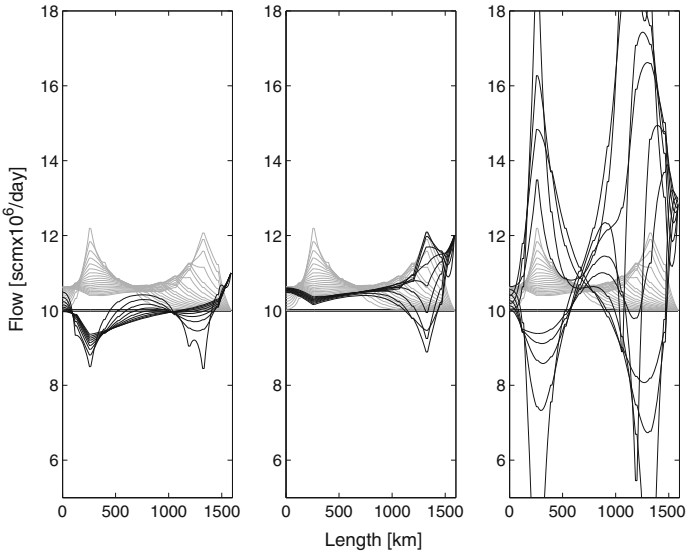


Fig. 2 Optimal axial flow profiles for low (*left*), medium (*middle*), and high (*right*) demand scenarios. Deterministic policy

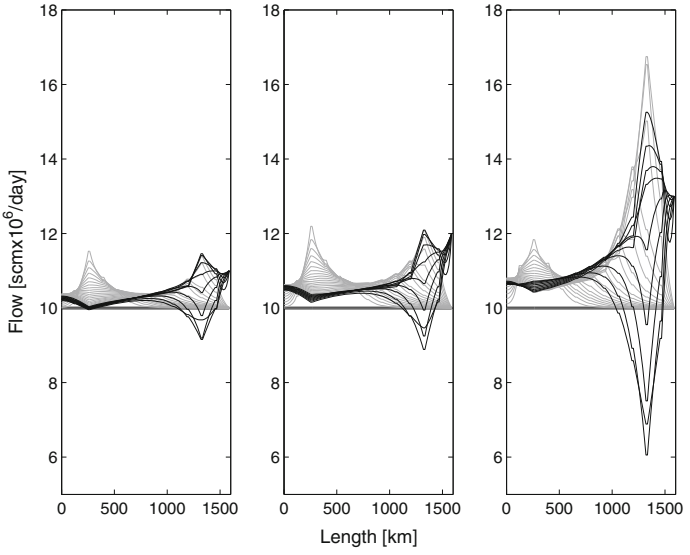


Fig. 3 Optimal axial flow profiles for low (*left*), medium (*middle*), and high (*right*) demand scenarios. Wait-and-see (perfect information) policy

ideal WS solution, however, presents less volatile flow profiles than those of the HN counterpart. This is because inventory can be planned differently for each scenario (we have perfect information). This is particularly evident in the low and high demand scenarios. The total power consumed in the WS scenarios is 88,245, 98,914 and 112,779 kWh, respectively. This is less than 1 % per scenario, compared with the HN solution. Clearly, significant resiliency can be gained by using the HN formulation over the deterministic one without sacrificing much performance over the ideal case of perfect information. These results illustrate that significant improvements in resiliency can be obtained by anticipating gas demand scenarios of power plants. Unfortunately, *deriving uncertainty characterizations for power plant demands is extremely difficult* because these are fully correlated to the decisions of the ISO on the power grid side. Consequently, coordination between gas and power grid sides would require the ISO to provide suitable demand scenarios that are compatible with its dispatch operations. We now use a case study in Illinois to illustrate this point and other economic issues. We also use this larger system to quantify the effect of increased control flexibility achieved by coordination.

The Illinois power grid transmission system comprises 2,522 lines, 1,908 nodes, 870 demands points, and 225 generators points (153 gas-fired generators). Because of the difficulty in obtaining natural gas infrastructure data, we construct a simulated natural gas network system using the basic topology reported by the EIA[4] and by using engineering insight to ensure gas supply to all the gas-fired power plants under nominal conditions. The gas network comprises 215 pipeline segments, 157 nodes, 12 compression stations, and 4 supply points. The resulting network is sketched in Fig. 4.

We compare economic performance for the infrastructures under coordinated and uncoordinated settings. The results are summarized in Table 1. In our simulations, the electrical loads were always satisfied; consequently, we report only the generation cost component of the grid cost (we denote this as φ_{grid}). From the results we make the following observations:

- Under a coordinated setting the power cost decreases by 0.38 % which represents a total of $140,000. The gas cost decreases by 7.54 %, which corresponds to a total of $1,020,000.
- Under an uncoordinated setting only 96 % of the gas requested is delivered. At a gas price of 3 $/MMBTU,[5] *the total undelivered gas has an economic value of* $605,000.
- Under a coordinated setting the compression cost increases by 17.4 %. This is the result of an increased amount of gas delivered to the power plants. In particular, 7.5 % *more gas is delivered under the coordinated setting.* At a gas price of 3 $/MMBTU, the value of the additional gas delivered is $1,080,000.

[4] http://tinyurl.com/cssgzr9.

[5] http://www.eia.gov/naturalgas/weekly/.

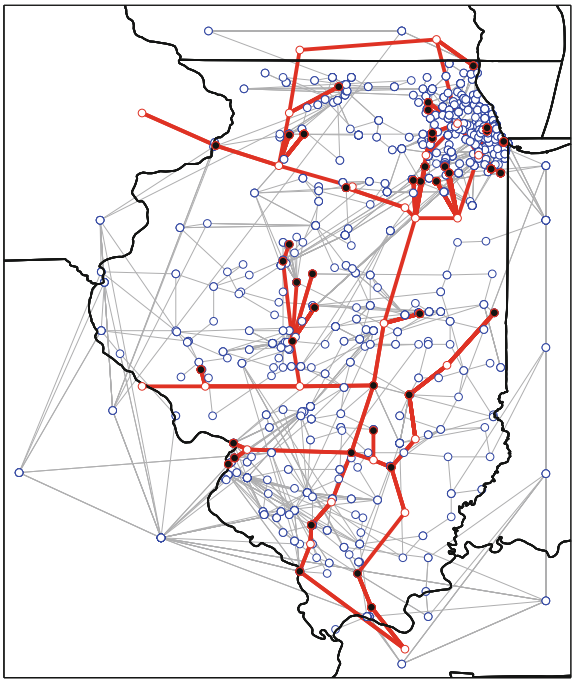


Fig. 4 Illinois electrical (*thin lines*) and gas (*thick lines*) transmission systems. *Black dots* are gas-fired power plants

Table 1 Economic performance under coordinated and uncoordinated settings

	φ_{grid} (M$)	φ_{gas} (M$)	$\varphi^{\text{gas,comp}}$ ($)	$d^{\text{gas,target}}$ (scm $\times 10^{-6}$)	d^{gas} (scm $\times 10^{-6}$)	$\mathcal{R}$ (M$)
Uncoord	36.54	−13.52	28,618	141.25	135.54	2.70
Coord	36.40	−14.54	33,600	145.74	145.74	3.50

(scm standard cubic meters and M$ million U.S. dollars)

Note that the total increase in compression cost is negligible compared to the additional value of the delivered demand.

- Under a coordinated setting the revenue for the gas-fired generators increases by 29.6 %, which corresponds to a total of $800,000. This is the result of the additional gas delivered and the decreased revenue penalties resulting from coordination.

It is rather surprising that both the gas and power grid sides benefit from coordination (i.e., *the objectives of the gas and grid operators do not compete*). Moreover, gas-fired generators increase their revenue. We can explain the decreased performance under an uncoordinated setting from the fact that the power grid operator cannot easily determine how much gas the gas network can deliver at different spatial locations and at different times. Thus, the power grid operator can

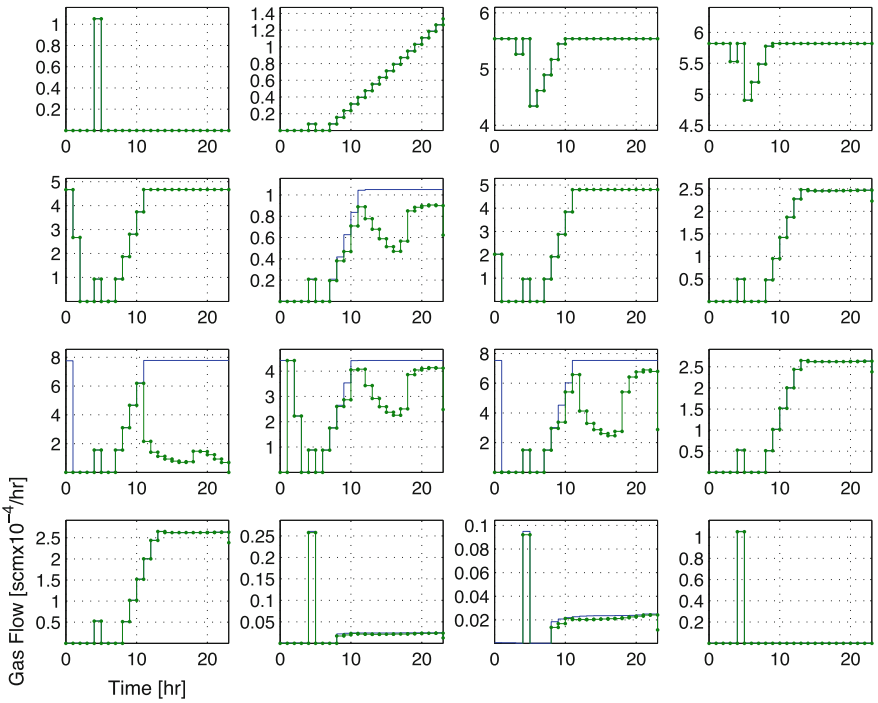


Fig. 5 Requested (*blue solid line*) and realized (*green dotted line*) gas demands for 16 power plants under an uncoordinated setting

be overly optimistic (as in the case presented) or pessimistic about the amount of gas that can actually be delivered. This situation is clearly illustrated in Fig. 5, where we present the target and realized gas demands for 16 different gas-fired generators under the uncoordinated setting. Note that the gas network cannot deliver the total amount of gas requested at four locations. The resulting error in the prediction introduces a penalty for both the power grid and the gas-fired generators. In particular, the power grid operator has to dispatch more expensive power plants, resulting in higher a generation cost and the gas-fired power plants have to pay for the unserved generation. Note also that, *even if the gas operator knows the gas demands of the power grid in advance*, it cannot guarantee to satisfy such demands due to physical constraints.

The presence of gas shortages at several power plant locations *would suggest that the gas pipeline system is constrained by design*. We now demonstrate that this perception is not correct, under a coordinated setting this *delivery bottleneck can in fact be avoided through better control*. In fact, as we have already seen in Table 1, 7 % *more gas can be delivered under the coordinated setting*; we now explain why this is the case. From Fig. 6 we see that dispatch under a coordinated setting is

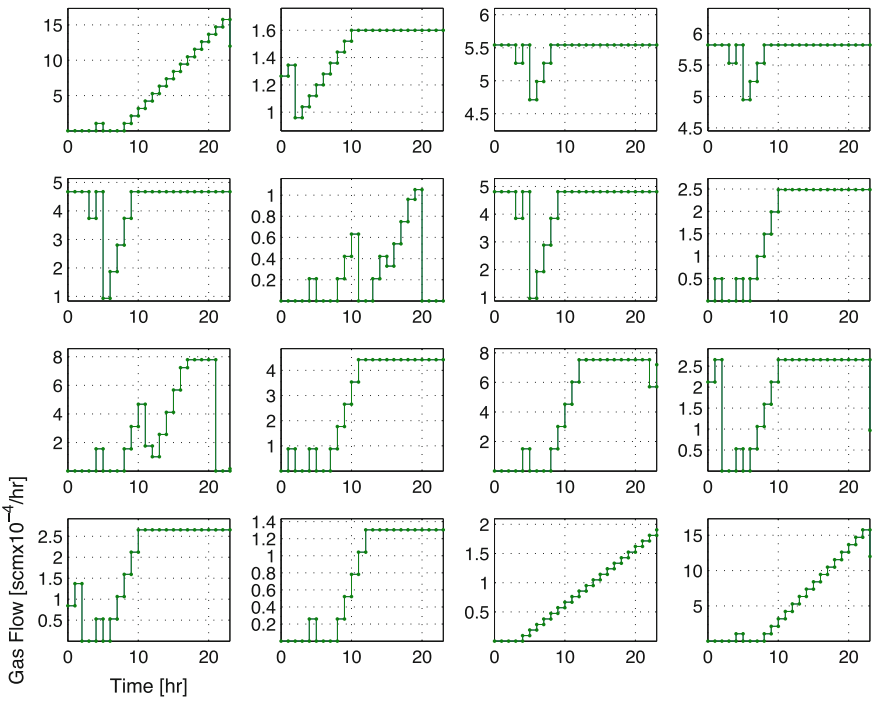


Fig. 6 Requested (*blue solid line*) and realized (*green dotted line*) gas demands for 16 power plants under coordinated setting

smoother (i.e., demands are ramps instead of aggressive withdrawals) and this significantly enhances the flexibility to deliver gas at other locations. In particular, all gas demands can be delivered. We can explain this increased flexibility by noticing that, under a coordinated setting, gas-fired generators act as *distributed demand response resources* that the gas operator can use to better control network pressures and flows and avoid delivery bottlenecks. In other words, *gas-fired power plants become assets rather than liabilities to the pipeline operator.*

Figures 5 and 6 also help us illustrate the rather arbitrary patterns of gas demands emanating from power plants, which are dictated by dispatch decisions of the ISO. These gas demands patterns are to be contrasted with gas and electricity demands of urban areas, which follow strong periodic patterns that are correlated to weather and behavior. The irregular patterns of power plants demands make it difficult to model their uncertainty and this can prevent the adoption of advanced stochastic optimal control solutions. The ISO plays a critical role here, as it can provide characterizations of uncertainty by using possible gas demand scenarios generated under different wind power, weather, and contingency scenarios.

4 Computational Issues

In this section we review problem structures arising in emerging optimal control models and discuss available solution strategies. We also highlight limitations of state-of-the-art solvers and discuss open challenges.

4.1 Emerging Model Structures

The deterministic (single scenario) grid-gas OCP can be expressed in the more general form

$$\min \quad \sum_{p\in\mathscr{P}} \varphi_p(x_p) \tag{22a}$$

$$\text{s.t.} \quad c_p(x_p) = 0, \quad p \in \mathscr{P} \quad (\lambda_p) \tag{22b}$$

$$x_p \geq 0, \quad p \in \mathscr{P} \quad (v_p) \tag{22c}$$

$$\sum_{p\in\mathscr{P}} \Pi_p x_p = 0 \qquad (\lambda_0) \tag{22d}$$

where $\mathscr{P} := \{\text{gas,grid}\}$. Here, $\lambda_0 \in \Re^{m_0}$ are the multipliers of the coupling constraints, $\lambda_p \in \Re^{m_p}$, $v_p \in \Re^{n_p}$ are the multipliers of the partition p constraints and bounds, $x_p \in \Re^{n_p}$ are the primal variables of the partition p, and $\Pi_p \in \Re^{m_0 \times n_p}$ are the coupling matrices. It is important to observe that this structure also arises in stochastic programming problems. For instance, the stochastic OCP for the gas side or for the power grid side can be expressed as (22).

Under a primal–dual interior point framework, the KKT system of problem (22) can be permuted into the block bordered diagonal (BBD) form:

$$\underbrace{\begin{bmatrix} K_0 & B_1^T & B_2^T & \cdots & B_{|\mathscr{P}|}^T \\ B_1 & K_1 & & & \\ B_2 & & K_2 & & \\ \vdots & & & \ddots & \\ B_{|\mathscr{P}|} & & & & K_{|\mathscr{P}|} \end{bmatrix}}_{:=M(\delta_w,\delta_c)} \begin{bmatrix} \Delta w_0 \\ \Delta w_1 \\ \Delta w_2 \\ \vdots \\ \Delta w_{|\mathscr{P}|} \end{bmatrix} = - \begin{bmatrix} r_0 \\ r_1 \\ r_2 \\ \vdots \\ r_{|\mathscr{P}|} \end{bmatrix}, \tag{23}$$

where $\Delta w_0 = \Delta\lambda_0$, $\Delta w_p = (\Delta x_p, \Delta\lambda_p)$,

$$K_0 = -\delta_c \mathbb{I}_{m_0}, \quad K_p = \begin{bmatrix} W_p(\delta_w) & J_p^T \\ J_p & -\delta_c \mathbb{I} \end{bmatrix}$$

$$B_p^T = \begin{bmatrix} \Pi_p & 0 \end{bmatrix}, \tag{24}$$

$J_p = \nabla_{x_p} c_p(x_p)$, $W_p(\delta_w) = \nabla_{x_p, x_p} \mathscr{L} + X_p^{-1} V_p + \delta_w \mathbb{I}$. Symbol $\mathscr{L}$ denotes the Lagrange function of (22) and $\delta_w, \delta_c \geq 0$ are regularization parameters. For more details in the derivation of the KKT system the reader is referred to (Kang et al. 2015). The BBD structure of the linear system can be exploited using a parallel Schur decomposition of the form

$$\underbrace{-\left(\delta_c \mathbb{I}_{m_0} + \sum_{p \in \mathscr{P}} B_p K_p^{-1} B_p^T\right)}_{S} \Delta w_0 = -r_0 + \sum_{p \in \mathscr{P}} K_p^{-1} B_p r_p \tag{25a}$$

$$K_p \Delta w_p = -r_p - B_p^T \Delta w_0, \quad p \in \mathscr{P}, \tag{25b}$$

where S is the Schur complement. One can exploit individual structures within each partition. For instance, we can express the optimization problem in the gas partition p = gas as,

$$\min \ \varphi(x, u) \tag{26a}$$

$$\text{s.t.} \quad c(x, u) = 0, \quad (\lambda) \tag{26b}$$

$$x \geq 0, \quad (v_x) \tag{26c}$$

$$u \geq 0, \quad (v_u). \tag{26d}$$

In this structure, u is assumed have the same dimension as the number of degrees of freedom of the problem (e.g., boost pressures and demands). In other words, if u is fixed, then $c(x, u) = 0$ is a square system of equations (in the gas system these correspond to the discretized transport and network equations). The KKT system of problem (26) is given by

$$\begin{bmatrix} W_{xx}(\delta_w) & W_{xu} & J_x^T \\ W_{ux} & W_{uu}(\delta_w) & J_u^T \\ J_x & J_u & \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta u \\ \Delta \lambda \end{bmatrix} = - \begin{bmatrix} r_x \\ r_u \\ r_\lambda \end{bmatrix} \tag{27}$$

where the coefficient matrix on the left hand side is also known as the augmented matrix. By construction, the Jacobian $J_x = \nabla_x c(x, u)$ is square; and, if it is non-singular and $\delta_c = 0$, then we can construct the following null space matrix:

$$Z = \begin{bmatrix} -J_x^{-1} J_u \\ I \end{bmatrix}. \tag{28}$$

The step for u can then be obtained by solving a reduced system of the form $Z^T W(\delta) Z \Delta u = r_Z$ where r_z is an appropriate right-hand side vector and $Z^T W(\delta) Z$ is the reduced Hessian. Having the step for u, we compute the step for x from $\Delta x = -J_x^{-1}(r_u + J_u \Delta u)$. Note that this approach requires factorizations of J_x and of the reduced Hessian $Z^T W(\delta) Z$ instead of factorizations of the entire augmented matrix. Because of this, this approach can yield significant speedups when the number of degrees of freedom u is small.

The expression of structures also facilitates model construction and can accelerate model processing overhead of algebraic modeling languages (e.g., generation of derivative information and sparsity structures) which is significant in large-scale models such as the ones arising in large infrastructures. For instance, the gas–electric coupled problem when discretized in space gives an NLP with 249,919 variables, 224,292 equality constraints, and 154,093 inequality constraints. If we replicate this model over multiple scenarios we can see that 100 scenarios already give an NLP with 25 million variables that cannot be processed with existing algebraic modeling languages such as AMPL, GAMS, or JuMP. Consequently, it is necessary to partition the problem to enable processing and model storage in memory.

We now demonstrate the benefits of identifying structures and exploiting them in high-performance computers. To do so, we use the parallel interior point solver PIPS-NLP developed by (Chiang et al. 2014) to solve a stochastic OCP for a gas pipeline system with 13 nodes and 12 pipelines and 10 compressors. We consider two settings: in the first case we exploit only the stochastic structure (Table 2) while in the second case we exploit the stochastic and the reduced space structure of the problem (Table 3). The NLP under study has 96 scenarios and a total of 1,930,752 variables. The problem is solved on the distributed-memory cluster Fusion at Argonne National Laboratory. As can be seen, strong scaling is observed in both cases and solution times can be brought down from an hour to less than 7 min. Moreover, these NLPs cannot be solved in a single processor because of the large

Table 2 Scalability of PIPS-NLP exploiting stochastic structure

Scenarios	n	Obj	Iter	Time (hh:mm:ss)	MPI proc
96	1,930,752	1.39×10^2	42	01:13:16	8
96	1,930,752	1.39×10^2	42	00:38:18	16
96	1,930,752	1.39×10^2	42	00:24:55	24
96	1,930,752	1.39×10^2	42	00:19:23	32
96	1,930,752	1.39×10^2	42	00:12:42	48
96	1,930,752	1.39×10^2	42	00:06:48	96

Table 3 Scalability of PIPS-NLP exploiting stochastic and reduced space structure

Scenarios	n	Obj	Iter	Time (hh:mm:ss)	MPI proc
96	1,930,752	1.39×10^2	42	00:29:54	8
96	1,930,752	1.39×10^2	42	00:14:45	16
96	1,930,752	1.39×10^2	42	00:10:00	24
96	1,930,752	1.39×10^2	42	00:07:36	32
96	1,930,752	1.39×10^2	42	00:05:14	48
96	1,930,752	1.39×10^2	42	00:02:54	96

memory requirements. We also observe that exploiting the reduced space structure inside each scenario decreases the computational time by a factor of nearly 3.

4.2 Dealing with Negative Curvature

A key difference between convex and nonconvex NLPs is the potential presence of *negative curvature*. The presence of negative curvature indicates that the Newton step computed from the solution of the KKT system (23) might not correspond to a minimum of the associated quadratic programming problem. In a line-search setting this is an important issue because the Newton step cannot be guaranteed to provide a descent direction for the objective function when the constraint violation is sufficiently small, which is a key requirement to ensure global convergence.

The presence of negative curvature can be checked by computing the inertia of the KKT system (23). This can be done using Haynsworth's formula:

$$\text{Inertia}(M(\delta_w, \delta_c)) = \sum_{p \in \mathscr{P}} \text{Inertia}(K_p) + \text{Inertia}(S). \tag{29}$$

We recall that $n = \sum_p n_p$ and $m = m_0 + \sum_p m_p$. Consequently, if we have that $\text{Inertia}(K_p) = \{n_p, m_p, 0\}$ for all $p \in \mathscr{P}$ then the inertia of $M(\delta_w, \delta_c)$ is correct if and only if $\text{Inertia}(S) = \{0, m_0, 0\}$. When the inertia is correct we can thus guarantee that the Newton step is a descent direction. One can obtain the inertia of the blocks K_p using LBL^T factorizations. If the problem has nested structures, one can also obtain the inertia of each block K_p by applying Haynsworth's formula recursively.

If the inertia of the linear system (23) (or a subsystem) obtained from (29) is not correct we progressively increase the regularization parameter δ_w until the KKT system has correct inertia. We call this the *inertia-based* regularization strategy. Note that every time we increase the regularization parameter we need to solve the linear system again. Consequently, the presence of negative curvature can make the computation of the Newton step very expensive.

Obtaining the inertia of Schur matrices using an LBL^T factorization can be inefficient because the Schur complement is often a dense matrix or contains dense

blocks. To avoid these limitations, we have recently proposed an inertia-free test of the form,

$$\Delta w^T W(\delta_w)\Delta w \geq \kappa \Delta w^T \Delta w \tag{30}$$

for $\kappa > 0$ (Chiang and Zavala 2014). Here, Δw is a Newton step computed along the null space of the constraint Jacobian and $W(\delta_w)$ is the *entire* Hessian matrix of the NLP (22). In the inertia-free approach we increase the regularization parameter until the curvature test (30) holds. The key observation is that this test implicitly requires the Newton step to be a descent direction directly which can occur even if the linear system does not have correct inertia. Consequently, the inertia-free test provides more flexibility to accept steps. To demonstrate this feature, we solve a large-scale stochastic optimal control instance and compare inertia-based regularization (IBR) and inertia-free regularization (IFR). This is an NLP with 128 uncertain scenarios, 1,024,651 variables, and 1,023,104 constraints. The results obtained with PIPS-NLP are presented in Table 4. We can see that, despite the high nonlinearity of the large-scale instances, the inertia-free approach IFR converges in all instances. In general IFR requires more iterations than does IBR but the number of factorizations is reduced, resulting in faster solutions.

4.3 Open Issues

We can pose the stochastic programming problem for the coupled grid-gas system (21a) in the general form:

$$\min \sum_{p\in\mathscr{P}} \sum_{j\in\mathscr{P}_p} \varphi_{p,j}(x_{p,j}) \tag{31a}$$

$$\text{s.t.} \quad c_{p,j}(x_{p,j}) = 0, \quad p \in \mathscr{P}, \quad j \in \mathscr{P}_p \quad (\lambda_{p,j}) \tag{31b}$$

$$x_{p,j} \geq 0, \quad p \in \mathscr{P}, \quad j \in \mathscr{P}_p \quad (v_{p,j}) \tag{31c}$$

Table 4 Performance of inertia-based and inertia-free regularization strategies on stochastic OCP for gas system

		IBR			IFR		
#MPI	Obj	Iter	Linear solves	Time (s)	Iter	Linear solves	Time (s)
8	1.26E-02	153	278	832	93	106	491
16	1.26E-02	136	251	363	109	122	315
32	1.26E-02	146	274	209	99	112	143
64	1.26E-02	157	286	123	101	114	79
128	1.26E-02	145	275	64	109	125	52

$$\sum_{j \in \mathscr{P}_p} \Pi_{p,j} x_{p,j} = 0 \quad (\lambda_{0,j}) \tag{31d}$$

$$\sum_{p \in \mathscr{P}} \Pi_p x_p = 0 \quad (\lambda_0) \tag{31e}$$

Here, $\mathscr{P}$ is the scenario set and $\mathscr{P}_p = \{\text{gas,grid}\}$ is the system partition. One can show that this nested problem yields an augmented system of the form (23) in which each diagonal block K_p has a BBD structure of the form

$$K_p = \begin{bmatrix} K_{p,0} & B_{p,1}^T & B_{p,2}^T & & B_{p,P_p}^T \\ B_{p,1} & K_{p,1} & & & \\ B_{p,2} & & K_{p,2} & & \\ \vdots & & & \ddots & \\ B_{p,P_p} & & & & K_{p,P_p} \end{bmatrix}, \quad p \in \mathscr{P}. \tag{32}$$

Consequently, we can also apply a Schur decomposition to perform solves with the block system K_p. Moreover, structures such as the reduced space structure of the gas side can be exploited as well. This would give a linear system with three nested structures.

The coupled stochastic problem is of interest because it would be desirable to understand the increasing resiliency gained by coordination. As we have seen, for instance, coordination enables the gas infrastructure to deliver significantly larger amounts of gas and this flexibility can be used to withstand abrupt variations of wind power. Solving coupled stochastic problems on realistic networks, however, is extremely challenging and defies the scope of state-of-the-art solvers. To give an idea of the complexity, for the Illinois system we have found that the solution time of a single scenario problem for the coupled gas–electric system is 40 min. While it is possible to partition the Illinois grid and gas systems using Schur decomposition, this is not always beneficial. In particular, we have found that performing a direct sparse factorization of the entire coupled system is more efficient that performing Schur decomposition. In other words, the benefits of Schur decomposition are only observed when the spatial domain (network size or spatial discretization resolution) of the gas and network systems increase. Consequently, the only alternative to accelerate solutions seems to coarsen the spatial discretization of the gas transport equations.

To illustrate the effect *coarsening*, we compare the economic and computational performance of coupled gas–electric problems with low- and high-resolution spatial discretizations. We compare the results using our base implementation with $N_x = 10$ spatial points per pipeline and a low-resolution implementation with $N_x = 3$ spatial points per pipeline. The low-resolution problem gives an NLP with 141,559 variables, 115,932 equality constraints, 157,765 inequality constraints, and 25,627

Table 5 Computational results for coupled problems for base and perturbed topologies

	Iter.	Time(s)	Linear solve (−)	Time/Linear solve (sec)
$N_x = 10$	232	2401.01	311	7.72
$N_x = 3$	157	573.68	188	3.05

Table 6 Economic performance under low- and high-resolution spatial discretizations

	φ_{grid} (M$)	φ_{gas} (M$)	$\varphi^{\text{gas,comp}}$ ($)	$d^{\text{gas,target}}$ (scm $\times 10^{-6}$)	d_{gas} (scm $\times 10^{-6}$)	$\mathcal{R}_{\text{gas}}$ (M$)
$N_x = 10$	36.40	−14.54	33,600	145.74	145.74	3.50
$N_x = 3$	36.39	−14.55	33,356	145.83	145.83	3.48

degrees of freedom. The high-resolution problem gives an NLP with 249,919 variables, 224,292 equality constraints, 154,093 inequality constraints, and 25,627 degrees of freedom. The number of degrees of freedom remains unchanged because all of these enter at the network nodes and are thus independent of the discretization resolution (Zavala 2014). This is an important structural property of gas optimal control formulations. The results comparing high and low resolutions are presented in Table 5. The solution time is reduced from 40 to about 10 min. Most notably, from Table 6 we can see that coarsening does not introduce large errors in economic performance. This behavior, however, cannot be guaranteed in general. It is thus necessary to devise linear algebra strategies that can perform coarsening adaptively at the linear algebra level to create preconditioners. Multigrid schemes are available for solving OCPs with embedded PDEs but these schemes are currently not general enough to handle hyperbolic PDEs as those arising in gas networks and to handle mixed sets of constraints. This is an important research area.

5 Conclusions

We have discussed characteristics of emerging optimal control models for interconnected natural gas and electrical networks. These models are motivated by the increasing interest in understanding the economic benefits of coordination and the need to understand system resiliency in the face of extreme weather events and renewable power adoption. We have seen that optimal control problems arising in infrastructures are highly structured and that these structures can be exploited to accelerate solutions and avoid memory bottlenecks. We have also demonstrated that existing state-of-the-art tools can handle large and highly nonlinear models but these capabilities are insufficient to handle geographical regions of practical interest.

Acknowledgments This material is based upon work supported by the U.S. Department of Energy, Office of Science, under contract number DE-AC02-06CH11357. Victor M. Zavala acknowledges the support of the Early Career Program of the U.S. Department of Energy. We also acknowledge the computing resources provided on Fusion and Blues, high-performance computing clusters operated by the Laboratory Computing Resource Center at Argonne National Laboratory.

References

January 2014 FERC data request. *ISO New England Technical Report*, 2014.

M. Abbaspour and K. Chapman. (2008) Nonisothermal transient flow in natural gas pipeline. *Journal of Applied Mechanics*, 750 (3):031018, 2008.

S. An, Q. Li, and T. W. Gedra. (2003) Natural gas and electricity optimal power flow. In *Transmission and Distribution Conference and Exposition, 2003 IEEE PES*, volume 1, pages 138–143. IEEE, 2003.

M. Arnold, R. R. Negenborn, G. Andersson, and B. De Schutter. (2010) Distributed predictive control for energy hub coordination in coupled electricity and gas networks. In *Intelligent Infrastructures*, pages 235–273. Springer, 2010.

B. Baumrucker and L. T. Biegler. (2010) MPEC strategies for cost optimization of pipeline operations. *Computers & Chemical Engineering*, 340 (6):900–913, 2010.

L. T. Biegler. (2010) *Nonlinear programming: concepts, algorithms, and applications to chemical processes*. Society for Industrial and Applied Mathematics, 2010.

D. C. Cafaro and I. E. Grossmann. (2014) Strategic planning, design, and development of the shale gas supply chain network. *AIChE Journal*, 600 (6):2122–2142, 2014.

R. G. Carter, T. F. Dupont, and H. H. Rachford Jr. (2004) Pack management and transient optimization of natural gas transmission lines. In *IGRC Conference, Vancouver*, 2004.

P. Centolella and A. Ott. (2009) The integration of price responsive demand into PJM wholesale power markets and system operations. *HEPG Technical Report*, 2009.

M. Chaudry, N. Jenkins, and G. Strbac. (2008) Multi-time period combined gas and electricity network optimisation. *Electric Power Systems Research*, 780 (7):1265–1279, 2008. ISSN 0378–7796.

N. Chiang and V. M. Zavala. (2014) An inertia-free filter line-search algorithm for large-scale nonlinear programming. 2014. Preprint ANL/MCS-P5197-0914, Argonne National Laboratory.

N. Chiang and V. M. Zavala. (2015) Large-scale optimal control of interconnected natural gas and electrical transmission systems. 2015. Preprint ANL/MCS-P5348-0515, Argonne National Laboratory.

N. Chiang, C. G. Petra, and V. M. Zavala. (2014) Structured nonconvex optimization of large-scale energy systems using PIPS-NLP. In *Proc. of the 18th Power Systems Computation Conference (PSCC), Wroclaw, Poland*, 2014.

M. Geidl and G. Andersson. (2007) Optimal power flow of multiple energy carriers. *Power Systems, IEEE Transactions on*, 220 (1):145–155, 2007.

J. Kang, N. Chiang, C. D. Laird, and V. M. Zavala. (2015) Nonlinear programming strategies on high-performance computers. In *Proc. of the IEEE Conference on Decision and Control, Osaka, Japan*, 2015.

C. Liu, M. Shahidehpour, Y. Fu, and Z. Li. (2009) Security-constrained unit commitment with natural gas transmission constraints. *Power Systems, IEEE Transactions on*, 240 (3):1523–1536, 2009.

C. Liu, M. Shahidehpour, and J. Wang. (2011) Coordinated scheduling of electricity and natural gas infrastructures with a transient model for natural gas flow. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 210 (2):025102, 2011.

D. Marqués and M. Morari. (1988) On-line optimization of gas pipeline networks. *Automatica*, 240 (4):455–469, 1988.

A. L. Ott. Experience with PJM market operation, system design, and implementation. *Power Systems, IEEE Transactions on*, 180 (2):528–534, 2003.

G. Pritchard, G. Zakeri, and A. Philpott. (2010) A single-settlement, energy-only electric power market for unpredictable and intermittent participants. *Operations Research*, 580 (4-part-2):1210–1219, 2010.

M. Qadrdan, M. Chaudry, J. Wu, N. Jenkins, and J. Ekanayake. (2010) Impact of a large penetration of wind generation on the GB gas network. *Energy Policy*, 380 (10):5684 – 5695, 2010. ISSN 0301–4215.

H. H. Rachford Jr, R. G. Carter, T. F. Dupont, et al. (2009) Using optimization in transient gas transmission. In *PSIG Annual Meeting*. Pipeline Simulation Interest Group, 2009.

H. R. Rachford Jr and R. G. Carter. (2004) Method and apparatus for determining optimal control settings of a pipeline, 2004. US Patent 6,701,223.

J. B. Rawlings, D. Angeli, and C. N. Bates. (2012) Fundamentals of economic model predictive control. In *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, pages 3851–3861. IEEE, 2012.

R. Z. Ros-Mercado and C. Borraz-Sánchez. (2015) Optimization problems in natural gas transportation systems: A state-of-the-art review. *Applied Energy*, 147:536–555, 2015. ISSN 0306–2619.

M. C. Steinbach. (2007) On PDE solution in transient optimization of gas networks. *Journal of Computational and Applied Mathematics*, 2030 (2):345–361, 2007.

V. M. Zavala. (2014) Stochastic optimal control model for natural gas networks. *Computers & Chemical Engineering*, 64:103–113, 2014.

V. M. Zavala and L. T. Biegler. (2009) Optimization-based strategies for the operation of low-density polyethylene tubular reactors: nonlinear model predictive control. *Computers & Chemical Engineering*, 330 (10):1735–1746, 2009.

Part III
Processing of Alternatives Raw Materials

Equation-Based Design, Integration, and Optimization of Oxycombustion Power Systems

Alexander W. Dowling, John P. Eason, Jinliang Ma, David C. Miller and Lorenz T. Biegler

Abstract The application of "systems-based tools" including exergy/pinch analysis and process simulation has facilitated increases in the thermal efficiency of ambient pressure oxycombustion coal-fired power systems with carbon capture from 36 to 39–40 $\%_{LHV}$, while also considering capital costs. This corresponds to a decrease in the energy penalty 10 %-points to 6–7 %-points (absolute), relative to reference air-fired coal power plants without CO_2 capture (46 $\%_{LHV}$). These efficiency improvements are primarily due to tailored next-generation air separation systems and plant-wide heat integration. Furthermore, oxycombustion power systems are an ideal candidate for numerical optimization, given the complex interactions between its five subsystems. This chapter extensively surveys the oxycombustion literature and summarizes four key design questions. A new, fully equation-based, flowsheet optimization framework is then introduced and applied to three oxycombustion-related case studies: design of a minimum energy air separation unit to produce an O_2 enriched stream for the boiler, optimization of the CO_2 polishing unit and compression train to minimize specific energy, and maximization of thermal efficiency in the oxy-fired steam cycle using a hybrid 1D/3D boiler model.

A.W. Dowling (✉) · J.P. Eason · L.T. Biegler
Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, USA
e-mail: awdowlin@alumni.cmu.edu; adowling2@wisc.edu

J.P. Eason
e-mail: jeason@andrew.cmu.edu

L.T. Biegler
e-mail: biegler@cmu.edu

J. Ma
AECOM & National Energy Technology Laboratory,
3610 Collins Ferry Road, P.O. Box 880, Morgantown, WV 26507, USA
e-mail: Jinliang.Ma@netl.doe.gov

D.C. Miller
National Energy Technology Laboratory, 626 Cochrans Mill Road, P.O. Box 10940,
Pittsburgh, PA 15236, USA
e-mail: David.Miller@netl.doe.gov

M. Martín (ed.), *Alternative Energy Sources and Technologies*,
DOI 10.1007/978-3-319-28752-2_5

1 Introduction

Motivated by the threat of climate change, global leaders have invested significant resources in developing CO_2 capture technologies for fossil fuel electricity generation systems to serve as bridge to a fully renewable power generation future. In general, there are three main strategies for CO_2 capture: post combustion from flue gas (N_2/CO_2 is the major separation), precombustion in Integrated Gasification Combined Cycle (IGCC) plants (H_2/CO_2 is the major separation) and oxycombustion (N_2/O_2 is the major separation). All three of these strategies produce a CO_2-rich product stream which may be used for enhanced oil recovery, and could be permanently stored in geological formations.[1] Cuèllar-Franca and Azapagic (2015) provide a complete life-cycle analysis of each of these carbon capture strategies. In oxycombustion (i.e., oxy-fired) systems, fossil fuel(s) are combusted in a N_2 lean atmosphere, producing a CO_2- and H_2O-rich flue gas, which is significantly easier to process into a high-purity CO_2 exit stream than flue gas in post-combustion capture schemes. Another significant benefit of oxycombustion is the possibility to retrofit existing power facilities, without significant changes to the power island (steam cycle). Several authors have compiled extensive reviews of oxycombustion technology and outstanding research challenges (Buhre et al. 2005; Chen et al. 2012; Scheffknecht et al. 2011; Toftegaard et al. 2010; Wall et al. 2009). Although this chapter focuses on optimization-based strategies for the design and integration of oxy-fired coal power systems, the concepts may be extended to oxycombustion of other fuels, including natural gas combined cycles (Amann et al. 2009; Bolland and Undrum 2003) and co-firing with biomass (Arias et al. 2008; Haykiri-Acma et al. 2010; López et al. 2014; Pawlak-Kruczek et al. 2013).

1.1 Process Overview

An oxycombustion power plant consists of five key subsystems, which are shown in Fig. 1: the boiler (i.e., furnace), pollution controls, power island, air separation unit, and CO_2 polishing unit. The latter two separation systems distinguish oxy- and air-fired power plants, and contribute to an approximately 10 %-point (absolute) penalty in thermal efficiency relative to a comparable air-fired pulverized power plant. The thermal efficiency of ambient pressure oxycombustion power plants are typically 30–31 $\%_{HHV}$ (36–38 $\%_{LHV}$), whereas air-fired supercritical PC plant thermal efficiencies are around 41 $\%_{HHV}$ (46 $\%_{LHV}$) without CO_2 capture and 31.2 $\%_{HHV}$ with post-combustion CO_2 capture.[2] Table 1 compares thermal

[1]Utilization applications for CO_2 from oxycombustion requires an additional CO_2/Ar, N_2, O_2 separation step.

[2]HHV and LHV abbreviate higher heating value and lower heating value, respectively.

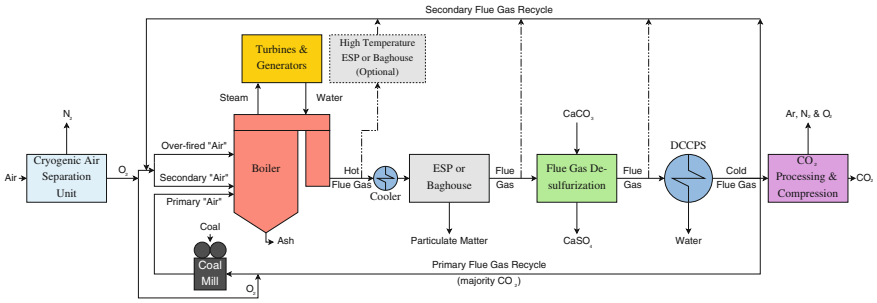


Fig. 1 The flowsheet of an oxycombustion process is shown above with three alternate secondary flue gas recycle locations (*dashed lines*, labeled Recycle Cases A, B, and C), which allow for secondary flue gas recycle at different temperature levels and higher cycle thermal efficiencies. For example, using the alternate and gas recycle immediately after the boiler (Case A, warmest option), the least amount of exergy is lost due to flue gas cooling, resulting in a 1 %-point (absolute) smaller energy penalty versus the standard recycle strategy (*solid line*) (Hagi et al. 2013, 2014a). These alternate strategies (Cases A–C), however, are not feasible for high sulfur content coals due to corrosion concerns

Table 1 Net thermal efficiency values for reference air-fired pulverized coal (PC) power plants without CO_2 capture, PC plants with post-combustion capture and various oxycombustion plant designs

Reference	Case description	Thermal efficiency
Air-fired pulverized coal power plants without and with post-combustion CO_2 capture		
Fout et al. (2015)	Subcritical PC w/o CO_2 capture	39.0 $\%_{HHV}$
NETL baseline report	Subcritical PC w/post-comb. CO_2 capture	31.2 $\%_{HHV}$
	Supercritical PC w/o CO_2 capture	40.7 $\%_{HHV}$
	Supercritical PC w/post-comb. CO_2 capture	32.5 $\%_{HHV}$
Oxy-fired pulverized coal power plants with CO_2 capture		
Andersson and Johnsson (2006)	Reference air-fired plant	42.6 %[a]
	Oxy-fired plant	33.5 %[a]
Fu and Gundersen (2010)	Reference air-fired plant w/o CO_2 capture	40.6 $\%_{HHV}$
	Oxy-fired w/o heat integration	30.4 $\%_{HHV}$
	Oxy-fired w/heat integration	30.5–30.8 $\%_{HHV}$
Fu and Gundersen (2012a)	Oxy-fired w/o heat integration	(not reported)
	Oxy-fired w/ASU + steam cycle integration	+0.38 $\%_{HHV}$ to +0.47 $\%_{HHV}$
Fu et al. (2015a)	Reference oxycombustion	30.4 $\%_{HHV}$
	Oxy. with optimal heat integration	30.9–40.0 $\%_{HHV}$
Hagi et al. (2014a)	Reference air-fired plant w/o CO_2 capture	46.1 $\%_{LHV}$
	Oxy-fired (Standard; 47 °C flue gas recycle)	36.1–37.8 $\%_{LHV}$

(continued)

Table 1 (continued)

Reference	Case description	Thermal efficiency
	Oxy-fired (Case C; 73 °C flue gas recycle)	35.8–37.3 $\%_{LHV}$
	Oxy-fired (Case B; 140 °C flue gas recycle)	36.1–38.1 $\%_{LHV}$
	Oxy-fired (Case A; 348 °C flue gas recycle)	35.7–38.3 $\%_{LHV}$
Hagi et al. (2014b)	Atmospheric oxy-fired (conservative)	36.1 $\%_{LHV}$
	Atmospheric oxy-fired (optimized)	39.1 $\%_{LHV}$
	Pressurized oxy-fired (ISOTHERM)	38.4 $\%_{LHV}$
	Pressurized oxy-fired (SPOC)	42.3 $\%_{LHV}$
Hong et al. (2010)	Pressurized oxycombustion plants	33.0–35.0 $\%_{LHV}$
	(same designs)	31.6–33.6 $\%_{HHV}$
Skorek-Osikowska et al. (2013)	Oxy-fired plant (600 °C/30 MPa steam)	30.5–33.4 $\%_{LHV}$
Soundararajan et al. (2014a)	Supercritical oxy. w/o CPU + steam cycle integration	34.8 $\%_{LHV}$
	Supercritical oxy. w/CPU + steam cycle integration	36.3 $\%_{LHV}$
	Subcritical oxy. w/o CPU + steam cycle integration	33.3–33.7 $\%_{LHV}$
	Subcritical oxy. w/CPU + steam cycle integration	34.4–36.3 $\%_{LHV}$
Tranier et al. (2011)	Air-fired supercritical reference plant	39.4 $\%_{HHV}$
	Base oxy-fired supercritical plant	29.3 $\%_{HHV}$
	Optimized oxy-fired supercritical plant	33.6 $\%_{HHV}$
	Air-fired ultrasupercritical reference plant	44.6 $\%_{HHV}$
	Base oxy-fired ultrasupercritical plant	33.0 $\%_{HHV}$
	Optimized oxy-fired ultrasupercritical plant	38.9 $\%_{HHV}$
Xiong et al. (2011a)	Air-fired reference plant	42.4 $\%_{LHV}$
	Oxy-fired plant	32.3 $\%_{LHV}$
Yan et al. (2015)	Air-fired reference (600 °C/26 MPa steam)	41.7 $\%_{LHV}$
	Optimized oxy-fired plant	31.5–31.8 $\%_{LHV}$

In the surveyed oxycombustion literature, most authors report their key findings in terms of a thermal efficiency penalty relative to a reference air-fired plant without CO_2 capture. It is important to note the use of higher versus lower heating values differs between studies and authors.
[a]Heating value basis not reported.

efficiency estimates reported in literature for various oxy-fired and air-fired coal power plant designs. Reducing the efficiency penalty of oxycombustion power systems (and their costs of electricity) requires careful integration of these five subsystems, which is aided by systems analysis and rigorous numerical optimization.

1.1.1 Air Separation

The first step of an oxy-fired power plant requires air to be separated into its primary components. The O_2-rich product stream is used for combustion. Membranes, pressure swing adsorption and cryogenic distillation technologies are industrially prevalent for O_2 production, but the former two are (currently) economically viable for only small or medium quantities of O_2 product (Smith and Klosek 2001). Thus, cryogenic air separation units are considered the most promising technology in the short term for oxy-combustion, although extensive research efforts have focused on the development of economically competitive membranes for O_2 production in power systems (Belaissaoui et al. 2014; Burdyny and Struchtrup 2010; Czyperek et al. 2010; Engels et al. 2010; Hashim et al. 2011; Mancini and Mitsos 2011a, b; Stadler et al. 2011).

In the classic double column cryogenic air separation unit (originally designed by Linde in 1910) (Foerg 2002), pressurized chilled air is fed into the bottom of the high pressure distillation column operating around 4–6 bar. The bottom product of this column is crude liquid O_2, and the distillate is high purity N_2. The bottom product and a fraction of the distillate are expanded and fed into the low-pressure distillation column operating near ambient pressure, which produces high purity O_2 (bottom product) and crude N_2 (top product). The temperature difference between the columns allows for the condenser of the former and reboiler of the latter to be heat integrated. Furthermore, the original feed air is cooled in a multistream heat exchanger using all of the cold products streams. A simplified schematic for a similar air separation unit is shown in Fig. 2. Low energy requirements (in the form of compression work) are achieved by tight heat integration in the process, with minimum temperature differences as low as 1 °C. Alternately, a third column may be used to achieve >97 mol% O_2 by removing Ar, although this is not necessary for most oxycombustion applications, as air leakage in the boiler and particulate removal system counteracts the benefit of producing high purity O_2.

Several researchers have proposed alternate air separation unit (ASU) designs tailored to produce medium purity (around 95 mol%) O_2 for oxycombustion processes. Fu and Gundersen (2012c) noted that the typical double column design requires 4.7 times the thermodynamic theoretical minimum work, and they employed exergy analysis to pinpoint the most inefficient parts of air separation units. This led them to propose different cycles that only compress the N_2 product to supply refrigeration, instead of the entire air feed (Fu and Gundersen 2013). Engineers from Air Products also proposed and compared several alternate ASU configurations for oxycombustion (Higginbotham et al. 2011), and researchers from Japan utilized self heat recuperation in yet another ASU configuration (Fu et al. 2015b). Overall, these advanced designs reduce the specific energy of separation 10–20 %, compared to the "standard" Linde double column configuration. Fu and Gundersen (2012c) estimate the ASU (before improvement) is responsible for a 6.6 $\%_{HHV}$-point thermal efficiency penalty, which is approximately 2/3 of the efficiency penalty for the entire oxycombustion process (10.2 $\%_{HHV}$-points). Due to the complex nature of multistream heat exchangers using in cryogenic air separation

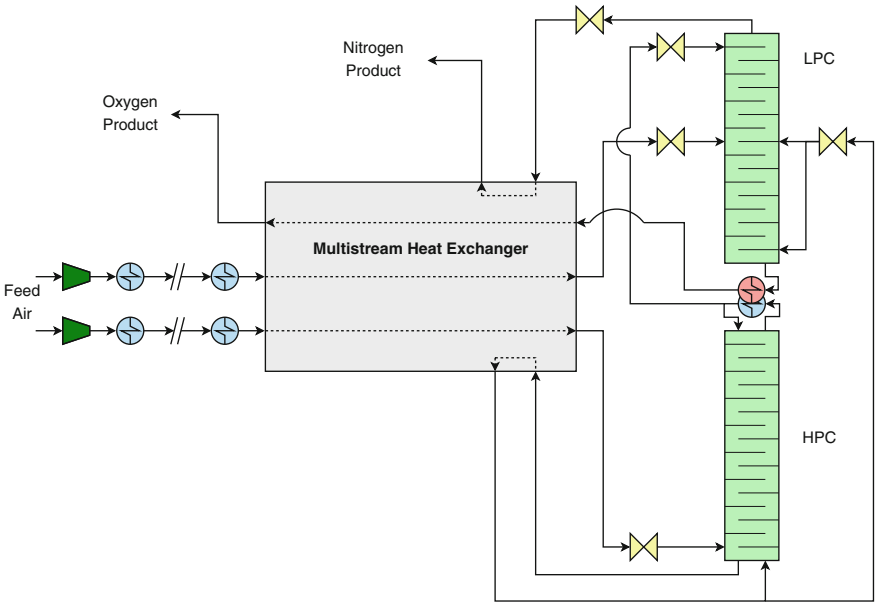


Fig. 2 A simplified schematic for an air separation unit is shown above. Similar to the Linde process (Foerg 2002), the *bottom* column operates at a higher pressure than the *top*, which allows for tight coupling between the condenser and reboiler. A multistream heat exchanger is used to cool the feed air using the product streams. This schematic includes a second air feed stream and extra recycle streams, which are not part of the simple Linde process

processes, it is nearly impossible to assess and compare the capital costs of these designs with meaningful accuracy using methods in open literature (as opposed to property costs models).

1.1.2 Furnace

In the boiler (i.e., furnace), pulverized coal (or another fuel) is combusted to produce heat, which is transferred to a working fluid (e.g., water/steam) through the furnace walls and in the secondary superheater. Additional heat from the hot flue gas is transferred to the working fluid in the primary superheater, reheaters (optional), economizer and air heater (The Babcock and Wilcox Company 2005a, b). In air-fired systems, pulverized coal is carried into boiler with the primary air stream. Additional air is injected into the burners (secondary air) or above them (over-fired air). In an oxycombustion system, recycled flue gas (typically 58–70 mol% CO_2, wet basis) is used to regulate temperatures, with CO_2 acting as a diluent. In this predominantly CO_2/O_2 atmosphere, approximately 28 mol% O_2 (wet basis) is needed to mimic the adiabatic flame temperatures in air-fired boilers (Wall et al. 2009). O_2 enriched air from the ASU is mixed with recycled flue gas before being fed into the boiler as

either primary, secondary or over-fired air, although care is required to reduce the explosion potential in the coal mill. Primary thrusts in oxycombustion research focus on heat transfer properties (Johansson et al. 2013; Krishnamoorthy et al. 2010; Rajhi et al. 2014; Smart et al. 2010a, b; Tan et al. 2006; Warzecha and Boguslawski 2014) and combustion phenomena (e.g., kinetics) (Bejarano and Levendis 2008; Duan et al. 2014; Gil et al. 2012; Molina and Shaddix 2007; Murphy and Shaddix 2006; Rathnam et al. 2009; Senneca and Cortese 2012; Tan et al. 2006) in the CO_2-rich atmosphere, which are significantly different from the N_2 rich atmospheres in traditional air-fired boilers. Research approaches are diverse, ranging from experimental campaigns (Bejarano and Levendis 2008; Croiset and Thambimuthu 2001; Fernández-Miranda et al. 2014; Gil et al. 2012; Kiga et al. 1997; Molina and Shaddix 2007; Murphy and Shaddix 2006; Rathnam et al. 2009; Roy and Bhattacharya 2014; Senneca and Cortese 2012; Smart et al. 2010a, b; Tigges et al. 2009; Zhuang and Pavlish 2012) to computational fluid dynamic simulations (Álvarez et al. 2011; Cao et al. 2010; Krishnamoorthy 2012; Myöhänen et al. 2009; Rajhi et al. 2014; Tigges et al. 2009; Warzecha and Boguslawski 2014). In addition to traditional PC boilers, many researchers have also studied fluidized bed furnaces operated under oxy-fired conditions (Bolea et al. 2012; Duan et al. 2014; Faé Gomes et al. 2015; Hultgren et al. 2014; Leckner and Gómez-Barea 2014; Myöhöanen et al. 2009; Zhou et al. 2013). Furthermore, several groups have assessed the potential to retrofit existing boilers for oxy-fired conditions (Chung et al. 2011; Faé Gomes et al. 2015; Fei et al. 2015; Fry et al. 2011a, b; Kupila et al. 2011; Tigges et al. 2009). In general, they conclude that retrofiting is feasible, although there are additional technical challenges compared to greenfield power plants, especially related to heat transfer in the furnace (Buhre et al. 2005).

In pressurized (i.e., "second generation") oxycombustion systems, the flue gas dew point is elevated such that the latent heat of condensation is upgraded to around 40 °C, making it valuable in the steam cycle. Hagi et al. (2014b) compared two pressurized processes, the flameless ISOTHERM technology (Malavasi and Rossetti 2013) and the staged pressurized oxycombustion (SPOC) process, against conservative and optimized ambient pressure oxycombustion processes. They found that the (absolute) energy penalties relative to an ultrasupercritical reference PC power plant without carbon capture were 7.7 $\%_{LHV}$ for ISOTHERM, 3.8 $\%_{LHV}$ for SPOC, and 10 and 7 $\%_{LHV}$, respectively, for conservative and optimized ambient pressure oxycombustion processes. Furthermore, they used exergy analysis to identified three main reasons for the superior energetic performance of the SPOC process. First, pressurized operation requires sealed combustors with no air leakage. As a result, the flue gas entering the CO_2 polishing unit (CPU) is 88–89 mol% CO_2, versus 78 mol% for ambient pressure oxycombustion, which allows for the CPU to be significantly smaller. Second, a staged configuration with four reactors is used to help regulate the flame temperature, which reduces the amount of flue gas recycle and decreases the auxiliary power demand by 35 %. Finally, the latent heat of condensation for the flue gas is utilized in the pressurized process. They anticipate further reduction in the energy penalty for the SPOC process with utilization of the waste heat from air compression. It is important to note this analysis focuses on

energetic metrics, and does not account for the required technology development or additional capital costs associated with pressurized oxycombustion (especially development of the sealed, staged reactor).[3] In other studies, several research groups attempted to optimize the operating pressure in the boiler by trading off heat recovered from the flue gas with the compression load for feed streams and/or exploring utilization of waste heat from the compressor(s) in the steam cycle (Gopan et al. 2014, 2015; Hong et al. 2009, 2010; Soundararajan and Gundersen 2013; Soundararajan et al. 2014b; Zebian and Mitsos 2013).

1.1.3 Power Island

The power island contains the steam turbines, where the hot working fluid (e.g., steam) is expanded to produce useful shaft work, which is converted into electricity. Four key principles are used in steam cycle design to increase thermal efficiency: heat recovery, steam reheat, steam extraction and supercritical operation. Regarding the first principle, there is significant thermal energy in the hot flue gas exiting the boiler (typically at 1300 °C). This heat is used to preheat water and air streams before they enter the furnace (in the economizer and air heaters, respectively). Heat from the flue gas is also used to reheat steam in the reheaters after the steam has partially expanded in the high pressure turbine. This is important, because as steam expands it cools, and reheat allows for more useful work to be extracted before it condenses. Up to two reheaters are common in modern steam cycles. In addition, small amounts of steam are extracted in several places through the steam cycle, which are used to preheat water before it goes into the economizer and furnace. The numbers of feedwater preheaters are determined by balancing improved theoretical thermodynamic performance (which prefers an infinite number of exchangers) against additional costs and pressure drops, and turbine design considerations. In modern plants, up to seven feed water preheaters are common.

Power plants are typically classified as subcritical or supercritical, depending on the condition of the steam produced in the boiler relative to the critical point of water (374 °C, 22 MPa). In subcritical steam cycles, the latent heat of vaporization supplied by the furnace cannot be converted to useful work, whereas almost all of the heat supplied above the critical point may be converted (The Babcock and Wilcox Company 2005a). Recent advances in materials have even enabled ultra-supercritical cycles, with maximum steam temperatures between 700 and 760 °C. For greenfield oxycombustion processes, supercritical or ultra-critical cycles are preferred due to their increased efficiency. In retrofit applications, the operating conditions of the steam are typically determined by the existing equipment. Finally,

[3]Another source of skepticism regarding pressurized oxycombustion comes from the small predicted energy penalties. Analysis of the theoretical minimum separation energy for pressurized oxycombustion would help address these concerns.

the power island also includes a condenser, which operates under vacuum (determined by the temperature of the available cooling water), and pumps to pressurize the liquid phase working fluid (water) to its operating pressure.

Unsurprisingly, many studies of oxy-fired processes (and CO_2 capture systems in general) mention heat integration as a promising strategy to increase overall plant efficiency (Harkin et al. 2010; Higginbotham et al. 2011; Jordal et al. 2004; Mc Donald et al. 2011; Perrin et al. 2014; Romeo et al. 2009; Tranier et al. 2011). The idea is to utilize low-grade waste heat, especially from compressors, in the steam cycle. This reduces the amount of low-pressure steam extracted for feedwater preheating, and/or increases the temperature of water/steam entering the boiler, both of which increase the overall power output of the turbines. A few of these studies have investigated the trade-off between compression work and generating heat for integration, with isothermal and adiabatic compression as the two extremes. Pinch analysis was used to estimate absolute thermal efficiency increases of 0.4–0.5 $\%_{HHV}$-points and 1.1–1.5 $\%_{LHV}$-points with proper integration of waste heat from compression for the ASU (Fu and Gundersen 2012a) and CPU (Soundararajan et al. 2014a), respectively. More recently, mixed integer nonlinear programming was used to optimize the compression ratio and intercooling temperature of the compression train for an ASU while accounting for heat integration with the steam cycle, resulting in a 0.5–0.6 $\%_{HHV}$-points (absolute) improvement in thermal efficiency (Fu et al. 2015a). These results are expected, as intercooling for compression accounts for approximately 18 % of the exergy destruction in the ASU (Fu and Gundersen 2012a), and the ASU contributes to around 2/3 of the energy penalty for oxycombustion. All of these cited studies assume the steam cycle remains unchanged, and they must approximate the value (i.e., ability to produce work) of extracted steam. Their methodologies cannot be used to rigorously optimize the flowrates, temperatures and pressures of extracted steam in order to utilize additional waste heat. Furthermore, these studies focus on energy and do not consider capital costs, which may be very high for complex compression schemes.

1.1.4 Pollution Controls and Flue Gas Recycle Strategies

After being cooled in the economizer and air heaters, particulate matter is removed from the flue gas using electrostatic precipitators (ESP) or fabric filters (i.e., baghouses). Next, SO_x is removed in flue gas desulfurization (FGD) units using lime ($CaCO_3$) to produce gypsum ($CaSO_4$). Finally, additional pollutants are removed using a direct contact cooler polishing scrubber (DCCPS). Typically, two flue gas recycle streams are used in oxycombustion processes. The primary recycle occurs after sulfur and water removal, and is used to dry and carry coal into the furnace. Placement of the secondary recycles (see Fig. 1), which is analogous to secondary or over-fired air in an air-fired system, depends on the sulfur content of the coal. Hagi et al. (2013, 2014a) discuss the efficiency implications of various locations of the secondary recycle, including before the ESP (which requires the addition of a high temperature particulate removal system for the secondary recycle), between

the ESP (or filter) and FGD, between the FGD and DCCPS, and after the DCCPS (same location as the primary recycle). In their study, hot flue gas recycle (Case A in Fig. 1) results in the best performance, with only a 7 $\%_{LHV}$-point (absolute) thermal efficiency penalty using an advanced ASU and CPU, whereas "standard" oxycombustion has a 10 $\%_{LHV}$-point (absolute) penalty. Due to corrosion concerns, the secondary recycle must be placed after the FGD when using high sulfur content coals. In all of these cases, it is important that the flue gas temperature remains well above its acid dew point before SO_x removal (see Okkes and Badge (1987) for a predictive correlation), which prevents acid condensation. Alternately with the CO_2RE (Combustion in O_2 Rich Environment) configuration (Zanganeh and Shafeen 2007), air is used to carry pulverized coal into the boiler, which eliminates the costs associated with primary recycle with an air-tight coal mill. In many oxy-fired systems, NO_x emissions are mitigated using low-NO_x burners, selective catalyst reduction (SCR) systems, and air-staging (over-fired air), which promotes the destruction of NO_x in high temperature/fuel rich areas of the furnace (Seltzer et al. 2010). This reburn phenomenon is important, and several research teams have experimentally characterized the formation of hazardous air pollutants in oxy-fired boilers with flue gas recycle (Croiset and Thambimuthu 2001; Fernández-Miranda et al. 2014; Zhuang and Pavlish 2012).

1.1.5 CO_2 Processing (Polishing) Unit

In the final section of an oxycombustion power system, the CO_2-rich flue gas is further dehydrated, purified and compressed to high pressures for transportation to utilization and/or sequestration sites. Typically, coal boilers operate under slight vacuum to minimize the escape of coal/ash particles and hot flue gases into the ambient through inevitable leaks. As a result, air leaks in the oxycombustion process introduce additional Ar, N_2, and O_2 into the flue gas. Although the final CO_2 product specifications are specific to the site and/or utilization application, 96 mol% CO_2 or greater purity is generally preferred (Pipitone and Bolland 2009; Toftegaard et al. 2010). However, flue gas after the pollution controls (but before water knock-out) typically contains 58–70 % CO_2, 8–18 % N_2, 2–5 % O_2, 2–3 % Ar, 15–17 % H_2O (all mole basis) (Darde et al. 2009). Several methods, including absorption, adsorption, membrane separation and cryogenic distillation, have been proposed to remove the remaining "non-condensable gases" (Songolzadeh et al. 2014). Due to the large difference in boiling points between CO_2 and these gases, sub-ambient flash separation is one of the most promising commercially available technologies.

Several researchers have studied CO_2 purification processes in oxy-combustion power plants. Pipitone and Bolland (2009) compared energy requirements for sub-ambient separation and compression of boiler flue gas for natural gas and pulverized coal fuels in three cases: no separation (baseline), a two flash unit separation system and a distillation column separation system. Similarly, Besong et al. (2013) compared sub-ambient separation using single flash and three flash

units with respect to various CO_2 recovery, purity and power requirements for the process, using feed gas compositions and thermodynamic data from Darde et al. (2009). They also analyzed the sensitivity of CO_2 recovery and purity to perturbations of the inlet pressure specifications for the flash vessels. Zanganeh et al. (2009) investigated the behavior of three flash unit pilot plant systems for a wide spectrum of inlet CO_2 concentrations. All of these studies found similar results: separations with more equilibrium units require less energy but have greater capital costs. A few integrated systems for CO_2 purification and hazardous air pollutant removal have also been proposed. For example, White et al. (2009) present experimental data for a system that removes NO_x and SO_x as HNO_3 and H_2SO_4 in condensate during compression before the CPU.

1.2 Key Design Decisions and Assumptions

For over 15 years, researchers have acknowledged the importance of plant-wide integration of the oxycombustion process to reduce the energy penalty and ensure competitiveness with other CO_2 capture pathways (e.g., post-combustion). Several key design questions are highlighted in literature:

1. What is the trade-off between the sizes of the ASU and CPU, and how does the optimum depend on the air leakage rate (Besong et al. 2013; Li et al. 2013, 2009)? For example, if the O_2 purity from the ASU is decreased, the size (and energy requirement) of the CPU will need to increase to maintain specified CO_2 product purity and recovery targets.
2. What is the best way to utilize low-grade waste heat from air and flue gas compression (ASU and CPU, respectively) (Darde et al. 2009; Fu et al. 2015a; Fu and Gundersen 2012c; Higginbotham et al. 2011; Romeo et al. 2009; Skorek Osikowska et al. 2013; Soundararajan et al. 2014a; Tranier et al. 2011)? This may require adjustment of the compression strategy and manipulation of steam extraction flowrates in the power island.
3. What is the best flue gas recycle strategy (out of several configuration options) and what is the optimal flue gas recycle rate (Hagi et al. 2013, 2014a)?
4. What are the benefits, costs, and risks of pressurized oxycombustion? Furthermore, how do performance metrics for *optimized* ambient and pressurized oxycombustion processes compare with each other and the analogous PC power plant without carbon capture (Gopan et al. 2014, 2015; Hagi et al. 2014b; Hong et al. 2009, 2010; Soundararajan and Gundersen 2013; Zebian et al. 2012; Zebian and Mitsos 2013)?

Addressing these design decisions requires system wide analysis of oxycombustion process. For example, manipulation of the O_2 purity requirements for the ASU not only impacts the composition of flue gas entering the CPU, but also impacts combustion efficiency in the furnace and the design of the steam cycle. Thus, these design questions are inter-related and should be addressed simultaneously. Furthermore,

the answers (i.e., optimal decisions) are highly sensitive to assumptions, especially the air leak rate. In fact, Yan et al. (2015) found the energy penalty for an oxycombustion process to be most sensitive to the air leakage assumption out of the eight parameters studied (air leakage rate, oxygen purity from the ASU, excess O_2 in the furnace, flue gas condensation temperature, gas leakage rate in the gas preheater, fan pressure rise, flue gas recirculation rate, and O_2 concentration after the primary fan). Air leakage rates are typically around 3 % (total flue gas mass flowrate) for new plants, and 8–16 % for older boilers (see Toftegaard et al. (2010) for a literature survey).

1.3 Review of Systems Engineering Literature

Given the complex trade-offs between each subsystem previously described, *holistic* analysis of the oxycombustion process is imperative. A majority of researchers use process simulators (e.g., Aspen Plus) to do this and to conduct techno-economic feasibility studies (Fu and Gundersen 2012b; Skorek-Osikowska et al. 2013; Xiong et al. 2009, 2012). This has led to a consensus that oxycombustion has around an 8–10 %-point absolute thermal efficiency penalty (both HHV and LHV bases) compared to a similar PC power plant without carbon capture (Andersson and Johnsson 2006; Buhre et al. 2005; Fu and Gundersen 2010). Several have concluded this penalty may be reduced to 6–7 %-points (absolute) through optimization of sub-ambient systems (Tranier et al. 2011), consideration of alternate (i.e., hot) flue gas recycle strategies (Hagi et al. 2014a), and heat integration to utilize waste heat from compression (Soundararajan et al. 2014a). For convenience, the reported net thermal efficiencies for these studies are summarized in Table 1. Furthermore, several research groups have used single variable perturbations to explore the sensitivity of the oxycombustion process to key assumptions and design variables, which aids in "by hand" process refinement and "optimization" (Besong et al. 2013; Hong et al. 2010; Skorek-Osikowska et al. 2013; Soundararajan et al. 2014a; Xiong et al. 2011b; Yan et al. 2015). Similarly, exergy analysis is employed in several studies to identify the sources of inefficiencies, which helps focus process refinement efforts (Fu and Gundersen 2012c, 2013; Hagi et al. 2013, 2014a; Xiong et al. 2011a). Pinch analysis is also used to estimate the potential for integration of waste heat into the steam cycle (Fu and Gundersen 2010; Harkin et al. 2010; Soundararajan et al. 2014a). However, advanced flowsheet optimization tools have seldom been applied to simultaneously optimize all of the subsystems in an oxycombustion process. The exception is the work of Zebian et al. (2012, 2013), in which 13 design variables are optimized for in a pressurized oxycombustion process.

Finally, it is important to acknowledge substantial contributions from industry, including Air Liquide (Darde et al. 2009; Kupila et al. 2011; McDonald et al. 2011; Perrin et al. 2014; Tranier et al. 2011), Air Products (Higginbotham et al. 2011; Smith and Klosek 2001; White et al. 2009), Babcock and Wilcox (McCauley et al. 2009; McDonald et al. 2011), Foster Wheeler (Hultgren et al. 2014; Myöhänen et al. 2009),

and Praxair (Shah et al. 2011). Furthermore, the National Energy Technology Laboratory has published a comprehensive baseline report which explores several oxycombustion configurations with detailed cost analysis (Matuszewski 2010).

2 Proposed Equation-Based Optimization Methodology

Zebian et al. (2012, 2013) found that multivariable optimization reveals complex trade-offs between key oxycombustion design variables, which are missed in single variable sensitivity studies. They employ sequential quadratic programming (SQP) techniques available in Aspen Plus, however, which are restricted to limited problem sizes.[4] For example, Aspen Plus does not allow the number of trays in a distillation column (e.g., the ASU) to be considered as a decision variable. Thus, to truly optimize the oxycombustion procession and answer the four previously stated design questions, it is necessary to consider new tools for large-scale equation-based flowsheet optimization.

Over the past 4 years, the authors have developed a new equation-based framework for flowsheet optimization, with the ultimate goal of simultaneously optimizing an entire oxycombustion process. In contrast with other equation-oriented tools (e.g., Aspen Custom Modeler, gPROMs), emphasis is placed on exploiting open formulations with exact first and second derivatives (efficiently tabulated via automatic differentiation). These exact derivatives enable the use of state-of-the-art large-scale nonlinear programming algorithms that are capable of considering optimization problems with 100,000 + variables and constraints. The new framework has five key parts:

1. Embedded cubic equation of state (e.g., Peng–Robinson) thermodynamics with complementary constraints to accommodate vanishing and reappearing phases (without using integer variables for each equilibrium calculation) (Dowling et al. 2015; Dowling and Biegler 2015; Kamath et al. 2010)
2. Simultaneous heat integration and process optimization using embedded pinch equation to rigorously calculate minimum utility requirements (Dowling and Biegler 2015a, b; Duran and Grossmann 1986; Kamath et al. 2012)
3. New distillation models that allows for feed tray location and column size optimization using tray bypass streams instead of integer variables (Dowling and Biegler 2014, 2015a, b)

[4]This limitation is primarily due to the nature of sequential-modular process simulators, such as Aspen Plus. Derivatives for optimization must be calculated with perturbations, which are noisy due to the iteration tolerances for convergence of individual unit models and flash calculations. See Biegler et al. (1997) for additional discussion. For equation-oriented process simulators, it is unclear if exact first derivatives are available with thermodynamic function evaluations and how non-ideal phase equilibrium calculations are formulated to ensure continuous derivatives across phase changes.

4. Equation-based steam cycle models and a hybrid 1D/3D boiler model (Dowling et al. 2014a; Ma et al. 2014, 2016)
5. Trust region optimization algorithm to consider subsystem models without exact derivatives (e.g., CFD boiler simulations) while maintaining the computational benefits of open equation models for a majority of the flowsheet (Biegler et al. 2014; Dowling et al. 2014a; Eason and Biegler 2015).

Regarding the first element of the framework, all of the thermodynamic calculations are expressed as algebraic equations, which allows for exact derivatives to be efficiently calculated via automatic differentiation. The streams in the air separation subsystems of the oxycombustion process are modeled using a cubic equation of state (e.g., Peng–Robinson), denoted $f(Z) = 0$. As the name implies, up to three real roots for Z, the compressibility factor, may exist.[5] Using the framework proposed by Kamath et al. (2010), each root is classified as vapor or liquid by examining the first and second derivatives of f with respect to Z.[6] Most importantly, this allows for the phase selection rules to be expressed as inequality constraints. For general flowsheet optimization problems, it is desirable for the number of phases to change as different temperatures, pressure and compositions are explored in the design space. Vanishing and reappearing phases introduces a nonsmooth feature in phase calculation problems, which are modeled using complementarity constraints.[7]

Let (L,x), (V,y), and (F,z^f) represent flowrate-composition pairs for the liquid outlet, vapor outlet and feed streams for a flash calculation, and Z_L and Z_V represent the compressibility factors of the outlet streams. The slack variables σ^L and σ^V are used to allow for either the vapor or liquid phase to vanish. A single flash calculation is modeled as follows:

$$L + V = F \tag{1a}$$

$$Lx_i + Vy_i = Fz_i^f, \quad \forall i \in \{Components\} \tag{1b}$$

$$H^l L + H^v V + H^f F \tag{1c}$$

$$f(Z_L) = 0,\ f(Z_V) = 0 \tag{1d}$$

$$f'(Z_L) \geq 0,\ f'(Z_V) \geq 0 \tag{1e}$$

$$f''(Z_L) \leq M\sigma^L,\ f''(Z_V) \geq -M\sigma^V \tag{1f}$$

[5]When three real roots exist, the smallest corresponds to the liquid phase; the largest corresponds to the vapor phase and the middle root is considered spurious.

[6]Because $f(Z)$ is a cubic polynomial with respect to Z, calculating $f'(Z)$ and $f''(Z)$ is straightforward.

[7]In essence, $0 \leq x \perp y \geq 0$ (x complements y) means either $x = 0$ and/or $y = 0$ (and $x \geq 0$, $y \geq 0$).

$$-\sigma^L \leq \beta - 1 \leq \sigma^V \tag{1g}$$

$$y_i = \beta x_i \frac{\phi^l}{\phi^v}, \quad \forall i \in \{Components\} \tag{1h}$$

$$\sum_{i \in \{Components\}} (y_i - x_i) = 0 \tag{1i}$$

$$0 \leq V \perp \sigma^V \geq 0, \; 0 \leq L \perp \sigma^L \geq 0 \tag{1j}$$

where H and ϕ are molar enthalpies and fugacity coefficients, which are calculated using a cubic equation of state model, and depend on temperature, pressure and stream composition. Equations (1a)–(1c) are standard mass and energy balances. Equation (1d) is the cubic equation of state for both outlet streams,[8] and (1e) eliminates the spurious middle root. Equation (1f) distinguishes between the vapor and liquid roots, and is relaxed using slack variables σ^L or σ^V. Similarly, the vapor–liquid equilibrium expression, (1h), is relaxed with slack variable β. Equation (1i) is the summation equation in the Rachford-Rice form. The complementarity constraints, (1j), require $L = 0$ or $\sigma^L = 0$ and $V = 0$ or $\sigma^V = 0$. Thus, if both phases exist, then $\sigma^L = \sigma^V = 0 \Longrightarrow \beta = 1$ and the standard flash equations are obtained. If the liquid phase vanishes, then $L = 0 \Longrightarrow \sigma^L \geq 0 \Longrightarrow \beta \leq 1$, and as consequence (1h) along with part of (1f) may be relaxed. Formulation (1) is analogous to the first order optimality conditions for the Gibbs free energy minimization of the two outlet stream system, and β is related to the Lagrange multipliers for $V \geq 0$ and $L \geq 0$, as derived in Biegler (2010). More recently, formulation (1) has been extended to correctly classify supercritical streams and check for phase stability (Dowling et al. 2015). The latter is important, as (1) may misclassify mixtures that are in the two phase region as single phase. This has been observed as a means to "cheat" thermodynamic calculations to improve heat integration (Dowling and Biegler 2015; Dowling et al. 2015).

In summary, with thermodynamic phase calculations formulated as (1), exact derivatives are available for optimization. In contrast many commercial flowsheet optimization software packages rely on "function calls" to evaluate thermodynamic properties (e.g., Multiflash in gPROMs), which do not offer a straightforward way to evaluate first derivatives. Moreover, when the number of phases changes, the phase equilibrium calculation becomes nonsmooth, which is modeled with complementarity constraints in (1).[9]

[8]The cubic equation of state model is also used to calculate the phase and enthalpy of the feed stream(s), but is omitted from (1) for brevity.

[9]Most nonlinear optimization algorithms rely on continuity assumptions for first (and possibly second) derivatives. When these assumptions are violated, solver performance typically deteriorates. Development of large-scale nonsmooth nonlinear optimization algorithms remains an open research challenge.

The remainder of this chapter describes application of the new framework to three parts of the oxycombustion flowsheet (the air separation unit, the CO_2 polishing unit and the steam cycle) and discusses future work to simultaneously optimize an entire oxycombustion process. Specific mathematical formulations of the models, omitted from this publication for brevity, are presented in several other papers (Dowling et al. 2014b, 2015; Dowling and Biegler 2013, 2014, 2015; Dowling et al. 2014a, c; Ma et al. 2014, 2015) and a PhD dissertation (Dowling 2015).

3 Air Separation Unit Optimization Case Study

Using the proposed framework, an Air Separation Unit (ASU) and accompanying multistream heat exchangers were optimized for oxycombustion processes to produce 90–97 mol% O_2 product. Instead of specifying a process configuration a priori, a superstructure that encompasses numerous possible recycle strategies is used (Fig. 3). In addition to manipulating the flowrate, temperature and pressure of each stream, the optimizer selects the best system configuration and manipulates the size of both the low and high pressure distillation columns by solving the following nonlinear program (NLP):

$$\begin{aligned} \min \quad & \text{Specific Compression Work} + \rho_s Q_s + \rho_w Q_w \\ & + \rho_\perp (\text{Complementarity Violations}) \\ \text{s.t.} \quad & \text{Double column superstructure (Fig. 3)} \\ & \text{Peng}-\text{Robinson Thermodynamics, including (1)} \\ & \text{MESH with tray Bypass Distillation Model} \\ & \text{Pinch-Location Heat Integration Equations} \\ & \text{Ideal Gas Compression Work Formula} \\ & \text{Additional Equilibrium unit Operation Models} \\ & O_2 \text{ purity} \geq \text{ specification} \end{aligned} \tag{2}$$

where ρ_s, ρ_w, and $\rho_\perp$ are sufficiently large positive weights on terms that should be forced to zero at the solution.[10] For the ASU design problems reported in this chapter, $\rho_s \geq 1$ and $\rho_w \geq 5$ works well, and were determined using trial and error.

The selected ASU structure is shown in Fig. 4, and is very similar to the Linde double column design with a Lachmann stream (e.g., feed to the low-pressure column). The multistream heat exchanger is modeled using a modification of the Duran and Grossmann (1986) formulation, which is a collection of nonlinear equations to calculate the minimum heating and cooling utilities, Q_s and Q_w, respectively. These terms are heavily penalized in the objective function, as shown

[10]In problems (2), (3) and (5), the complementarity constraints, such as (1j), are accommodated using the penalty formulation. See Biegler (2010) for an overview of theoretical aspects of optimization problems with complementarity constraints, along with applications in chemical engineering and solution techniques.

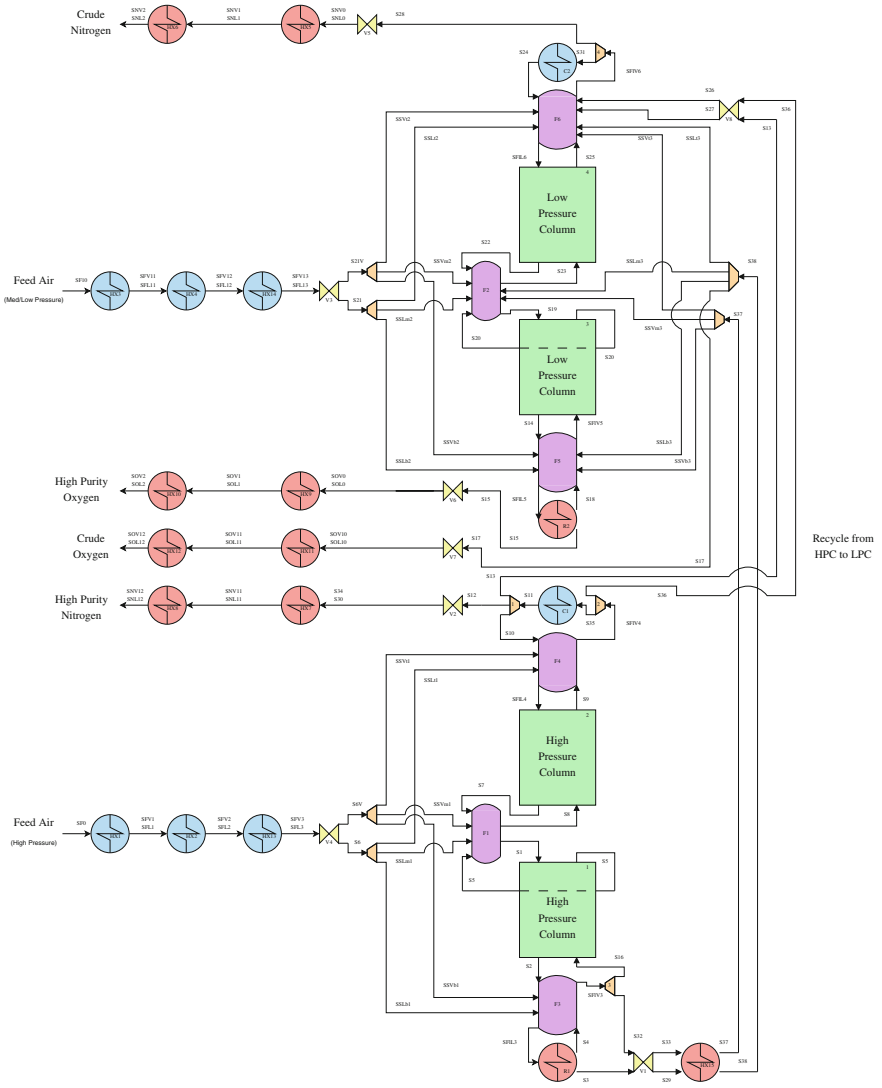


Fig. 3 Air is pressurized, cooled and fed into the *top* (low pressure) and *bottom* (high pressure) distillation column. The *bottom* products and a fraction of the distillate from the higher pressure column are recycled to the low-pressure column. Several different recycle configurations are encoded in the double column ASU superstructure. The *bottom* product and distillate of the low-pressure column are the O_2 and N_2 products, respectively. Heat exchanger halves are individually modeled, where *red* units provide heat and *blue* units require heat. The halves are implicitly paired together by the pinch-based heat integration model (Duran and Grossmann 1986). The number of theoretical trays in the distillation columns (*green* units) are decision variables, and modeled using only continuous variables (Dowling and Biegler 2014, 2015a, b). The *orange* shapes are splitters, and the feed air compressors are not shown, but are included in (2) using an ideal gas formula

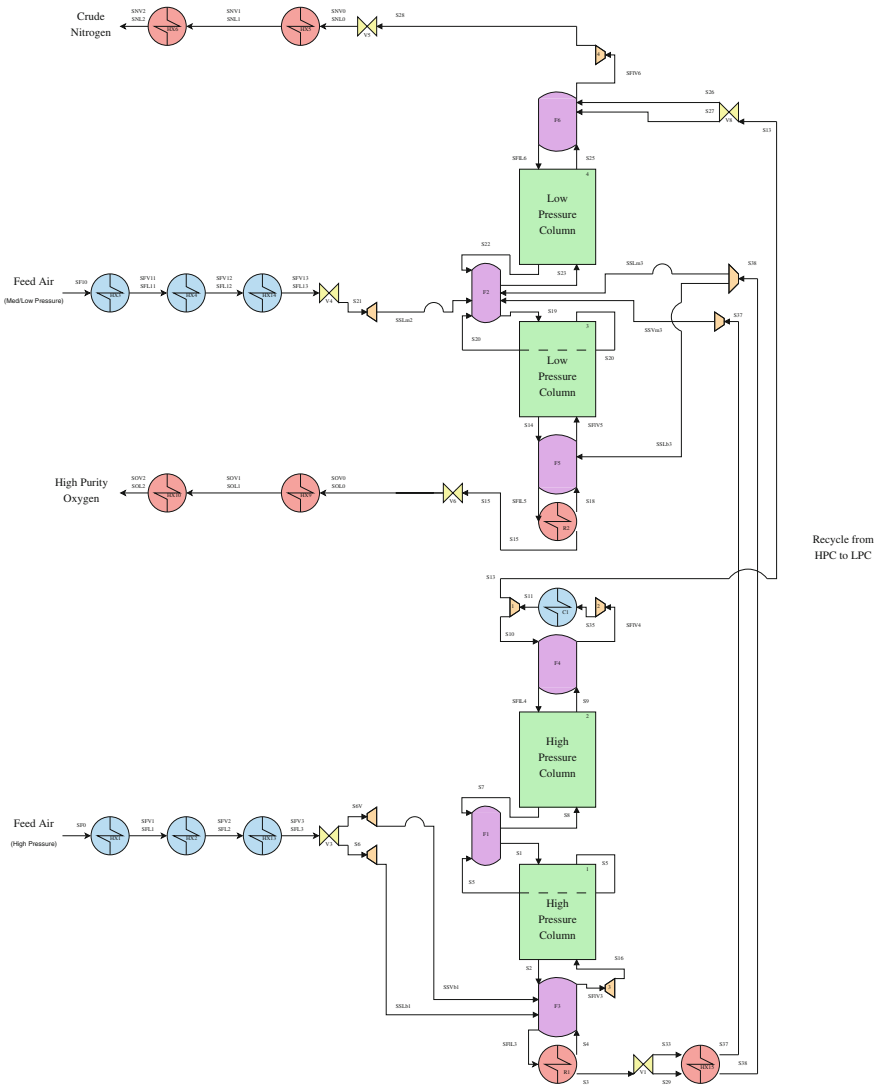


Fig. 4 The optimal ASU flowsheet topology, determined by solving (2) with $\Delta T_{\min} = 1.5$ °C and 95 mol% O_2 product, has several notable features: air is fed into the low-pressure columns (also known as a Lachman stream); some of the liquid from the high pressure condenser is expanded and fed into the top of the low-pressure column, which provides reflux without the need for a second condenser; high purity N_2 is not extracted from the high pressure column. Streams not shown in this figure but included in Fig. 3 have zero flowrates at the optimal solution

in (2), and thus at the optimal solution, $Q_s = Q_w = 0$, which indicates the multistream heat exchanger is self integrated. This approach, however, requires specification of a minimum temperature difference, $\Delta T_{\min}$, which implicitly fixes some of the multistream heat exchanger's capital costs. The optimization procedure was

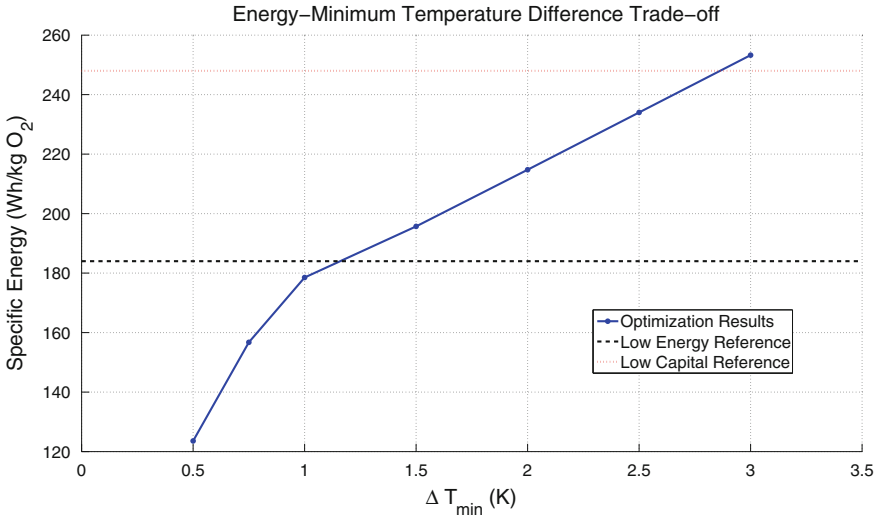


Fig. 5 Comparison of minimized specific energy usage for the ASU versus $\Delta T_{\min}$ specification, and two designs from Matuszewski (2010). The *blue curve* is generated by adjusting $\Delta T_{\min}$ and resolving (2). Between $\Delta T_{\min}$ of 1 and 3 K, the energy usage varies linearly with the temperature difference in the multistream heat exchanger. Below $\Delta T_{\min}$ = 1 K decreases in the temperature difference provide progressively larger reductions in energy usage (a nonlinear relationship). It is likely impractical, however, to construct multistream heat exchangers with $\Delta T_{\min}$ < 1 K given the large area requirements (thus high costs) and other design complexities (e.g., pressure drops)

repeated for $\Delta T_{\min}$ values ranging from 1 to 3 °C and these results are compared against two designs from American Air Liquide (Matuszewski 2010) in Fig. 5. Unfortunately, none of the necessary assumptions for the American Air Liquide designs are published (including $\Delta T_{\min}$), which prohibits a direct comparison. Nevertheless, Fig. 5 shows how the specific energy consumption for the optimized results compares with the low energy and low capital industry designs for $\Delta T_{\min}$ values of 1.3 and 2.8 °C, respectively, which are very reasonable for cryogenic systems. For 95 mol% O_2 and $\Delta T_{\min}$ of 1.5 °C, the optimized ASU requires 196 Wh/kg of pure O_2. Furthermore, development of the detailed ASU model is important for the final goal: to simultaneously optimize an entire oxycombustion process and to study sizing trade-offs with the CPU and the potential to integrate waste heat from compression.

4 CO_2 Processing Unit Optimization Case Study

In the CO_2 processing unit (CPU), "non-condensible" gases (Ar, N_2, O_2) are separated from CO_2 by liquefaction, and the purified CO_2 stream is compressed/pumped to 150 bar for utilization and/or sequestration. The selected CPU process is based on Fu and Gundersen (2012b) and consists of a feed compression train,

mulit-stream heat exchanger, two flash vessels, and a CO_2 product compression train, as shown in Fig. 6.

The energy requirement of the system for a fixed flue gas composition is optimized by solving the following NLP:

$$\begin{aligned} \min \quad & \text{Specific Compression Work} + \rho_s Q_s + \rho_w^{\text{MHEX}} Q_w^{\text{MHEX}} + \hat{\rho}_w \hat{Q}_w \\ & + \rho_{\perp}(\text{Complementarity Violations}) \\ \text{s.t.} \quad & \text{Peng}-\text{Robinson Thermodynamics, including (1)} \\ & \text{Pinch-Location Heat Integration Equations} \\ & \text{Compressor Model with Fixed Isentropic Efficiency} \\ & \text{Additional Equilibrium Unit Operation Models} \end{aligned} \tag{3}$$

where Q_w^{MHEX} is the cooling demand for the multistream heat exchanger, and $\hat{Q}_w$ is the cooling utility serviced by chilled sea water. Similar to the ASU, ρ_s and ρ_w^{MHEX} are large weights such that $Q_s = Q_w^{\text{MHEX}} = 0$ at the optimum, indicating the multistream heat exchanger is completely self heat integrated. Furthermore, the above optimization problem includes a modification of the original cubic equation of state phase classification constraints from Kamath et al. (2010) to identify the correct phase above the critical pressure (Dowling et al. 2015). Repeated solution of (3) for different CO_2 purity and recovery specifications leads to the Pareto optimal frontiers shown in Fig. 7. Interestingly, as CO_2 purity increases, the specific energy decreases, which initially is counter-intuitive. The reason for this trend is that as CO_2 purity increases, the total amount of gas that must be compressed from the flash vessel operating pressure (around 20–40 bar) to the final product pressure (150 bar) decreases. Likewise, the specific separation work per mass of CO_2 captured is higher for low recovery systems (with constant purity). Thus, for high recovery systems, the energy penalty from additional CO_2 product compression has less impact than the larger CO_2 mass flowrate in the denominator for the specific energy calculation. Interestingly, increasing product purity has a more dramatic impact on energy requirements (objective function) when CO_2 product purity is greater than 96 mol%. Further analysis of the results reveals lower flash vessel operating pressures are required in this region, and more of the separation is due to partial liquefaction in the multistream heat exchanger. For CO_2 purities beyond the extremes shown in Fig. 7, optimization problem (3) is infeasible.[11] Additional CPU optimization results (including operating conditions for each stream) are presented in Dowling et al. (2015).

Investigation of a heat integrated ASU-CPU is ongoing using the optimization formulation. First, the CPU was re-optimized with incremental amounts of waste cooling available from the N_2 product stream from the ASU. At best, for each single unit of cooling (energy), the shaft work requirement for the CPU is reduced by 0.6

[11]Problem (3) is infeasbile for the following CO_2 recovery-purity combinations: (94.6 % recovery, 97.0 mol% purity), (94.6 % recovery, 97.5 mol% purity), (92.0 % recovery, 97.5 mol% purity), and (90.0 mol% recovery, 98.0 mol% purity).

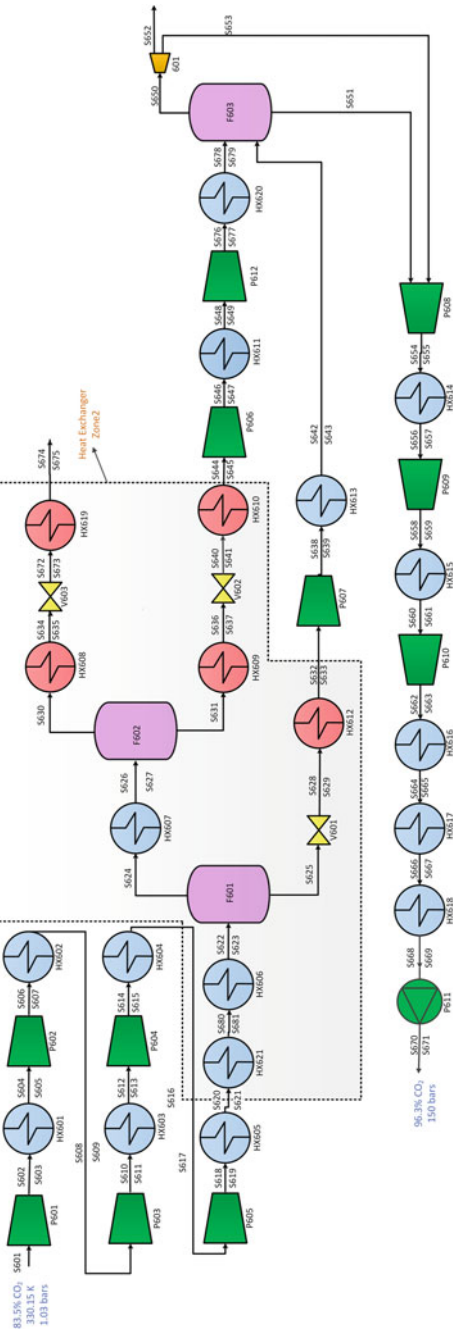


Fig. 6 The CO_2 processing unit (CPU) flowsheet includes pre- and post-separation compression trains, two flash vessels and a multistream heat exchanger. The model of Duran and Grossmann (1986) is extended to include multiple exclusive heat integration zones, with the heat exchanger halves shown in the shaded region corresponding to the multistream heat exchanger. Compressors and pumps are colored *green*, and the single splitter is shaded *orange*

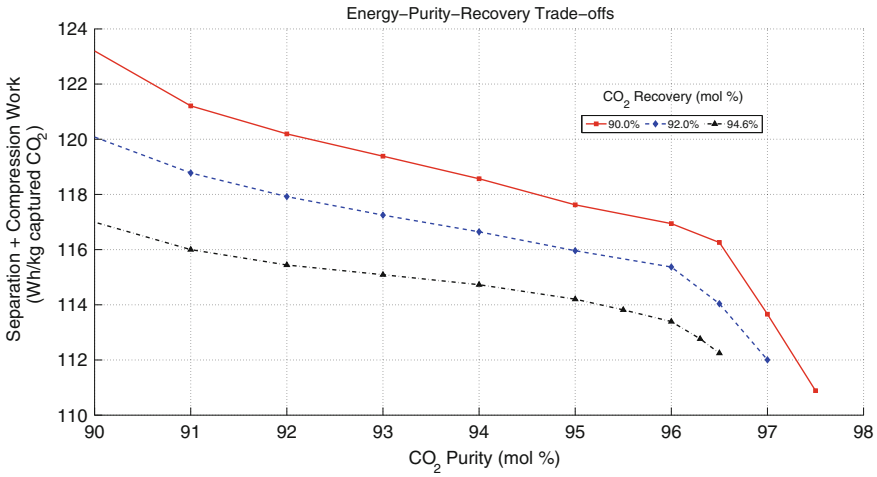


Fig. 7 The Pareto optimality surfaces show the trade-off between CO_2 recovery and purity (problem specifications) and specific separation plus compression energy (design objective function) in the CPU system. The constant recovery *curves* may not be extended to higher CO_2 purities, otherwise (3) becomes infeasible

units of energy. The less than 1:1 cooling-energy trade-off is a consequence of the assumed CPU topology; after separation, the liquid CO_2 product is vaporized to balance the multistream heat exchanger (e.g., provide cooling). Instead, it may be possible to couple refrigeration from the ASU with a modified CPU, in which the liquid CO_2 product can be pumped to the desired pressure, rather than compressed into the dense supercritical phase. A preliminary flowsheet illustrating this concept is shown in Fig. 8.

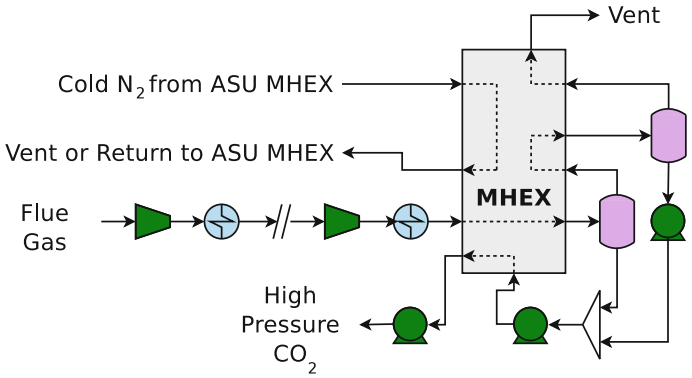


Fig. 8 In this alternate concept, product liquid N_2 from the ASU is used to liquify CO_2, which allows for CO_2 to be pumped instead of compressed. Evaluation of this design requires simultaneous optimization of the ASU and CPU because producing additional cooling in the ASU increases power requirements of the main air compressors

5 Steam Cycle Optimization

In most of the studies reviewed above, either a single subsystem of the oxycombustion process (e.g., ASU) is analyzed and improved, or simplified models (e.g., Gibbs reactors) are used for the boiler when simulating the entire plant (Pei et al. 2013; Xiong et al. 2011b). The latter is undesirable for several reasons. Most notably, heat transfer in the boiler is very complex and mainly radiative. Without a detailed furnace model, it is very difficult to predict both heat transfer into the working fluid at specific locations (e.g., through the boiler wall, in the superheater) and the furnace exit gas temperature (FEGT) with good accuracy. Furthermore, the combustion and heat transfer properties are dramatically different in CO_2/O_2 versus N_2/O_2 atmospheres; thus correlations for air-fired boilers are not appropriate. Finally, when thermal efficiency is maximized using an embedded Gibbs free energy minimization boiler model, the optimizer elects not to feed excess O_2 into the boiler because the model predicts near complete combustion (Dowling and Biegler 2013). This behavior, however, is not reflective of actual furnaces, in which excess oxygen is provided to reduce unburned carbon in the fly ash. As a consequence, modelers are forced to specify an amount of excess oxygen as part of the boiler model. Instead, computational fluid dynamics (CFD) models are preferred for detailed analysis of oxycombustion systems, especially because design correlations are not currently available for oxy-fired boilers, due to a limited number of industrial scale demonstration projects.

Embedding a CFD model into a process simulator (e.g., co-simulation) is a difficult task for several reasons. First, CFD simulations typically require hours to days to converge, and optimization requires repeated evaluation at different values for the design variables. Worse, if the furnace is located in a recycle loop (such as in the oxycombustion process), convergence of a process simulation in a sequential-modular mode requires repeated evaluation of the CFD model for a single instance of the design variables. Surrogate (e.g., reduced-order) models are a popular approach to address these concerns. For example, Edge et al. (2013, 2012) first fit a surrogate model for boiler performance based on CFD data and then used the surrogate model in a process simulator (gPROMs). Fei et al. (2015) use a similar approach to evaluate the potential for oxy-fired retrofits. Optimization with this workflow is flawed, however, as the surrogate models are not reevaluated or reconstructed at the optimal solution. In this section, a trust region optimization algorithm is introduced, which automatically updates the surrogate model during optimization while ensuring convergence. To reduce computational time, a hybrid 1D/3D boiler model is also introduced. The trust region strategy is used in conjunction with algebraic steam cycle models. Ultimately, this allows for optimization of all flow rates, temperatures and pressures in the steam cycle, including steam extraction, and enables rigorous integration of waste heat from compressors. In contrast, most other studies assume the steam cycle remains fixed and approximate the value of extracted low-pressure steam (Fu et al. 2015a; Fu and Gundersen 2012a; Soundararajan et al. 2014a).

5.1 Hybrid 1D/3D Boiler Model

High fidelity computational fluid dynamic simulations to resolve complex 3D flow and flame profiles are not necessary and are too computationally expensive for overall oxycombustion system design and optimization. In contrast, Gibbs minimization calculations oversimplify the reaction kinetics and bulk heat transfer in the boiler. This motivates the development of a medium fidelity hybrid 1D/3D boiler model with the following key characteristics:

1. Accurate prediction of key performance outputs, including heat transfer to the wall and secondary superheater, furnace exit gas temperature (FEGT), unburned carbon in fly ash, and flue gas composition.
2. Sensitivity of performance with respect to boiler geometry, burner configuration, and inlet air/recycle stream temperature and composition.
3. Low computational requirements (i.e., 1–2 minutes to converge for a single design point on a desktop computer).

In the hybrid zonal model, the boiler is discretized into a user specified number of vertical zones, as shown in Fig. 10. Each zone is modeled as a well-stirred reactor and the gas phase is assumed to be in chemical equilibrium. Reaction kinetics are only considered for char oxidation and gasification by CO_2 and O_2. Uniform properties (velocity, temperature, etc.) are assumed in each zone, except for radiation intensity, which is calculated using a discrete ordinate method on a 3D mesh consisting of unstructured hexagonal cells (see Fig. 10). Sample temperature profiles are shown in Fig. 9. As expected, the gas temperature is uniform in each zone, but the wall temperature varies in multiple dimensions, determined by the

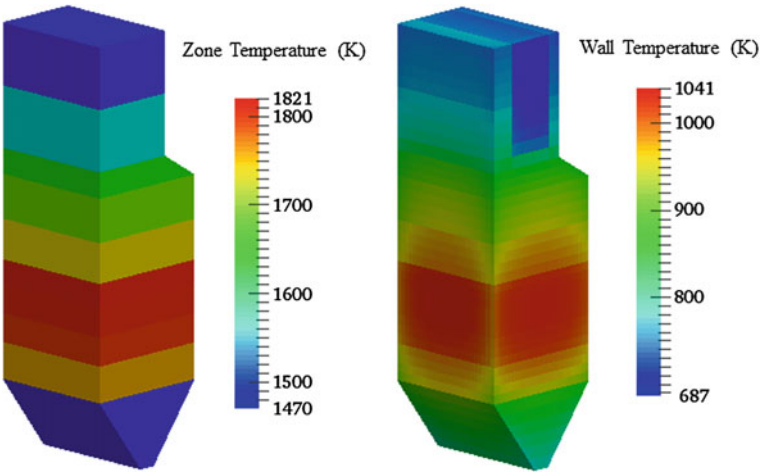


Fig. 9 Zone and wall temperatures from the hybrid boiler model. Notice the gas phase temperature is resolved in 1D and wall temperature is resolved in 3D

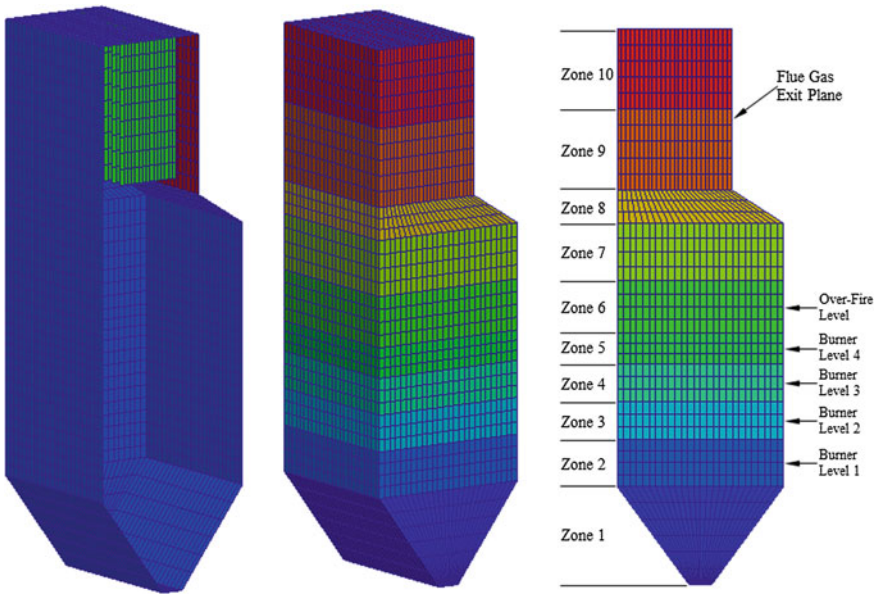


Fig. 10 1D zones and 3D mesh for the hybrid boiler model. In each of the 1D zones, the gas phase is approximated as a well-stired control volume. In contrast, a 3D mesh is used for radiative heat transfer calculations

3D radiative heat transfer calculations. The wall temperature profiles also include radiant superheaters at the top of the boiler. Additional reheaters, convective superheaters and the economizer are modeled as (convective) heat exchangers. Additional details for this hybrid boiler model are given in Ma et al. (2014, 2015).

The hybrid boiler model was validated against air-fired and oxy-fired 3D simulations for an existing utility boiler (PacifiCorp's Hunter 3 unit). Details of the CFD simulations are available in an NETL/Reaction Engineering International (REI) technical report (Adams et al. 2013). Comparisons between this hybrid boiler model and the CFD simulations are summarized in Table 2. As expected, the hybrid boiler model's heat losses to the enclosures are less than with the reference CFD model. This occurs because the hybrid boiler model does not include the section beyond the vertical nose plane. This part of the boiler is modeled using convective heat exchange equations.

5.2 Trust Region Optimization Algorithm

Although the hybrid boiler model offers great improvements balancing accuracy and computational expense, the procedural nature of the model is still unsuitable for integrated optimization approaches. As mentioned above, surrogate models are

Table 2 Comparison between hybrid boiler model and CFD simulations (Ma et al. 2015)

	Air-fired	
	Hybrid model	CFD model
Horiz. Nose flue gas exit temp.	1684 K	1674 K
Heat transfer to enclosure wall (Q^{wall})	436 MW	436 MW
Heat transfer to platen superheater (Q^{sup})	98.1 MW	98.5 MW
	Oxy-fired	
	Hybrid model	CFD model
Horiz. nose flue gas exit temp.	1655 K	1656 K
Heat transfer to enclosure wall (Q^{wall})	405 MW	407 MW
Heat transfer to platen superheater (Q^{sup})	112 MW	112 MW

often used to integrate simulations into optimization problems; however, the solution is not guaranteed to be accurate (Biegler et al. 2014). This often necessitates updating the surrogate during the optimization procedure. To rigorously integrate the boiler model into the equation-oriented optimization framework, one may use trust region concepts from nonlinear programming. The goal is to guarantee convergence to the true optimum of the underlying problem by building surrogate models that can be trusted in a local region of the design space.

The basic structure of trust region methods is as follows: At each iteration k, a surrogate model is constructed in a trust region around the current values of decision variables x_k. The optimization problem is solved while keeping the decision variables within the trust region. Then, the accuracy of the surrogate is checked at this subproblem solution. Depending on the success of this solution in making progress to the true solution, the step is either accepted or rejected and the trust region may be increased or decreased accordingly. Convergence is guaranteed by accuracy conditions imposed on the surrogate model. For a black-box function $d(x)$, and trust region radius Δ_k, the following accuracy condition is required for the surrogate model $r(x)$:

$$\|d(x) - r(x)\| \leq \kappa_f \Delta_k^2 \quad \|\nabla d(x) - \nabla r(x)\| \leq \kappa_g \Delta_k \tag{4}$$

This is the so called "fully linear property" from derivative free optimization (Conn et al. 2009). κ_f and κ_g are bounding constants that hold over all iterations k. Any choice of surrogate models may be used as long as it satisfies this condition. The fully linear property requires that as the trust region radius approaches zero, the surrogate model and its gradient approach their true values. Popular surrogate choices such as Kriging and polynomial interpolation are acceptable, given simple geometry assumptions on the sample set.

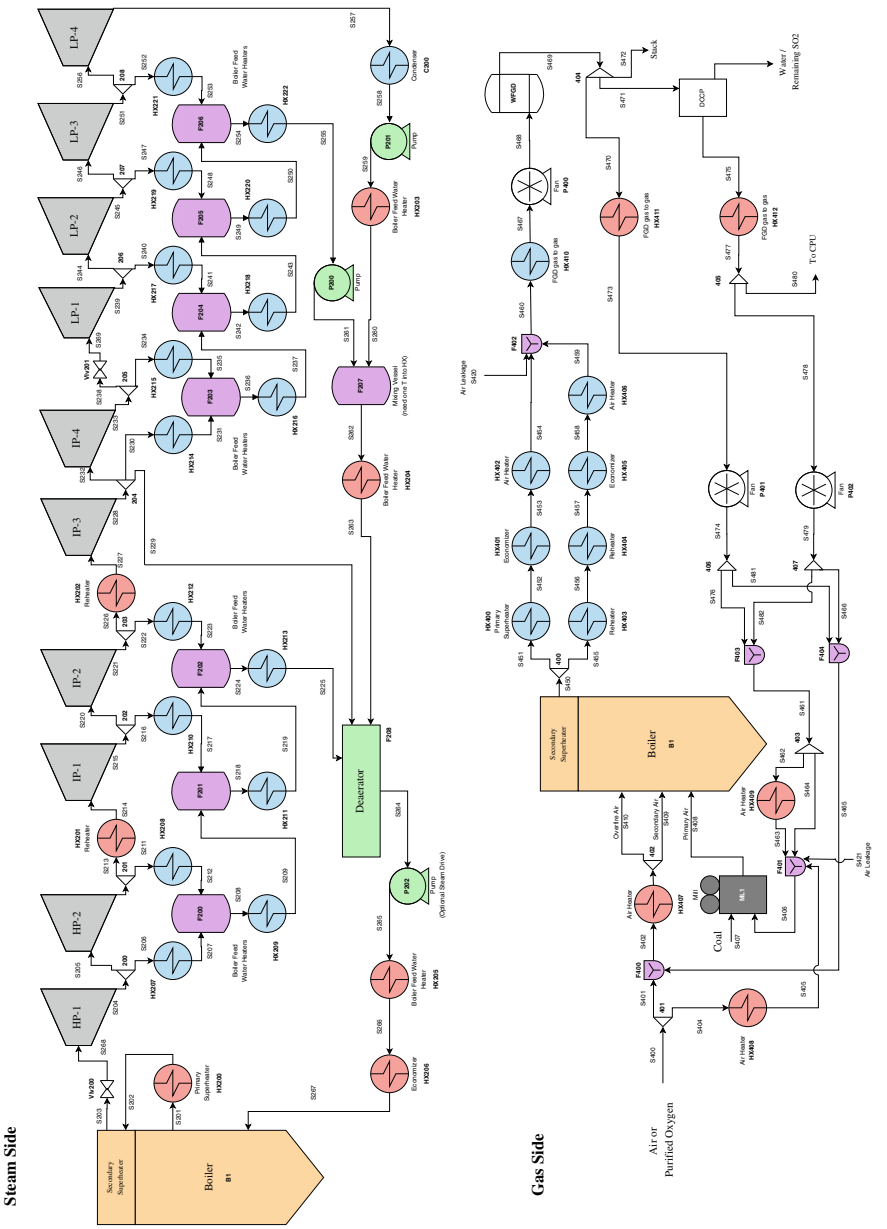
A trust region filter method has been developed, motivated by the black-box equality constraints imposed by the boiler model. By adaptively updating the surrogate model and trust region radius, this algorithm will provably converge to a local optimum of the true problem. The algorithm has been compared against

existing penalty function-based approaches (see Biegler et al. 2014), and the filter method was demonstrated to reduce calls to the black-box model by about an order of magnitude. For details of this comparison and the full trust region filter algorithm, see Eason and Biegler (2015) and Fletcher et al. (2002).

5.3 Case Study

Using the trust region optimization algorithm and hybrid boiler model, the thermal efficiency of a double reheat oxy-fired steam cycle is maximized by solving (5). The water and gas sides of the steam cycle are shown in Fig. 11. The high, intermediate and low-pressure turbine sections are divided into stage groups. Assumed pressures drops and turbine efficiencies are listed in Table 3. An air leakage rate of 3 % (mole basis) relative to the air flowrate entering the boiler is assumed; 1/3 of the air enters the system in the coal mill, and the remainder, in the baghouse. In order to maintain adequate air flow to carry the pulverized coal into the boiler, the flowrate of primary air into the boiler (*S408* in Fig. 11) has a lower bound of 2520 mol/s, which is roughly 2 kg of gas per 1 kg of coal. The primary air stream also has upper bounds on temperature and O_2 mol fraction of 400 K and 35 %, respectively, in order to reduce the risk of combustion in the coal mill. The power required for the ASU and CPU are approximated using correlations from Dowling and Biegler (2015) and Darde et al. (2009) (assuming 90 % CO_2 capture), respectively. Finally, the coal feed rate (i.e., thermal input rate) is kept constant in the optimization problem.

$$
\begin{aligned}
\max \quad & \text{Thermal Efficiency} + \rho_s Q_s + \rho_w Q_w + \rho_\perp (\text{Complementarity Violations}) \\
\text{s.t.} \quad & \text{Thermal Efficiency} = \frac{\sum W_{\text{turbine}} - \sum W_{\text{pump}} - \sum W_{\text{fan}} - W_{\text{CPU}} - W_{\text{ASU}}}{\text{Thermal Input Rate}} \\
& \text{Fixed Thermal Input Rate} \\
& \text{Steam Turbine Model} \\
& \text{Pump and Fan Models} \\
& \text{Pinch-location Heat Integration Equations} \\
& \text{Correlation-based Fuel Gas Thermodynamics Model} \\
& \text{Steam Thermodynamics TableLook-ups} \\
& \text{Hybrid Boiler Model} \\
& W_{\text{CPU}} = [-1.9091(\text{mol}\%\,\text{CO}_2 \text{ into CPU}) + 275.45\ \text{Wh/kg CO}_2] \\
& \qquad \times (90\,\%\,\text{CO}_2 \text{ capture rate}) \\
& \qquad \times (\text{mass flowrate of CO}_2 \text{ into CPU}) \\
& W_{\text{ASU}} = [2.2762(\text{mol}\%\,\text{O}_2 \text{ from ASU}) - 20.288\ \text{Wh/kg pure O}_2] \\
& \qquad \times (\text{mass flowrate of O}_2 \text{ from ASU})
\end{aligned}
\tag{5}
$$


Steam Side
Gas Side
Boiler
B1
Secondary Superheater
Deaerator
Air or Purified Oxygen
Coal
Stack
To CPU
Water / Remaining SO2

◀ **Fig. 11** The water/steam and gas sides of the steam cycle includes two steam reheaters and options for primary and secondary recycle before and/or after the DCCP. Heat exchangers are modeled as halves, where red units receive heat and blue units provide heat. The halves are implicitly matched as part of the pinch-based heat integration model. Several heat integration zones are used to distinguish the primary superheat, reheaters, economizer, air heaters, boiler feed water heaters and condenser. See Dowling (2015) for additional details

Table 3 Assumptions for oxy-fired steam cycle optimization case study

Steam side pressure drops									
Primary superheater (HX200)			2.5 %	Reheaters (HX201 & 202)					8 %
Feed water heaters (HX203–205)			4 %	Economizer (HX206)					4 %
Throttle valve (Vlv200)			4 %	IP/LP crossover (Vlv201)					3 %
Gas side pressure drops									
Primary superheater (HX400)			2 %	Reheaters (HX403)					2 %
Economizer (HX401 & 405)			2 %	Air heaters (HX402 & 406)					2 %
FGD gas-to-gas (HX410–412)			2 %	Air heaters (HX407–HX409)					2 %
FGD			0 %	DCCPS					0 %
Steam turbine stage group isentropic efficiencies									
HP-1	91.5 %	HP-2	90.5 %	IP-1	94.0 %	IP-2	93.8 %	IP-3	94.7 %
IP-4	93.5 %	LP-1	89.2 %	LP-2	88.9 %	LP-3	88.5 %	LP-4	88.2 %

Both steam thermodynamics (using steam tables) and the hybrid boiler model are incorporated into (5) using linear surrogate models. For example, steam enthalpy for each stream is approximated with

$$H = a(T - T_0) + b(P - P_0) + H_0, \tag{6}$$

where T_0, P_0 and H_0 are the temperature, pressure and enthalpy of the stream at the center of the trust region, and a and b are fitted using two additional samples from IAPWS IF-97 steam table routine (Holmgren 2007). A similar surrogate model is used to calculate each boiler output as a linear combination of the inputs.

Boiler inputs:

1. Primary air temperature
2. Secondary/over-fired air temperature
3. Average temperature of boiling water inside water wall tubes
4. Average secondary superheater steam temperature
5. Primary air component flowrates (O_2, N_2, H_2, CO, CO_2, H_2O, SO_2, H_2S, CH_4, Ar)
6. Secondary air component flowrates (same components as primary air, but different compositions)
7. Over-fired air total flowrate (same composition as secondary air)

Boiler outputs:

1. Boiler enclosure water wall heat duty
2. Secondary superheater heat duty
3. Flue gas component flowrates (same components as primary air)
4. Flue gas temperature

Optimization problem (5) was solved with two different models for ASU behavior. In Case A, the oxygen purity supplied by the ASU was fixed at 97 mol%, and the power requirement of the ASU only depended on the oxygen demand (≈200 Wh/kg pure O_2). In Case B, the oxygen purity was allowed to vary between 90 and 98 mol% with power demands varying according to the correlation from Dowling and Biegler (2015). The optimization results are shown in Table 4. Case A yields an oxy-fired power plant with a net power output of 437.4 MW and net efficiency of 33.0 %, whereas Case B yields an oxy-fired plant with net power output of 440.4 MW and efficiency of 33.2 %.[12] Interestingly, in Case B the oxygen composition is pushed to its lower bound of 90 mol%. In both cases the optimizer pushes the steam temperatures leaving the secondary superheater and reheaters to their upper bounds of 835 and 867 K, respectively, as expected. Similarly, the lower bound of 0.068 bar for the condenser operating pressures is active. The optimizer also pushes the temperature and oxygen content of the primary flue gas recycle streams (S408) to their upper bounds of 400 K and 35 mol%. This trend is consistent with the conclusions of Hagi et al. (2014a): hotter flue gas recycle involves less exergy destruction in heat exchangers and leads to higher efficiency. The temperature of the secondary and over-fired air was near 441 K for both cases.

The difference in efficiency between Cases A and B can be attributed primarily to the boiler performance; in Case B, 10.9 MW of additional heat is transferred through boiler water wall. This is in part due to the difference in O_2 concentration in the secondary air between Cases A (23.1 mol% O_2) and B (21.6 mol% O_2), as the reduced oxygen concentration in Case B has a complicated effect on the boiler performance. Gasification reactions are more favored and the emissivity of the gas mixture changes significantly. Thus, in Case B, there are higher temperatures in the boiler and hence more radiative heat transfer. Another difference is related to the utilization of heat in the flue gas. In Case B, an additional 19.5 MW of heat are transferred in the primary superheater, whereas an additional 27.6 MW of heat are transferred in Case A in the economizer. Although more total heat is extracted from the flue gas in Case A, heat is transferred at higher temperatures in Case B, and is more valuable to convert to useful mechanical work. Thus, the value of heat (i.e., exergy) is implicitly considered in the optimization problem (5), through the thermodynamic and pinch-based heat integration models. Although detailed analysis (e.g., CFD simulations for boiler, optimization with detailed models for

[12]Thermal efficiencies are calculated on a higher heating value basis. These calculation does not consider all of the auxillary loads, such as the coal mill.

Table 4 The overall heat and work balance for steam side of Fig. 11 are broken down by major components, including shaft work for the ASU and CPU

	Case A	Case B
Work from turbines (MW)	568.2	570.3
HP	94.5	94.5
IP	267.3	267.7
LP	206.4	208.1
Pumping work (MW)	12.4	12.4
Fan work (MW)	3.6	3.7
Heat from boiler (MW)	520.6	531.3
Boiler walls	446.4	457.3
Secondary superheater	74.2	74
Heat from flue gas (MW)	659.2	653.0
Primary superheater	201.0	220.5
Reheater (HX201)	168.6	168.7
Reheater (HX202)	146.6	148.4
Economizer	143.0	115.4
Heat rejected (MW)	620.9	623.3
Fuel heat rate (MW)	1325.5	1325.5
ASU power (MW)	71.2	65.6
CPU power (MW)	43.6	48.2
Net power (MW)	437.4	440.4
Thermal efficiency (HHV)	**33.0 %**	**33.2 %**
Flue gas recycle distribution		
Bypasses DCCP, to secondary rec.	29.7 %	31.3 %
To CPU after DCCP	28.0 %	29.9 %
To primary recycle after DCCP	23.6 %	25.3 %
To secondary recycle after DCCP	18.7 %	13.5 %

At the solution of (5), 20 % of the work from the steam turbines is used to power the ASU and CPU. In Case A, the O_2 purity of the ASU effluent is fixed to 97 mol%, whereas in Case B the purity is an optimization variable and bounded between 90 and 98 mol%. Finally, the distribution of flue gas recycle is reported using an inert (Ar) mole basis and the exit of the baghouse (S460, after air leakage entrance) as the reference point.

cryogenics systems) is necessary to validate these results, solving optimization problem (5) using the hybrid 1D/3D boiler model highlights complex interactions that take place between components in the oxycombustion system.

6 Conclusions and Outlook

Through the use of process simulators and exergy analysis, several research groups have identified heat integration opportunities, alternate flue gas recycle strategies, and sub-ambient gas separation system enhancements that increase the net thermal

efficiency of ambient pressure oxycombustion power plants by up to 4 %-points (absolute, either HHV or LHV bases). Although helpful, this type of analysis has a few key limitations. In almost all of the surveyed literature, the authors considered a fixed steam cycle to make the analysis tractable, and approximated the value of saved low-pressure steam through heat integration (e.g., pinch analysis). Few rigorously model the changes necessary to the steam cycle to incorporate waste heat from compression. Moreover, the coal boilers were either modeled using overly simplistic "Gibbs reactors" or analyzed a fixed design point (from experiments or CFD simulations). The former approach does not adequately capture nonlinear sensitivity of the boiler to perturbations in combustion conditions (e.g., recycle stream composition and temperature, etc.), and the latter overlooks key design decision variables. Finally, "optimization" performed in many of the reviewed studies relied on single variable sensitivity studies and manual design iterations guided by engineering insights and exergy analysis.

In contrast, we present a new equation-based framework for flowsheet optimization. Implemented in GAMS, the framework utilizes exact derivatives and state-of-the-art numerical optimization solvers. Its capabilities are demonstrated in several case studies: first, a cryogenic air separation unit (ASU) is custom designed to procedure O_2 for oxy-fired coal power plants. Both rigorous thermodynamic (Peng–Robinson) and multistream heat exchanger models are considered. The predicted energy requirements are similar to modern industrial designs. Moreover, the model is used to create Pareto optimal surfaces showing the impact of O_2 product purity on separation energy. Next, the CO_2 processing unit and compression train is optimized, and Pareto curves showing the impact of CO_2 recovery and purity on separation and compression energy are generated. In another set of case studies, the thermal efficiency of an entire oxycombustion plant is optimized. Steam table models and a new hybrid 1D/3D zonal boiler model are incorporated into the optimization problem using a trust region framework. In these case studies, correlations are used for the ASU and CPU. Two designs are analyzed: one in which the ASU effluent purity is fixed to 97 mol% O_2, and another where the ASU effluent purity is bounded between 90 and 98 mol% O_2 and optimized. By varying the ASU product purity, the net thermal efficiency increases by 0.2 $\%_{HHV}$, which is attributed to complex combustion phenomena in the boiler and higher temperatures for heat transfer to the steam cycle.

Ongoing work focuses on integration of the sub-ambient separation systems and steam cycle models, with the ultimate goal of simultaneously optimizing the entire oxycombustion process using detailed models. This will provide an unprecedented ability to balance trade-offs between subsystems in the oxycombustion process, and (1) optimize the size of the ASU and CPU, (2) optimize the purity of O_2 supplied to the boiler, and (3) explore waste heat utilization by both manipulating the pressure ratios and initial temperatures in the compressors (which impacts the quantity and quality of waste heat produced) while also modifying the temperature and quantity of extracted low-pressure steam. Although the mathematical models summarized in this chapter are sufficient for full plant optimization, implementation issues remain, including tuning of the trust region algorithm and its parameters. Finally, comparison

of optimized oxy-combustion and air-fired plant designs using the developed modeling framework are planned.

Furthermore, the new framework should be a very useful tool for evaluating, designing and optimizing other thermal energy systems in addition to coal oxy-combustion. The equation-based approach allows for large flowsheets to be optimized using detailed thermodynamic and heat integration models. Moreover, the trust region algorithm enables "black-box" unit operations to be included in the optimization problem, without losing the computational benefits of the equation-based models for the remainder of the flowsheet.

Acknowledgements This work was supported by the U.S. Department of Energy, Office of Fossil Energy as part of the Carbon Capture Simulation Initiative (CCSI). This technical effort was performed in support of the National Energy Technology Laboratory's ongoing research under the RES contract DE-FE0004000. This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

References

Adams, B. R., David, K., Senior, C., Shim, H. S., Otten, B. V., Fry, A., Wendt, J., Eddings, E., Paschedag, A., Zurek, S., Shaddix, C., & Cox, W. 2013. *Characterizaton of Oxy-combustion Impacts in Existing Coal-Fired Boilers*. Tech. rept. National Energy Technology Labratory. NT0005288.

Álvarez, L., Gharebaghi, M., Pourkashanian, M., Williams, a., Riaza, J., Pevida, C., Pis, J. J., & Rubiera, F. 2011. CFD modelling of oxy-coal combustion in an entrained flow reactor. *Fuel Processing Technology*, **92**(8), 1489–1497.

Amann, J.-M., Kanniche, M., & Bouallou, C. 2009. Natural gas combined cycle power plant modified into an O_2/CO_2 cycle for CO_2 capture. *Energy Conversion and Management*, **50**(3), 510–521.

Andersson, Klas, & Johnsson, Filip. 2006. Process evaluation of an 865 MW_e lignite fired O_2/CO_2 power plant. *Energy Conversion and Management*, **47**(18–19), 3487–3498.

Arias, B., Pevida, C., Rubiera, F., & Pis, J. J. 2008. Effect of biomass blending on coal ignition and burnout during oxy-fuel combustion. *Fuel*, **87**(12), 2753–2759.

Bejarano, Paula A., & Levendis, Yiannis A. 2008. Single-coal-particle combustion in O_2/N_2 and O_2/CO_2 environments. *Combustion and Flame*, **153**(1–2), 270–287.

Belaissaoui, Bouchra, Le Moullec, Yann, Hagi, Hayato, & Favre, Eric. 2014. Energy efficiency of oxygen enriched air production technologies: Cryogeny vs membranes. *Separation and Purification Technology*, **125**, 142–150.

Besong, Marvine Tambe, Maroto-Valer, M. Mercedes, & Finn, Adrian J. 2013. Study of design parameters affecting the performance of CO_2 purification units in oxy-fuel combustion. *International Journal of Greenhouse Gas Control*, **12**(0), 441–449.

Biegler, L. T., Grossmann, I. E., & Westerberg, A. W. 1997. *Systematic Methods of Chemical Process Design*. Prentice Hall PTR.

Biegler, Lorenz T. 2010. *Nonlinear Programming: Concepts, Algorithms, and Applications to Chemical Processes*. Society of Industrial and Applied Mathematics.

Biegler, Lorenz T., Lang, Yi-dong, & Lin, Weijie. 2014. Multi-scale Optimization for Process Systems Engineering. *Computers & Chemical Engineering*, **60**(Aug.), 17–30.

Bolea, Irene, Romeo, Luis M., & Pallarés, David. 2012. The role of external heat exchangers in oxy-fuel circulating fluidized bed. *Applied Energy*, **94**, 215–223.

Bolland, Olav, & Undrum, Henriette. 2003. A novel methodology for comparing CO_2 capture options for natural gas-fired combined cycle plants. *Advances in Environmental Research*, **7**(4), 901–911.

Buhre, B. J. P., Elliott, L. K., Sheng, C. D., Gupta, R. P., & Wall, T. F. 2005. Oxy-fuel combustion technology for coal-fired power generation. *Progress in Energy and Combustion Science*, **31**(4), 283–307.

Burdyny, Thomas, & Struchtrup, Henning. 2010. Hybrid membrane/cryogenic separation of oxygen from air for use in the oxy-fuel process. *Energy*, **35**(5), 1884–1897.

Cao, Huali, Sun, Shaozeng, Liu, Yinghui, & Wall, Terry F. 2010. Computational Fluid Dynamics Modeling of NOx Reduction Mechanism in Oxy-Fuel Combustion. *Energy & Fuels*, **24**(1), 131–135.

Chen, Lei, Yong, Sze Zheng, & Ghoniem, Ahmed F. 2012. Oxy-fuel combustion of pulverized coal: Characterization, fundamentals, stabilization and CFD modeling. *Progress in Energy and Combustion Science*, **38**(2), 156–214.

Chung, Timothy S., Patiño Echeverri, Dalia, & Johnson, Timothy L. 2011. Expert assessments of retrofitting coal-fired power plants with carbon dioxide capture technologies. *Energy Policy*, **39**(9), 5609–5620.

Conn, Andrew R, Scheinberg, Katya, & Vicente, Luis N. 2009. *Introduction to derivative-free optimization*. Vol. 8. Siam.

Croiset, E., & Thambimuthu, K. V. 2001. NOx and SO_2 emissions from O_2/CO_2 recycle coal combustion. *Fuel*, **80**(14), 2117–2121.

Cuéllar-Franca, Rosa M., & Azapagic, Adisa. 2015. Carbon capture, storage and utilisation technologies: A critical analysis and comparison of their life cycle environmental impacts. *Journal of CO2 Utilization*, **9**, 82–102.

Czyperek, M., Zapp, P., Bouwmeester, H. J. M., Modigell, M., Ebert, K., Voigt, I., Meulenberg, W. A., Singheiser, L., & Stöver, D. 2010. Gas separation membranes for zero-emission fossil power plants: MEM-BRAIN. *Journal of Membrane Science*, **359**(1–2), 149–159.

Darde, Arthur, Prabhakar, Rajeev, Tranier, Jean-Pierre, & Perrin, Nicolas. 2009. Air separation and flue gas compression and purification units for oxy-coal combustion systems. *Energy Procedia*, **1**(1), 527–534.

Dowling, Alexander W. 2015. *An Equation-based Framework for Large-Scale Flowsheet Optimization and Applications for Oxycombustion Power System Design*. Ph.D. thesis, Carnegie Mellon University.

Dowling, Alexander W., & Biegler, Lorenz T. 2013. Optimization-based Process Synthesis for Sustainable Power Generation. *Chemical Engineering Transactions*, **35**, 1–12.

Dowling, Alexander W., & Biegler, Lorenz T. 2014. Rigorous Optimization-based Synthesis of Distillation Cascades without Integer Variables. *Pages 55–60 of:* Klemeš, Jiří Jaromír, Varbanov, Petar Sabev, & Liew, Peng Yen (eds), *24th European Symposium on Computer Aided Process Engineering*. Computer Aided Chemical Engineering, vol. 33. Elsevier.

Dowling, Alexander W., & Biegler, Lorenz T. 2015. A framework for efficient large scale equation-oriented flowsheet optimization. *Computers & Chemical Engineering*, **72**, 3–20.

Dowling, Alexander W., Eason, John P., Ma, Jinliang, Miller, David C., & Biegler, Lorenz T. 2014a. Coal Oxycombustion Power Plant Optimization using First Principles and Surrogate Boiler Models. *Energy Procedia*, **63**, 352–361.

Dowling, Alexander W., Balwani, Cheshta, Gao, Qianwen, & Biegler, Lorenz T. 2014b. Equation-oriented Optimization of Cryogenic Systems for Coal Oxycombustion Power Generation. *Energy Procedia*, **63**, 421–430.

Dowling, Alexander W., Gao, Qianwen, & Biegler, Lorenz T. 2014c. Equation-Oriented Optimization of Cryogenic Systems for Coal Oxycombustion Power Plants. *Pages 501–506 of:* Eden, Mario R., Siirola, John D., & Towler, Gavin P. (eds), *Proceedings of the 8th International Conference on Foundations of Computer-Aided Process Design*. Computer Aided Chemical Engineering, vol. 34. Elsevier.

Dowling, Alexander W., Balwani, Cheshta, Gao, Qianwen, & Biegler, Lorenz T. 2015. Optimization of sub-ambient separation systems with embedded cubic equation of state thermodynamic models and complementarity constraints. *Computers & Chemical Engineering,* **81**(Oct.), 323–343.

Duan, Lunbo, Sun, Haicheng, Zhao, Changsui, Zhou, Wu, & Chen, Xiaoping. 2014. Coal combustion characteristics on an oxy-fuel circulating fluidized bed combustor with warm flue gas recycle. *Fuel*, **127**, 47–51.

Duran, M. A., & Grossmann, I. E. 1986. Simultaneous optimization and heat integration of chemical processes. *AIChE Journal*, **32**(1), 123–138.

Eason, John P., & Biegler, Lorenz T. 2015. Reduced model trust region methods for embedding complex simulations in optimization problems. *Pages 773–778 of:* Gernaey, Krist V., Huusom, Jakob K., & Gani, Rafiqul (eds), *12th International Symposium on Process Systems Engineering and 25th European Symposium on Computer Aided Process Engineering*. Computer Aided Chemical Engineering, vol. 37. Elsevier.

Edge, P.J., Heggs, P.J., Pourkashanian, M., Stephenson, P.L., & Williams, a. 2012. A reduced order full plant model for oxyfuel combustion. *Fuel*, **101**(Nov.), 234–243.

Edge, P.J., Heggs, P.J., Pourkashanian, M., & Stephenson, P.L. 2013. Integrated fluid dynamics-process modelling of a coal-fired power plant with carbon capture. *Applied Thermal Engineering*, **60**(1–2), 242–250.

Engels, S., Beggel, F., Modigell, M., & Stadler, H. 2010. Simulation of a membrane unit for oxyfuel power plants under consideration of realistic BSCF membrane properties. *Journal of Membrane Science*, **359**(1–2), 93–101.

Faé Gomes, Gabriel M, Vilela, Antônio C F, da Silva Priebe, Guilherme P, & Dalla Zen, Leandro. 2015. Retrofit of a Bubbling Fluidized Bed Pilot Plant From Air Combustion to Oxyfuel Combustion. *Journal of Energy Resources Technology*, **137**(3), 034501/1–034501/7.

Fei, Y., Black, S., Szuhánszki, J., Ma, L., Ingham, D.B., Stanger, P.J., & Pourkashanian, M. 2015. Evaluation of the potential of retrofitting a coal power plant to oxy-firing using CFD and process co-simulation. *Fuel Processing Technology*, **131**, 45–58.

Fernández-Miranda, Nuria, Lopez-Anton, M Antonia, Daz-Somoano, Mercedes, & Martnez-Tarazona, M. Rosa. 2014. Effect of Oxy-Combustion Flue Gas on Mercury Oxidation. *Environmental Science & Technology*, **48**, 7164–7170.

Fletcher, Roger, Gould, Nicholas IM, Leyffer, Sven, Toint, Philippe L, & Wächter, Andreas. 2002. Global convergence of a trust-region SQP-filter algorithm for general nonlinear programming. *SIAM Journal on Optimization*, **13**(3), 635–659.

Foerg, Wolfgang. 2002. History of cryogenics: the epoch of the pioneers from the beginning to the year 1911. *International Journal of Refrigeration*, **25**(3), 283–292.

Fout, Tim, Zoelle, Alexander, Keairns, Dale, Turner, Marc, Woods, Mark, Kuehn, Norma, Shah, Vasant, Chou, Vincent, & Pinkerton, Lora. 2015 (July). *Cost and Performance Baseline for Fossil Energy Plants Volume 1a: Revision 3*. Tech. rept. National Energy Technology Labratory. DOE/NETL-2015/1723.

Fry, Andrew, Adams, Brad, Davis, Kevin, Swensen, Dave, Munson, Shawn, & Cox, William. 2011a. An investigation into the likely impact of oxy-coal retrofit on fire-side corrosion behavior in utility boilers. *International Journal of Greenhouse Gas Control*, **5**(Supplemental), S179–S185.

Fry, Andrew, Adams, Brad, Paschedag, Alan, Kazalski, Paul, Carney, Casey, Oryshchyn, Danylo, Woodside, Rigel, Gerdemann, Steve, & Ochs, Thomas. 2011b. Principles for retrofitting coal burners for oxy-combustion. *International Journal of Greenhouse Gas Control*, **5** (Supplemental), S151–S158.

Fu, Chao, & Gundersen, Truls. 2010. Heat Integration of an Oxy-Combustion Process for Coal-Fired Power Plants with CO_2 Capture by Pinch Analysis. *Chemical Engineering Transactions*, **21**, 181–186.

Fu, Chao, & Gundersen, Truls. 2012a. Integrating the Compression Heat in Oxy-combustion Power Plants with CO_2 Capture. *Chemical Engineering Transactions*, **29**, 781–786.

Fu, Chao, & Gundersen, Truls. 2012b. Techno-economic analysis of CO_2 conditioning processes in a coal based oxy-combustion power plant. *International Journal of Greenhouse Gas Control*, **9**, 419–427.

Fu, Chao, & Gundersen, Truls. 2012c. Using exergy analysis to reduce power consumption in air separation units for oxy-combustion processes. *Energy*, **44**(Aug.), 60–68.

Fu, Chao, & Gundersen, Truls. 2013. Recuperative vapor recompression heat pumps in cryogenic air separation processes. *Energy*, **59**(Sept.), 708–718.

Fu, Chao, Anantharaman, Rahul, & Gundersen, Truls. 2015a. Optimal integration of compression heat with regenerative steam Rankine cycles in oxy-combustion coal based power plants. *Energy*, **84**, 612–622.

Fu, Qian, Kansha, Yasuki, Song, Chunfeng, Liu, Yuping, Ishizuka, Masanori, & Tsutsumi, Atsushi. 2015b. A cryogenic air separation process based on self-heat recuperation for oxy-combustion plants. *Applied Energy*.

Gil, M. V., Riaza, J., Álvarez, L., Pevida, C., Pis, J. J., & Rubiera, F. 2012. Oxy-fuel combustion kinetics and morphology of coal chars obtained in N_2 and CO_2 atmospheres in an entrained flow reactor. *Applied Energy*, **91**(1), 67–74.

Gopan, Akshay, Kumfer, Benjamin M., Phillips, Jeffrey, Thimsen, David, Smith, Richard, & Axelbaum, Richard L. 2014. Process design and performance analysis of a Staged, Pressurized Oxy-Combustion (SPOC) power plant for carbon capture. *Applied Energy*, **125**, 179–188.

Gopan, Akshay, Kumfer, Benjamin M., & Axelbaum, Richard L. 2015. Effect of operating pressure and fuel moisture on net plant efficiency of a staged, pressurized oxy-combustion power plant. *International Journal of Greenhouse Gas Control*, **39**, 390–396.

Hagi, Hayato, Nemer, Maroun, Le Moullec, Yann, & Bouallou, Chakib. 2013. Assessment of the Flue Gas Recycle Strategies on Oxy-Coal Power Plants using an Exergy-based Methodology. *Chemical Engineering Transactions*, **35**.

Hagi, Hayato, Le Moullec, Yann, Nemer, Maroun, & Bouallou, Chakib. 2014a. Performance assessment of first generation oxy-coal power plants through an exergy-based process integration methodology. *Energy*, **69**, 272–284.

Hagi, Hayato, Le, Yann, & Bouallou, Chakib. 2014b. Towards Second Generation Oxy-Pulverized Coal Power Plants : Energy Penalty Reduction Potential of Pressurized Oxy-Combustion Systems. *Energy Procedia*, **63**, 431–439.

Harkin, Trent, Hoadley, Andrew, & Hooper, Barry. 2010. Reducing the energy penalty of CO_2 capture and compression using pinch analysis. *Journal of Cleaner Production*, **18**(9), 857–866.

Hashim, S. S., Mohamed, A. R., & Bhatia, S. 2011. Oxygen separation from air using ceramic-based membrane technology for sustainable fuel production and power generation. *Renewable and Sustainable Energy Reviews*, **15**(2), 1284–1293.

Haykiri-Acma, H., Turan, A. Z., Yaman, S., & Kucukbayrak, S. 2010. Controlling the excess heat from oxy-combustion of coal by blending with biomass. *Fuel Processing Technology*, **91**(11), 1569–1575.

Higginbotham, Paul, White, Vince, Fogash, Kevin, & Guvelioglu, Galip. 2011. Oxygen supply for oxyfuel CO_2 capture. *International Journal of Greenhouse Gas Control*, **5**(Supplemental), S194–S203.

Holmgren, Magnus. 2007. *X Steam: Thermodynamic properties of water and steam.*

Hong, Jongsup, Chaudhry, Gunaranjan, Brisson, J. G., Field, Randall, Gazzino, Marco, & Ghoniem, Ahmed F. 2009. Analysis of oxy-fuel combustion power cycle utilizing a pressurized coal combustor. *Energy*, **34**(9), 1332–1340.

Hong, Jongsup, Field, Randall, Gazzino, Marco, & Ghoniem, Ahmed F. 2010. Operating pressure dependence of the pressurized oxy-fuel combustion power cycle. *Energy*, **35**(12), 5391–5399.

Hultgren, Matias, Ikonen, Enso, & Kovács, Jenö. 2014. Oxidant control and air-oxy switching concepts for CFB furnace operation. *Computers and Chemical Engineering*, **61**, 203–219.

Johansson, Robert, Leckner, Bo, Andersson, Klas, & Johnsson, Filip. 2013. Influence of particle and gas radiation in oxy-fuel combustion. *International Journal of Heat and Mass Transfer*, **65**, 143–152.

Jordal, Kristin, Anheden, Marie, Yan, Jinying, & Strömberg, Lars. 2004. Oxyfuel Combustion for Coal-Fired Power Generation with CO_2 Capture Opportunities and Challenges. *In: 7th International Conference on Greenhouse Gas Technologies*.

Kamath, Ravindra S., Biegler, Lorenz T., & Grossmann, Ignacio E. 2010. An equation-oriented approach for handling thermodynamics based on cubic equation of state in process optimization. *Computers & Chemical Engineering*, **34**(12), 2085–2096.

Kamath, Ravindra S, Biegler, Lorenz T, & Grossmann, Ignacio E. 2012. Modeling Multistream Heat Exchangers with and without Phase Changes for Simultaneous Optimization and Heat Integration. *AIChE Journal*, **58**(1).

Kiga, T., Takano, S., Kimura, N., Omata, K., Okawa, M., Mori, T., & Kato, M. 1997. Characteristics of pulverized-coal combustion in the system of oxygen/recycled flue gas combustion. *Energy Conversion and Management*, **38**(Supplemental), S129–S134.

Krishnamoorthy, Gautham. 2012. A new weighted-sum-of-gray-gases model for oxy-combustion scenarios. *International Journal of Energy Research*, **31**, 1752–1763.

Krishnamoorthy, Gautham, Sami, Muhammad, Orsino, Stefano, Perera, Anura, Shahnam, Mehrdad, & Huckaby, E. David. 2010. Radiation modelling in oxy-fuel combustion scenarios. *International Journal of Computational Fluid Dynamics*, **24**(3–4), 69–82.

Kupila, Kati, Dernjatin, Pauli, Sormunen, Risto, Sumida, Tadashi, Kiyama, Kenji, Briglia, Alain, Sanchez-Molinero, Ivan, & Darde, Arthur. 2011. Risk analysis of FORTUMs 560 MW_e net power plant retrofit to oxyfuel combustion. *Energy Procedia*, **4**, 1820–1827.

Leckner, Bo, & Gómez-Barea, Alberto. 2014. Oxy-fuel combustion in circulating fluidized bed boilers. *Applied Energy*, **125**, 308–318.

Li, H., Yan, J., Yan, J., & Anheden, M. 2009. Impurity impacts on the purification process in oxy-fuel combustion based CO_2 capture and storage system. *Applied Energy*, **86**(2), 202–213.

Li, Hailong, Hu, Yukun, Ditaranto, Mario, Willson, David, & Yan, Jinyue. 2013. Optimization of Cryogenic CO_2 Purification for Oxy-coal Combustion. *Energy Procedia*, **37**, 1341–1347.

López, R., Fernández, C., Fierro, J., Cara, J., Martnez, O., & Sánchez, M.E. 2014. Oxy-combustion of corn, sunflower, rape and microalgae bioresidues and their blends from the perspective of thermogravimetric analysis. *Energy*, **74**, 845–854.

Ma, Jinliang, Dowling, Alex, Eason, John, Biegler, Lorenz, & Miller, David. 2014. Development of First Principle Boiler Model and Its Reduced Order Model for the Optimization of Oxy-combustion Power Generation System. *In: 39th International Technical Conference on Clean Coal & Fuel Systems*.

Ma, Jinliang, Eason, John P., Dowling, Alexander W., Biegler, Lorenz T., & Miller, David C. 2016. Development of a first-principles hybrid boiler model for oxy-combustion power generation system. *International Journal of Greenhouse Gas Control*, **46**, 136–157.

Malavasi, M., & Rossetti, E. 2013. *High-efficiency combustors with reduced environmental impact and processes for power generation derivable therefrom*. US Patent 8,453,583.

Mancini, N. D., & Mitsos, A. 2011a. Ion transport membrane reactors for oxy-combustion - Part I: Intermediate-fidelity modeling. *Energy*, **36**(8), 4701–4720.

Mancini, N. D., & Mitsos, A. 2011b. Ion transport membrane reactors for oxy-combustion-Part II: Analysis and comparison of alternatives. *Energy*, **36**(8), 4721–4739.

Matuszewski, Michael. 2010. *Cost and Performance for Low-Rank Pulverized Coal Oxycombustion Energy Plants DOE/NETL-401/093010*. Tech. rept. September. National Energy Technology Labratory.

McCauley, K. J., Farzan, H., Alexander, K. C., McDonald, D. K., Varagani, R., Prabhakar, R., Tranier, J. P., & Perrin, N. 2009. Commercialization of oxy-coal combustion: Applying results of a large 30 MWth pilot project. *Energy Procedia*, **1**(1), 439–446.

McDonald, D. K., Moorman, Steve, Darde, Arthur, & de Limon, Sebastien. 2011. Oxy-Combustion: A Promising Technology for Coal-Fired Plants. *Power Magazine*.

Molina, Alejandro, & Shaddix, Christopher R. 2007. Ignition and devolatilization of pulverized bituminous coal particles during oxygen/carbon dioxide coal combustion. *Proceedings of the Combustion Institute*, **31**(2), 1905–1912.

Murphy, Jeffrey J., & Shaddix, Christopher R. 2006. Combustion kinetics of coal chars in oxygen-enriched environments. *Combustion and Flame*, **144**(4), 710–729.

Myöhänen, Kari, Hyppänen, Timo, Pikkarainen, Toni, Eriksson, Timo, & Hotta, Arto. 2009. Near Zero CO_2 Emissions in Coal Firing with Oxy-Fuel Circulating Fluidized Bed Boiler. *Chemical Engineering & Technology*, **32**(3), 355–363.

Okkes, A. G., & Badge, B. V. 1987. Get acid dew point of flue gas. *Hydrocarbon Processing*, **66**(7), 53–55.

Pawlak-Kruczek, Halina, Ostrycharczyk, Michal, Baranowski, Marcin, Czerep, Michal, & Zgóra, Jacek. 2013. Co-firing of biomass with pulverised coal in oxygen enriched atmosphere. *Chemical and Process Engineering*, **34**(2), 215–226.

Pei, Xiaohui, He, Boshu, Yan, Linbo, Wang, Chaojun, Song, Weining, & Song, Jingge. 2013. Process simulation of oxy-fuel combustion for a 300MW pulverized coal-fired power plant using Aspen Plus. *Energy Conversion and Management*, **76**(Dec.), 581–587.

Perrin, Nicolas, Dubettier, Richard, Lockwood, Frederick, Tranier, Jean-Pierre, Bourhy-Weber, Claire, & Terrien, Paul. 2014. Oxycombustion for coal power plants: Advantages, solutions and projects. *Applied Thermal Engineering*, **74**, 75–82.

Pipitone, Gabriele, & Bolland, Olav. 2009. Power generation with CO_2 Capture: Technology for CO_2 purification. *International Journal of Greenhouse Gas Control*, **3**(5), 528–534.

Rajhi, M. A., Ben-Mansour, R., Habib, M. A., Nemitallah, M. A., & Andersson, K. 2014. Evaluation of gas radiation models in CFD modeling of oxy-combustion. *Energy Conversion and Management*, **81**, 83–97.

Rathnam, Renu Kumar, Elliott, Liza K., Wall, Terry F., Liu, Yinghui, & Moghtaderi, Behdad. 2009. Differences in reactivity of pulverised coal in air (O_2/N_2) and oxy-fuel (O_2/CO_2) conditions. *Fuel Processing Technology*, **90**(6), 797–802.

Romeo, Luis M., Bolea, Irene, Lara, Yolanda, & Escosa, Jesús M. 2009. Optimization of intercooling compression in CO_2 capture systems. *Applied Thermal Engineering*, **29**(8–9), 1744–1751.

Roy, Bithi, & Bhattacharya, Sankar. 2014. Oxy-fuel fluidized bed combustion using Victorian brown coal: An experimental investigation. *Fuel Processing Technology*, **117**, 23–29.

Scheffknecht, Günter, Al-Makhadmeh, Leema, Schnell, Uwe, & Maier, Jörg. 2011. Oxy-fuel coal combustion A review of the current state-of-the-art. *International Journal of Greenhouse Gas Control*, **5**(July), S16–S35.

Seltzer, Andrew, Fan, Zhen, Hack, Horst, & Shah, Minish. 2010 (June). Simulation and Prediction of Pollutants in a Flexi-BurnTM Oxyfuel Pulverized Coal Power Plant. *In: 35th International Technical Conference on Coal Utilization & Fuel Systems*.

Senneca, Osvalda, & Cortese, Luciano. 2012. Kinetics of coal oxy-combustion by means of different experimental techniques. *Fuel*, **102**, 751–759.

Shah, Minish, Degenstein, Nick, Zanfir, Monica, Kumar, Ravi, Bugayong, Jennifer, & Burgers, Ken. 2011. Near zero emissions oxy-combustion CO_2 purification technology. *Energy Procedia*, **4**, 988–995.

Skorek-Osikowska, Anna, Bartela, Lukasz, Kotowicz, Janusz, & Job, Marcin. 2013. Thermodynamic and economic analysis of the different variants of a coal-fired, 460 MW power plant using oxy-combustion technology. *Energy Conversion and Management*, **76**, 109–120.

Smart, J. P., O'Nions, P., & Riley, G. S. 2010a. Radiation and convective heat transfer, and burnout in oxy-coal combustion. *Fuel*, **89**(9), 2468–2476.

Smart, John P., Patel, Rajeshriben, & Riley, Gerry S. 2010b. Oxy-fuel combustion of coal and biomass, the effect on radiative and convective heat transfer and burnout. *Combustion and Flame*, **157**(12), 2230–2240.

Smith, A. R., & Klosek, J. 2001. A review of air separation technologies and their integration with energy conversion processes. *Fuel Processing Technology*, **70**(2), 115–134.

Songolzadeh, Mohammad, Soleimani, Mansooreh, Takht Ravanchi, Maryam, & Songolzadeh, Reza. 2014. Carbon dioxide separation from flue gases: A technological review emphasizing reduction in greenhouse gas emissions. *The Scientific World Journal*, **2014**, 1–34.

Soundararajan, Rengarajan, & Gundersen, Truls. 2013. Coal based power plants using oxy-combustion for CO_2 capture: Pressurized coal combustion to reduce capture penalty. *Applied Thermal Engineering*, **61**(1), 115–122.

Soundararajan, Rengarajan, Anantharaman, Rahul, & Gundersen, Truls. 2014a. Design of Steam Cycles for Oxy-combustion Coal based Power Plants with Emphasis on Heat Integration. *Energy Procedia*, **51**(1876), 119–126.

Soundararajan, Rengarajan, Gundersen, Truls, & Ditaranto, Mario. 2014b. Oxy-Combustion Coal Based Power Plants: Study of Operating Pressure, Oxygen Purity and Downstream Purification Parameters. *Chemical Engineering Transactions*, **39**, 229–234.

Stadler, Hannes, Beggel, Franz, Habermehl, Martin, Persigehl, Bernhard, Kneer, Reinhold, Modigell, Michael, & Jeschke, Peter. 2011. Oxyfuel coal combustion by efficient integration of oxygen transport membranes. *International Journal of Greenhouse Gas Control*, **5**(1), 7–15.

Tan, Yewen, Croiset, Eric, Douglas, Mark A., & Thambimuthu, Kelly V. 2006. Combustion characteristics of coal in a mixture of oxygen and recycled flue gas. *Fuel*, **85**(4), 507–512.

The Babcock and Wilcox Company. 2005a. Boilers, Superheaters and Reheaters. *Chap. 19, pages 1–22 of: Steam: its generation and use*, 41 edn.

The Babcock and Wilcox Company. 2005b. Economizers and Air Heaters. *Chap. 20, pages 1–18 of: Steam: its generation and use*, 41 edn.

Tigges, K. D., Klauke, F., Bergins, C., Busekrus, K., Niesbach, J., Ehmann, M., Kuhr, C., Hoffmeister, F., Vollmer, B., Buddenberg, T., Wu, Song, & Kukoski, Allan. 2009. Conversion of existing coal-fired power plants to oxyfuel combustion: Case study with experimental results and CFD-simulations. *Energy Procedia*, **1**(1), 549–556.

Toftegaard, Maja B., Brix, Jacob, Jensen, Peter A., Glarborg, Peter, & Jensen, Anker D. 2010. Oxy-fuel combustion of solid fuels. *Progress in Energy and Combustion Science*, **36**(5), 581–625.

Tranier, Jean-Pierre, Dubettier, Richard, Darde, Arthur, & Perrin, Nicolas. 2011. Air separation, flue gas compression and purification units for oxy-coal combustion systems. *Energy Procedia*, **4**(Jan.), 966–971.

Wall, Terry, Liu, Yinghui, Spero, Chris, Elliott, Liza, Khare, Sameer, Rathnam, Renu, Zeenathal, Farida, Moghtaderi, Behdad, Buhre, Bart, Sheng, Changdong, Gupta, Raj, Yamada, Toshihiko, Makino, Keiji, & Yu, Jianglong. 2009. An overview on oxyfuel coal combustion-State of the art research and technology development. *Chemical Engineering Research and Design*, **87**(8), 1003–1016.

Warzecha, Piotr, & Boguslawski, Andrzej. 2014. Simulations of pulverized coal oxy-combustion in swirl burner using RANS and LES methods. *Fuel Processing Technology*, **119**, 130–135.

White, Vince, Torrente-Murciano, Laura, Sturgeon, David, & Chadwick, David. 2009. Purification of oxyfuel-derived CO_2. *Energy Procedia*, **1**(1), 399–406.

Xiong, Jie, Zhao, Haibo, Zheng, Chuguang, Liu, Zhaohui, Zeng, Lingda, Liu, Hao, & Qiu, Jianrong. 2009. An economic feasibility study of O_2/CO_2 recycle combustion technology based on existing coal-fired power plants in China. *Fuel*, **88**(6), 1135–1142.

Xiong, Jie, Zhao, Haibo, & Zheng, Chuguang. 2011a. Exergy Analysis of a 600 MW_e Oxy-combustion Pulverized-Coal-Fired Power Plant. *Energy*, **25**, 3854–3864.

Xiong, Jie, Zhao, Haibo, Chen, Meng, & Zheng, Chuguang. 2011b. Simulation Study of an 800 MW_e Oxy-combustion Pulverized-Coal-Fired Power Plant. *Energy & Fuels*, **25**, 2405–2415.

Xiong, Jie, Zhao, Haibo, & Zheng, Chuguang. 2012. Thermoeconomic cost analysis of a 600 MW_e oxy-combustion pulverized-coal-fired power plant. *International Journal of Greenhouse Gas Control*, **9**, 469–483.

Yan, Kai, Wu, Xiaojiang, Hoadley, Andrew, Xu, Xueyuan, Zhang, Jianwen, & Zhang, Lian. 2015. Sensitivity analysis of oxy-fuel power plant system. *Energy Conversion and Management*, **98**, 138–150.

Zanganeh, Kourosh E., & Shafeen, Ahmed. 2007. A novel process integration, optimization and design approach for large-scale implementation of oxy-fired coal power plants with CO_2 capture. *International Journal of Greenhouse Gas Control*, **1**(1), 47–54.

Zanganeh, Kourosh E., Shafeen, Ahmed, & Salvador, Carlos. 2009. *CO_2 Capture and Developement of an Advanced Pilot-Scale Cryogenic Separation and Compression Unit*. *Energy Procedia*, **1**, 247–252.

Zebian, Hussam, & Mitsos, Alexander. 2013. Pressurized oxy-coal combustion: Ideally flexible to uncertainties. *Energy*, **57**, 513–526.

Zebian, Hussam, Gazzino, Marco, & Mitsos, Alexander. 2012. Multi-variable optimization of pressurized oxy-coal combustion. *Energy*, **38**(1), 37–57.

Zebian, Hussam, Rossi, Nicola, Gazzino, Marco, Cumbo, Danila, & Mitsos, Alexander. 2013. Optimal design and operation of pressurized oxy-coal combustion with a direct contact separation column. *Energy*, **49**(Jan.), 268–278.

Zhou, Wu, Zhao, Changsui, Duan, Lunbo, Chen, Xiaoping, & Liu, Daoyin. 2013. A simulation study of coal combustion under O_2/CO_2 and O_2/RFG atmospheres in circulating fluidized bed. *Chemical Engineering Journal*, **223**, 816–823.

Zhuang, Ye, & Pavlish, John H. 2012. Fate of Hazardous Air Pollutants in Oxygen-fired Coal Combustion with Different Flue Gas Recycling. *Environmental Science and Technology*, **46**(8), 4657–4665.

Wind Energy

Jason Ganley, Jie Zhang and Bri-Mathias Hodge

Abstract Wind energy is a variable and uncertain renewable resource that has long been used to produce mechanical work, and has developed into a large producer of global electricity needs. As renewable sources of energy and feedstocks become more important globally to produce sustainable products, many different processes have started adopting wind power as an energy source. Many times this is through a conversion to hydrogen through electrolysis that allows for a more continuous process input. Other important pathways include methanol and ammonia. As the demand for sustainable products and production pathways increases, and wind power capital costs decrease, the role of wind power in chemical and energy production seems poised to increase significantly.

1 Introduction

Mankind has long harnessed wind energy to help perform work, with the Persians first building vertical axis windmills approximately around 900 AD. During the Middle Ages horizontal axis windmills were used extensively in Europe for many mechanical tasks, such as water pumping and grain milling. Similar windmill variants with many fan-type blades came into widespread usage in the Western United States during the nineteenth century, pumping water for farms and steam

J. Ganley (✉)
Department of Chemical and Biological Engineering, Colorado School of Mines, Golden, CO, USA
e-mail: jganley@mines.edu

J. Zhang
Department of Mechanical Engineering, University of Texas at Dallas, Richardson TX, USA
e-mail: jiezhang@utdallas.edu

B.-M. Hodge
Power Systems Engineering Center, National Renewable Energy Laboratory, Golden, CO, USA
e-mail: bri.mathias.hodge@nrel.gov

M. Martín (ed.), *Alternative Energy Sources and Technologies*,
DOI 10.1007/978-3-319-28752-2_6

railroads (Manwell et al. 2002). With the invention of electrical generators in the late nineteenth century, wind power began to be utilized for producing electricity, instead of just mechanical work. Most of these early wind turbines had a low power capacity compared to modern turbines, with a notable exception being the Smith–Putnam 1.25 MW machine built at Grandpa's Knob in Vermont (Putnam 1948). Sustained and widespread interest in wind power did not occur until the oil crisis of the 1970s (Burton et al. 2001), after which commercial scale turbines started to appear regularly, especially in Denmark, Germany, and California.

The modern wind energy industry has developed dramatically since 2000, as wind power capacity worldwide has grown substantially, reaching over 318 GW in 2013 (Bilgili et al. 2015), and providing over 4 % of electricity production in the US (IEA 2014) and almost 8 % of electricity demand in the EU (EWEA 2014). This dramatic growth has been driven both by policies enacted in response to climate change (Wiser et al. 2011; Hurlbut et al. 2013), and to the decreasing costs of wind power (Lantz et al. 2012). Wind energy has been one of the most widely deployed alternative energy sources, and therefore is likely one of the most technologically mature. However, wind power output is also variable and uncertain, in contrast to the thermal power plants that it is often replacing, because of the meteorological processes that drive wind power generation. Since these meteorological phenomena vary over time and space, wind power output from geographically dispersed plants is not perfectly correlated. In practice this means that the variability and uncertainty of wind power is reduced as it is aggregated over large spatial and temporal scales. An example of this variability and uncertainty over 1 week in the power system of Texas is shown in Fig. 1. At small penetration rates in large interconnected power systems, the impacts of this variability and uncertainty are limited, as the power system was designed to handle variable and uncertain load. It is only when the wind penetration rate increases, or when a small or islanded power system is considered that this variability and uncertainty can present challenges.

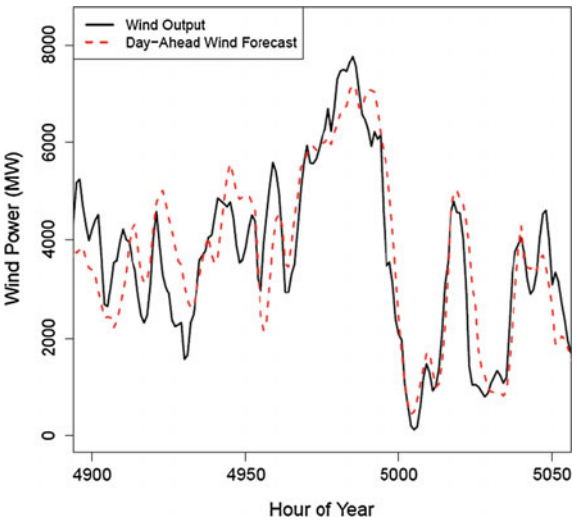


Fig. 1 Aggregated wind power output and day-ahead forecasts, at hourly resolution, for 1 week in the ERCOT system (Hodge et al. 2012)

Wind energy's variability and uncertainty present both challenges and opportunities for its utilization in non-power system operations, such as powering industrial processes and basic chemical manufacturing. Many continuous processes require a constant or adjustable supply of power to operate at maximum efficiencies. In most locations this is difficult to provide with wind power without relying on some form of electricity storage or a grid connection to supplement wind power product during times of low wind speed. The majority of costs associated with a wind power plant are capital costs, with operational costs typically being very low, and hence the marginal cost of generation is near zero. This leads to a fundamentally different operating philosophy than thermal plants that have relatively high fuel, and hence marginal, costs. Wind power therefore often operates as must take power; meaning that any power it is capable of producing at a certain time is utilized. The low marginal cost of wind and fast start-up and shut-down times also dictates that wind is one of the first generators to be turned off when there are over-generation conditions or negative system prices. This is known as wind power curtailment and presents an opportunity for very low cost power from the wind producer, as it would otherwise not generate during those time periods. This excess power can then be used for energy storage applications, or as an inexpensive input into an energy conversion process that produces another valuable product, for example, electrolysis of water to produce hydrogen.

2 Wind Energy

2.1 Wind Resources

Wind resource assessment is the science of measuring and quantifying the quality of wind power potential at a particular location. This is a critical preliminary step in developing new wind power locations, and is necessary because wind resources can vary significantly even over relatively small geographic scales. Wind speeds can also have strong inter-annual variability, meaning that wind resource assessments should ideally be conducted over a multiple year span to provide more realistic potential power production estimates. An example of a national wind resource assessment from the United States, and the resulting wind map, is shown in Fig. 2. This map shows national 80 m land-based and offshore annual wind resource potential for the United States, in terms of mean annual wind speed.

The available energy from a wind resource varies appreciably, even inter-annually. The uncertainty in wind resource potential is significant, and therefore thorough wind resource assessment is necessary for determining long-term wind conditions at a location of interest to wind developers. Determining and forecasting long-term wind conditions serves two important objectives: (i) analyzing the quality of a wind plant site, and (ii) designing an optimum wind plant layout, including selecting appropriate turbine types for the site.

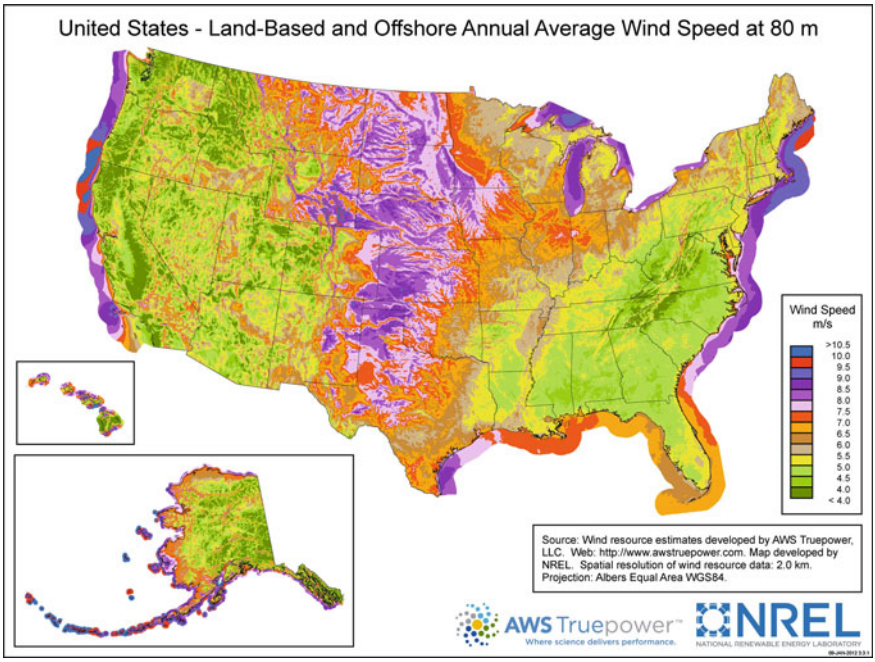


Fig. 2 Wind resource map at 80 m for the United States (NREL 2015)

A comprehensive wind resource assessment usually entails the following tasks (NYSERDA 2010; Rodrigo 2010; Drunsic 2012; Zhang et al. 2015):

- Site prospecting: identification of a suitable site using cartography (e.g., wind maps, political maps, power system transmission line maps, etc.).
- Measurement campaign: characterization of the on-site wind resource by recording the winds for 1–4 years continuously at or near hub height, with temporary meteorological masts (possibly completed or supplemented with remote sensing instruments).
- Microscale vertical extrapolation: transfer of the measurements to hub heights.
- Long-term extrapolation: extension of the measurements to the 10- to 30-year-long operating lifetime using historical observations (e.g., nearby tall towers, surface weather stations, and modeled data sets such as reanalysis datasets) and statistical methods.
- Wind plant layout design: establishment of turbine locations relative to wind resource estimation (e.g., using standard wake models or computational fluid dynamic models).
- Gross energy production estimation: calculation of the potential wind energy production over a year, or over the projected project lifetime, for the entire site.

- Energy losses assessment: evaluation of losses due to various causes (e.g., equipment scheduled and unscheduled maintenance, collection array losses, etc.).
- Uncertainty estimation: careful evaluation of the uncertainty associated with every step above. Physical uncertainties in wind energy may be broadly classified into long-term and short-term uncertainties. Long-term uncertainties are mainly introduced by (i) variation of wind conditions, (ii) turbine design, and (iii) other environmental, operational and financial factors. Short-term uncertainties are mainly introduced by boundary layer turbulence and other flow variations that occur over small time scales (order of minutes to hours).

Evaluating and reducing the uncertainties is of particular importance to secure financing and ensuring the wind plant investor's confidence. The uncertainties drive the probability distribution of the expected energy production. A recent comparison of pre-construction energy assessments for about 200 North American utility-scale wind plants to the actual power output highlighted the breakdown of contributions to the total energy production uncertainty (Fig. 3) (KEMA and Wind 2012). These values depend on many factors, including the project size, the topography complexity, and the availability of historical wind data.

Wind power density (WPD) is a useful way to evaluate the available wind resource at a potential site. WPD, measured in watts per square meter, indicates how much energy is available at the site. WPD (W/m^2) is a nonlinear function of the probability density function (*pdf*) of wind speed, which is expressed as:

$$\mathrm{WPD} = \frac{1}{2}\rho \int_0^{U_{\max}} U^3 f(U)\mathrm{d}U \tag{1}$$

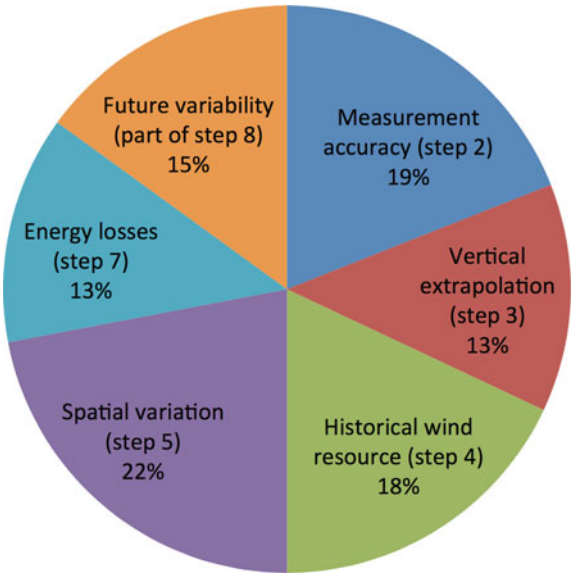


Fig. 3 Breakdown of contributions to the total energy production uncertainty (KEMA and Wind 2012)

where U represents the wind speed; $U_{\max}$ is the maximum possible wind speed at that location; ρ represents the air density; and $f(U)$ is the *pdf* of the wind speed.

Existing wind distribution modeling approaches can be broadly classified into: (i) univariate and unimodal distributions of wind speed (such as Weibull, Rayleigh, and Gamma distributions) (Carta et al. 2008), and (ii) bivariate and unimodal distributions of wind speed and wind direction (Vega 2008; Erdem and Shi 2011; Zhang et al. 2013). For univariate wind speed distribution, the most widely used is the 2-parameter Weibull distribution. Other distributions used to characterize wind speed include the 1-parameter Rayleigh distribution, 3-parameter generalized Gamma distribution, 2-parameter Lognormal distribution, 3-parameter Beta distribution, 2-parameter inverse Gaussian distribution, singly truncated normal Weibull mixture distribution, and the maximum entropy probability density function For joint distribution of wind speed and direction, there exists the Weibull distribution (Vega 2008), angular–linear, Farlie–Gumbel–Morgenstern, and anisotropic log-normal approaches (Erdem and Shi 2011), Multivariate and Multimodal Wind Distribution (MMWD) model (Zhang et al. 2013).

2.2 Wind Power Technology

The most common wind power technology today is the horizontal axis wind turbine. The utility-scale turbines that dominate the total wind energy production range in size from 500 kW to 5 MW, depending on application and location. Power output from these wind turbines is inherently variable and uncertain, as it corresponds very closely to the local meteorological conditions. Large wind power plants combine the output from up to hundreds of individual turbines. Each turbine consists of a number of subsystems, including: a rotor, drivetrain, nacelle, tower, and electrical system. A rotor consists of a hub and blades, with three blades having become the de facto standard, and is utilized to extract energy from the wind. The drivetrain is composed of shafts, couple, a gearbox, mechanical brake, and electrical generator. The geartrain steps up the shaft speed from the slow-turning rotor to the high speed required to drive the generator. Induction generators are used to convert the mechanical energy into electrical energy, and are typically asynchronous in modern turbines since maintaining the constant speed necessary for synchronous generation is difficult with variable wind speeds. Power electronics can be used to convert the DC power output to AC for grid injection, and are also used to regulate real and reactive power output.

Four basic types of utility-scale wind turbines appear widely in practice (Ellis et al. 2011): Type 1 (fixed speed), Type 2 (variable-slip), Type 3 (doubly-fed induction generator), and Type 4 (full converter). Many of the original utility-scale wind turbines were designed with fixed speed rotors (Type 1). These employed squirrel-cage AC induction machines with a direct grid connection, and many did not have the ability to feather their blades. While these machines were relatively reliable, the fixed speed rotor limited energy capture from the wind and required

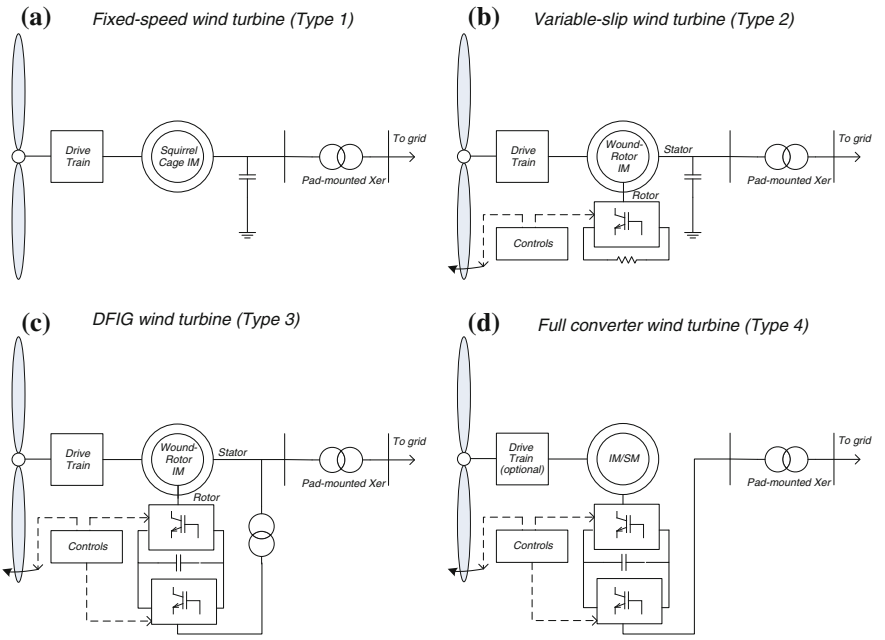


Fig. 4 Wind turbine technologies (Singh et al. 2014)

reactive power compensation, often in the form of capacitor banks. The other three turbine types were designed with variable speed rotors to allow for improved wind energy capture, and utilize blade pitching for this purpose. Variable-slip (Type 2) turbines offer a wider range of operation speed variations, but lose power because of the rotor resistance required to do so. Type 3 (Doubly-fed induction generator) turbines can recover some of this lost power using an AC/DC/AC converter. This also allows for the provision of both real and reactive power output. A Type 4 turbine (full converter) utilizes an AC/DC/AC converter as the sole connection to the grid. This disconnects the generator, whether synchronous or induction, from direct grid connection and allows for independent real and reactive power control. Figure 4 provides a schematic of the different wind technology types. While other wind technologies exist, the majority of installed wind turbines follow one of these basic four types. Minor variations of the technologies may be designed to provide unique capabilities needed for a particular installation.

2.3 Wind Power and Electricity System Integration

Wind power output is linked to the current meteorological conditions; a wind plant cannot produce more power at any time than the current wind conditions allow.

Most electricity systems operate so as to dispatch power generation to match variable and uncertain load, over which they often have very little control, and which is also driven by weather conditions. Thermal generators have historically been used to provide the generation necessary to meet demand, but they often have capacity factors of 90 % or greater when used for baseload power. As a comparison, wind generators have capacity factors ranging between 10–50 %, with those in wind rich regions with the most current turbine technology above 40 %. Thus the variability and uncertainty of wind power present a number of issues for its successful integration into different electricity systems.

A number of different technologies and operational changes have been proposed as means of better integrating wind power into power systems. Which option is best for a particular system is often determined by a number of factors including: the size of the power system, the wind plant locations, the wind penetration rate, and the balance of generation. Energy storage is an often proposed solution to reduce the impacts of wind power variability and uncertainty, and a review of using energy storage for wind power applications is provided in Diaz-Gonzalez et al. (2012). Wind power forecasting can help reduce the uncertainty in wind power output at multiple timescales useful for power system operations, and a review of methods is available in Foley et al. (2012a, b).

2.3.1 Wind Plant Power Generation

The power generated by a wind plant is an intricate function of the configuration and location of the individual wind turbines. The flow pattern inside a wind plant is complex, primarily because of the wake effects, and the highly turbulent flow. The engineering planning of a wind plant generally includes, but is not limited to, critical decision-making regarding: (1) the nameplate capacity of the wind plant; (2) the land area and land-shape of the plant site; (3) the type(s) of turbine to be installed; (4) the placement of turbines in the wind plant (farm layout). The key inputs and outputs of a generalized wind farm design methodology are illustrated in Fig. 5.

The annual energy production (AEP) of a wind plant in kWh (E_{plant}) at a particular location can be expressed as:

$$E_{\text{plant}} = (365 \times 24) \int_0^{360^\circ} \int_0^{U_{\max}} P_{\text{plant}}(U, \theta) p(U, \theta) \mathrm{d}U \mathrm{d}\theta \tag{2}$$

where $U_{\max}$ is the maximum possible wind speed at that location, $P_{\text{plant}}(U, \theta)$ represents the power generated by the plant (in kW) for a wind speed U and a wind direction θ, and $p(U, \theta)$ represents the probability of occurrence of a wind condition defined by speed U and direction θ. The *capacity factor* of a wind plant can be defined as the ratio of "the actual or expected output of the plant over a time period" and "the potential output if the plant was operating at full nameplate capacity

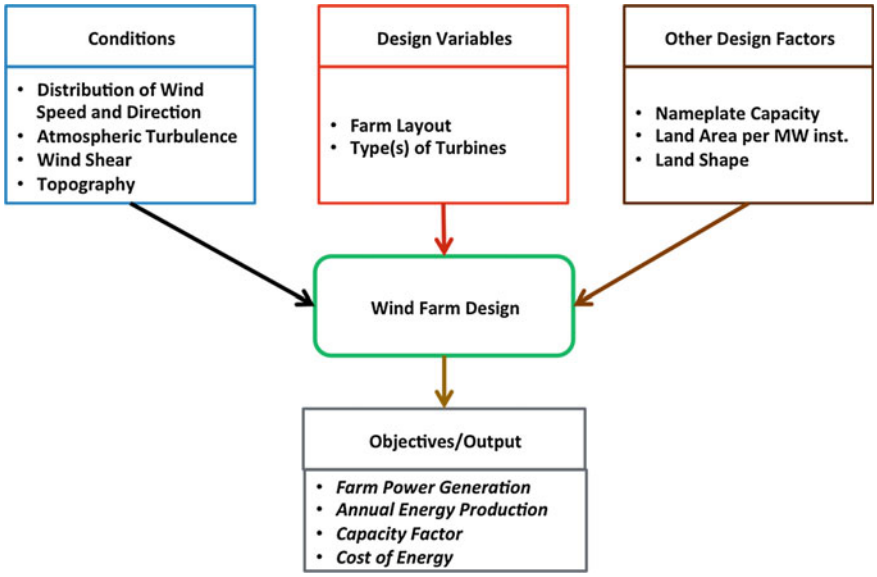


Fig. 5 Key factors (inputs and outputs) in wind plant design (Chowdhury et al. 2014)

throughout that time period". The annual wind plant *capacity factor* (CF) can be expressed in percentage as:

$$\mathrm{CF} = \frac{E_{\text{plant}}}{(365 \times 24)\sum_{j=1}^{N} P_{rj}} \times 100 \tag{3}$$

where P_{rj} represents the rated power of the jth turbine, and the expression $\sum_{j=1}^{N} P_{rj}$ represents the nameplate capacity of the wind plant.

The principal basis for estimating the conversion of wind speed to wind power is the wind power curve, examples of which can be seen in Fig. 6. Wind turbines generally have a cut-in speed, that is, a minimum speed at which power generation begins, generally around 4 m/s. The power output then increases very quickly with increasing wind speeds until full power output between 10–15 m/s. Full power output continues, with blade feathering to control rotor speeds, until a cut-out speed, after which no power is produced and the rotor is fixed to avoid excessive mechanical stress. However, in practice the conversion from wind speed to wind power has many sources of uncertainty, and is thus a more stochastic process because of turbine shading, measurement errors, and numerous other factors.

Energy deficits due to mutual shading effects can be determined using analytical or numerical wake models. These wake models provide a measure of the growth of the wake and the velocity deficit in the wake as functions of the distance downstream from a given wind turbine. The Park wake model, originally developed by Jensen (1983) and later by Katic et al. (1986), is one of the most popular analytical

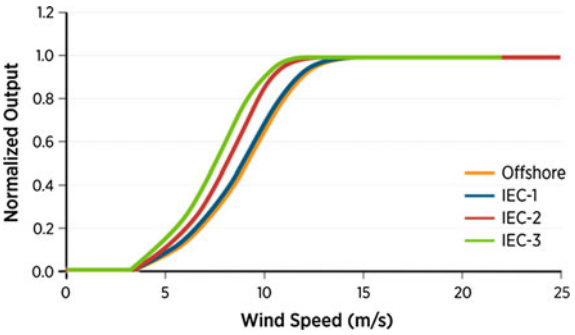


Fig. 6 Examples of power curves for different wind turbine classes (Draxl et al. 2015)

wake models used in wind plant modeling. Other standard analytical wake models include Frandsen's model (Frandsen et al. 2006), Larsen's model (Larsen et al. 1996), and Ishihara's model (Ishihara et al. 2004). To address this wake-induced energy deficiency, the arrangement of turbines over the plant site is optimized. Two primary classes of turbine arrangement (or layout optimization) methods exist in the literature: (i) methods that divide the wind plant into a discrete grid in order to search for the optimum grid locations of turbines (Grady et al. 2005; González et al. 2010; Şişbot et al. 2010), or (ii) more recent methods that define the turbine location co-ordinates as continuous variables, thereby allowing turbines to take up any feasible location within the plant (Kusiak and Song 2010; Chowdhury et al. 2012, 2013). Various classes of algorithms have also been leveraged for optimizing wind plant layouts, including evolutionary and genetic algorithms (Grady et al. 2005; González et al. 2010; Şişbot et al. 2010), swarm-based algorithms (Chowdhury et al. 2012, 2013), and pattern search algorithms (Du Pont and Cagan 2012).

The physical uncertainties affecting the performance of a wind plant are primarily of two types: (1) *short-term uncertainties:* uncertainties introduced by boundary layer turbulence and by other flow variations that occur over the minutes to hour timescale, as well as by short-lived extreme weather conditions; and (2) *long-term uncertainties:* uncertainties introduced by the long-term variations in wind conditions and by other environmental, operational, and financial factors. The physical sources of long-term uncertainties can be further broadly classified into the categories shown in Fig. 6 (Fig. 7).

2.3.2 Grid Integration of Wind Power

With increasing penetration of wind generation into interconnected power systems, system operators are faced with increased levels of variability and uncertainty. Bulk power system operations span a range of different timescales that can be summarized, from longest to shortest, as: unit commitment, load-following, economic dispatch, and regulation. Figure 8 shows a general power system load pattern for a single day. Unit commitment and scheduling are performed throughout the day to

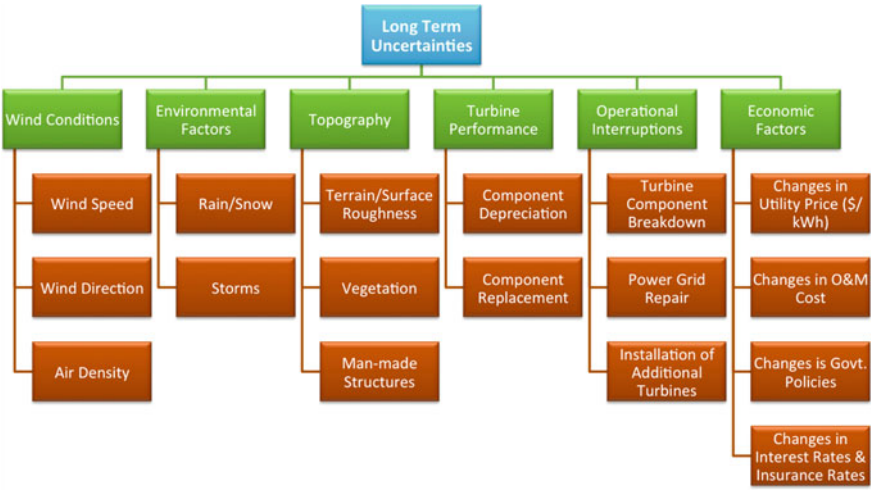


Fig. 7 Classification of the long-term uncertainties in wind energy (Messac et al. 2012)

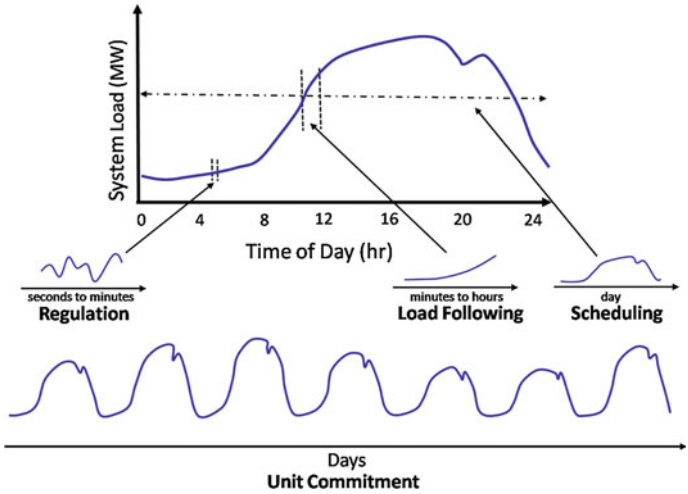


Fig. 8 Power system operation timescales (Modified from Ela et al. 2011)

economically commit the units in the system to meet the general system load pattern of the day. During shorter periods of time (minutes to hours), the system re-dispatches its units, and sometimes starts and stops quick-start units, to counteract deviations from the schedule through load following. Regulation is the balancing of fast second-to-second and minute-to-minute random variations in load or generation (Ela et al. 2011). These strategies represent the balancing during normal conditions of the power system. Wind power has a maximum upper availability

limit that is both variable and uncertain at multiple timescales, which increases the difficulty for power system operators to accurately schedule generation to meet demand. Thus, wind power forecasting plays an important part in power system operations, especially given the increasing amounts of wind power. Improved short-term wind power forecasts assist operators to make better day-ahead market operation and unit-commitment decisions, help real-time operations in the hour ahead, and warn operators about severe weather events.

The forecasts of load and renewables are never 100 % accurate and different types of reserves are used to help mitigate the effects of forecast errors. Typical ancillary service markets are: spinning contingency reserve, non-spinning contingency reserve, and regulating reserve. Variable generation such as wind can increase the requirements for normal balancing reserve, such as regulating reserve, which can increase the prices for those services. Improved wind power forecasting is expected to help reduce reserve requirements. In general, nuclear, biomass, and coal power plants must be committed in the day-ahead timeframe, while combined cycle and steam turbines can be committed in the 4-h-ahead or hour-ahead timeframe, respectively. Improved short-term wind forecasts allow operators make better DA market operation and unit-commitment decisions, help real-time operations in the hour ahead, and inform operators about extreme events.

Wind forecast models are commonly divided into three categories based on the data utilized (Foley et al. 2012a, b). Statistical forecasting is based on the analysis of historical time series of wind, whereas physics-based forecasting uses numerical weather prediction (NWP) models that may include statistical corrections. The first type of forecast model generally provides reasonable results in the estimation of long-term horizons, such as mean monthly, quarterly, and annual wind speed, or for very short timescales, such as less than 1 h (Monteiro et al. 2009). The impact of atmospheric dynamics becomes more important for short-term horizons of more than 1 h, such as day-ahead forecasts, and NWP models—such as the Advanced Regional Prediction System (ARPS) (Xue et al. 2000, 2001), the Weather Research and Forecasting Model (WRF) (Skamarock et al. 2005), Mesoscale Atmospheric Simulations System (MASS) (Manobianco et al. 1996), and Global Forecast System (Kanamitsu 1989)—often produce more accurate forecasts on these timescales. The third category is the combination of the physics-based and statistical forecasting, which generally consists of (i) an ensemble of high-resolution rapid-update NWP models, and (ii) a model output statistics (MOS) procedure to correct systematic errors of relevant NWP meteorological variables (e.g., wind speed and direction) at forecast sites.

Figure 9 shows an example of a hybrid physics-based and statistical forecasting system (Freedman et al. 2014), which includes: (i) the National Oceanic and Atmospheric Administration's 3-km High-Resolution Rapid Refresh model; (ii) three configurations of the ARPS; (ii) three configurations of the WRF model; and (iii) three configurations of the MASS. A MOS procedure was applied to the forecasts from each NWP system.

There are two types of forecasts in terms of information representation: deterministic and probabilistic. Deterministic forecasting provides look-ahead single

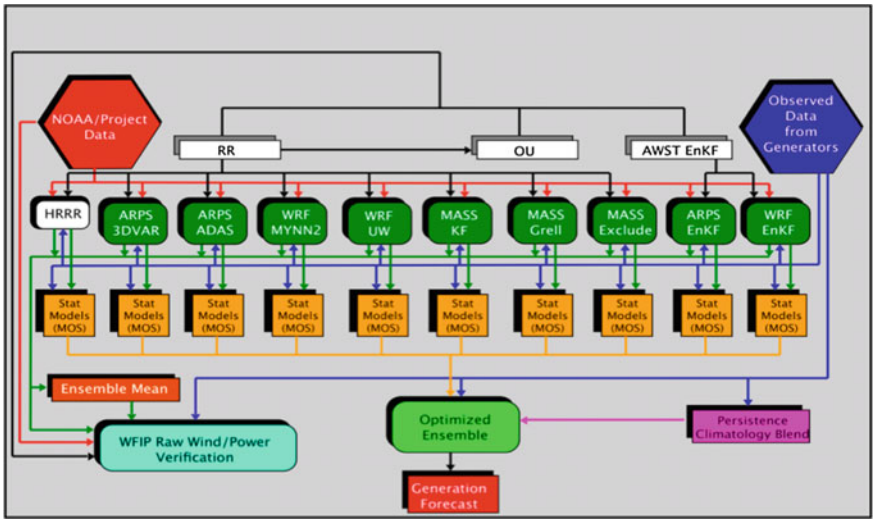


Fig. 9 Overall framework of the wind power forecasting system (Zhang et al. 2014)

point values at different forecast horizons, while probabilistic forecasting provides uncertainty information to the end user, such as probability of power generation scenarios, and confidence intervals for a given hour of the day, which can be quite useful for system operations. Deterministic wind power forecasting can be used in the day-ahead security constrained unit commitment, the real-time security constrained unit commitment, or the real-time security constrained economic dispatch in power system scheduling. With the probabilistic wind power forecasting, stochastic optimization can be applied to any of those same timeframes to gain improvements in reliability as compared to traditional scheduling methods (Wu et al. 2015).

Severe fluctuation incidents with large magnitudes and short durations, so-called "ramping events", are another significant concern for power system operators. Although the current power system is capable of handling small amounts of uncertainty and variability, with integrating a large amount of wind energy into the power grid, it is important to enhance the ability to handle large ramps in the power output. Moreover, forecasting and preparing for these ramping events requires not only determining the magnitude of wind ramps, but also their duration and timing.

Energy storage is another solution to reduce the impacts of wind power variability and uncertainty. Energy storage can be an important bridge to ensure that the power supply meets demand and to allow wind energy to serve as a continuous power source, while maintaining the frequency and voltages at close to normal values to ensure the integrity of customer loads. Typically, energy storage starts with electricity and converts it into some other form of energy that is stored and later converted back into electricity. Direct electricity storage stores ions or

electrons in a battery or capacitor. Other storage methods convert electric energy into some form of potential or mechanical energy, as in pumped storage hydropower, compressed-air energy storage, or flywheels.

3 Electrolytic Production of Hydrogen from Wind Energy

The energy produced from variable and uncertain renewable energy sources may also be readily stored in the chemical bonds of simple molecules, and later used as fuel. The simplest of these fuels is molecular hydrogen, which may be produced by the electrolysis of water. Water electrolysis utilizes electrical power to dissociate water into its constituent elements—hydrogen and oxygen. Other methods for the production of hydrogen, such as coal gasification, biomass processing, or natural gas reforming, do not involve electrical power as a primary process input. Hydrogen is a popular candidate medium for the storage and transport of renewably generated electricity because of its high gravimetric energy density and high autoignition temperature (Alazemi and Andrews 2015). Further, hydrogen produces only water upon combustion, and is the best fuel for most contemporary high-efficiency fuel cell systems (Ahluwalia and Wang 2005; Moore et al. 2006). Hydrogen-powered internal combustion engines operate with mechanical efficiencies superior to their gasoline-powered counterparts (Yamada and Mohamad 2010), and hydrogen is also a suitable replacement for current aviation fuels (Contreras et al. 1997). Modern-day electrolyzers may reach operating efficiencies of greater than 70 % (Dieguez et al. 2008; Brigulgio et al. 2013), although the capital costs for large-throughput electrolyzer systems remain high (Li et al. 2009).

Hydrogen fuel is the subject of numerous studies and reviews regarding sustainable energy economies and the firming of renewable energy sources (Agbossou et al. 2000; Turner 2004; Yilanci et al. 2009). In the case of pairing hydrogen production, storage, and transport with wind power, many case studies have been performed that take into account wind turbine, electrolyzer, fuel cell/engine generator set, and hydrogen storage system characteristics. The capital costs for wind-powered electrolyzer systems are dominated by turbine size and electrolyzer type and capacity, while the operating costs are largely controlled by the efficiency of the electrolyzer and the hydrogen compressors used in the storage systems (Linnemann and Steinberger-Wilckens 2007; Jorgensen and Ropenus 2008). Additionally, if a system is also tied to an external electricity grid, operating costs will be affected by the purchase of electricity from the grid to power an electrolyzer during a period of insufficient wind (Greiner et al. 2007). In-depth economic analyses of various wind-to-hydrogen system configurations indicate that grid connection of the system for the continuous production of hydrogen via electrolysis may reduce the impact of hydrogen storage system capital costs (Bartholmy 2005; Korpas and Greiner 2008), although even moderate advances in hydrogen storage technology could significantly reduce the cost of hydrogen for the end user (Menanteau et al. 2011). Other general concerns which determine the design of

wind-driven hydrogen production systems are the potential detrimental effects of variable operation on electrolyzer lifetime and efficiency, the relative capacity factors of the wind turbine and the electrolyzer system, and finally the method of hydrogen storage (Sherif et al. 2005). Pure hydrogen may be stored as a compressed gas, as a cryogenic liquid, or on metal surfaces in hydride form (Niaz et al. 2015). The drawbacks associated with each of these methods in their current states of development are major limitations in the adoption of molecular hydrogen as an energy vector. In the following section, another method of hydrogen storage is discussed: its incorporation into synthetic fuels.

4 Wind for the Production of Hydrogen-Bearing Fuels

Molecular hydrogen has the highest gravimetric (or specific) energy density of any chemical fuel. However, it must either be cryogenically liquefied or compressed to very high (<690 bar) pressures in order to approach the volumetric energy density of most commonly used liquid fuels. The problems associated with handling and storing hydrogen in its liquid state, or at extremely high pressure, are perhaps the greatest obstacles for the adoption of hydrogen as an energy storage medium (Niaz et al. 2015). In addition, the embrittlement of steel and other iron-based alloys by hydrogen (Bernstein 1970) indicates that hydrogen would not be easily incorporated into existing fuel storage and distribution systems, necessitating high infrastructure investments for widespread usage.

The limitations associated with hydrogen fuel have spurred interest in the storage of hydrogen as a part of hydrogen-rich fuels that are more easily stored in liquid form, and which are compatible with conventional fuel infrastructures. These fuels include hydrocarbons (alkanes, alcohols, ethers, etc.), boranes, and nitrogen/hydrogen compounds. Carbonaceous fuels may be synthesized by the traditional Fischer-Tropsch processing of hydrogen and carbon monoxide or the catalytic reaction of hydrogen with carbon dioxide. While the required hydrogen may still be obtained from water electrolysis, carbon monoxide and carbon dioxide may be renewably produced from biomass gasification or from the capture and conversion of atmospheric carbon dioxide. Among the boranes, presently only ammonia borane (NH_3BH_3) has received significant consideration for use as a hydrogen storage medium. However, significant technical difficulties related to the use of ammonia borane as a fuel remain—in particular, the sluggish rate of hydrogen release and poor processing economics for regeneration of the ammonia borane must be addressed (Bowden et al. 2008). Nitrogen/hydrogen compounds include ammonia (NH_3) and hydrazine (N_2H_4), the latter of which is dangerously unstable. For the purpose of this discussion, we will consider methanol as a typical synthetic hydrocarbon fuel, and anhydrous ammonia as the most developed carbon-free fuel candidate.

The synthesis of methanol from hydrogen and carbon monoxide or carbon dioxide is typically carried out over a copper/zinc oxide catalyst at pressures above 50 bar and temperatures in excess of 220 °C (Herman et al. 1979; Stiles 1987):

$$2H_2 + CO \leftrightarrow CH_3OH \tag{1}$$

$$3H_2 + CO_2 \leftrightarrow H_2O + CH_3OH \tag{2}$$

Both reactions (1) and (2) are exothermic, and the thermal sensitivity of the copper catalyst requires careful control of the reactor temperature. If wind-generated electricity is used for hydrogen generation via electrolysis of water, it may also be used to drive the electrical loads of gas compressors, cooling loops, and reactant gas preheaters. Methanol has many advantages as a hydrogen-bearing fuel; it is a liquid at room temperature and pressure, it may be used in internal combustion engines and gas turbines (Reed and Lerner 1973), as well as in direct methanol fuel cells (McGrath et al. 2004; DeLuca and Elabd 2006). Methanol shares many of the demonstrated advantages of other alcohol fuels versus gasoline, including having a higher octane rating and increased fire safety. The hygroscopic nature of methanol and its toxicity are two of the primary concerns regarding its potential use in fuel infrastructures. Using the bulk atmosphere as a CO_2 source for methanol synthesis is a poor prospect, considering its low concentration in the atmosphere on the whole. The practicality of producing methanol from electrolyzer hydrogen supplies and carbon dioxide sequestered from fossil fuel combustion has been shown to be limited mostly by the cost of the renewable electricity supply (Boretti 2013), with capital costs for wind energy making up the largest single contribution (Davis and Martin 2014). This is due in part to the high compressor and cooling loop loads demanded by methanol synthesis units (Soltanieh et al. 2012).

Another fuel that does not involve carbon for hydrogen fixation is ammonia (NH_3). Ammonia is produced in larger quantities than any other synthetic chemical (Blarigan 2000), and the production process requires only hydrogen (which may be sourced from renewable electricity-driven electrolysis) and nitrogen:

$$3H_2 + N_2 \leftrightarrow 2NH_3 \tag{3}$$

This reaction is only mildly exothermic, and has been proven industrially on very large scales; many modern ammonia production facilities generate more than 3000 metric tons per day. The small energy of formation for ammonia allows it to be readily reformed into a hydrogen-rich gas (75 % H_2 at 100 % conversion). Alternatively, molecular hydrogen may be produced from the ammonia by ammonia electrolysis (Vitse et al. 2005). The nitrogen required is sourced from air separation processes such as cryogenic distillation and pressure swing adsorption. Both of these methods require only air and electricity as process inputs. Once again, wind power is a suitable source of electricity (Tallaksen et al. 2015). Therefore, ammonia generation may be carried out using existing synthesis techniques with only water, air, and (renewable) electricity.

Ammonia has several advantages when compared to hydrogen fuel. First, it may be stored and transported as a liquid under moderate pressures (10 bar and above), or as a refrigerated liquid at atmospheric pressure (below −33 °C). Ammonia has a greater atomic hydrogen density than cryogenically liquefied hydrogen, and has a much narrower range of flammability in air: 16–25 % ammonia in air, compared to 4–75 % hydrogen in air. In addition, a stoichiometric ammonia–air mixture has a higher autoignition temperature (651 °C) than a hydrogen–air mixture (536 °C). While a large ammonia distribution and storage infrastructure related to its use as a fertilizer already exists, it is the toxicity of ammonia which has proved to be the most significant barrier to its adoption as a fuel. In spite of this, ammonia remains a popular fuel/energy storage medium for intermittent renewable energy systems (Morgan et al. 2014; Patil et al. 2014; Tallaksen et al. 2015).

Another interesting example of molecular hydrogen storage is the reversible hydrogenation/dehydrogenation of polyaromatic hydrocarbons (PAHs) (Teichmann et al. 2011, 2012; Dong et al. 2015). These systems would again require hydrogen from wind-powered electrolysis units. While these materials hold little value for use as direct transportation fuels when compared to the examples mentioned previously (due to the necessary storage of spent fuel), they do offer an interesting aspect of transporting and storing hydrogen somewhat densely as a thermally stable, low vapor pressure intermediate. Properly selected PAHs will be liquids at the appropriate temperatures during hydrogenation/dehydrogenation processing, and will provide high-purity hydrogen gas to energy systems. Further, many of these materials may store hydrogen at volumetric and gravimetric densities superior to those available in contemporary metal hydride systems (Malyshenko et al. 2015).

Hydrogen generated from wind power systems may also be used to upgrade hard-to-process carbonaceous fuel materials, such as sand oil bitumen and biomass solids or heavy oils (Samolada et al. 1998; Perez-Navarro et al. 2010; Olateju and Kumar 2011; Sharifzadeh et al. 2015). These processes, which generally employ catalytic pyrolysis of the raw material, aim to produce feedstock that is easily transported and refined into liquid fuels. The regions associated with the raw, unprocessed fuels are also often associated with stranded wind resources (Olateju and Kumar 2011). On-site generation of the required hydrogen gas with wind-generated electricity enables local pyrolysis processing of the fuel, resulting in a product that is lower in sulfur and oxygen, as well as of reduced viscosity for easier transfer via pipeline.

5 Conclusions

Wind energy has long been harnessed to produce mechanical work, and its expansion as a significant source of global electricity production has expanded enormously over the past decade. However, wind power's variable and uncertain nature, and the high costs of storing electricity, makes it a challenging primary input for alternative chemical and fuel production processes. This dictates that the energy produced by wind must often be converted to hydrogen by water electrolysis to be

used as a feedstock. From hydrogen the energy from wind can be used as a direct fuel source, as in most high-efficiency fuel cells, or as an input into producing other energy products, such as methanol or ammonia. Hydrogen from wind power can also be used in the upgrading of other fuels, such as biomass or heavy oils. Currently the capital costs of wind make up a very large portion of most of the described processes, and decreasing wind energy prices should make other chemical production routes more economically viable in the future. The variable and uncertain nature of wind energy is very localized in nature, meaning that for the time being the best application may be to offset the required electricity of these numerous processes with local wind, backed up by a reliable grid connection.

References

Agbossou, K., R. Chahine, J. Hamelin, F. Laurencelle, A. Anouar, J.-M. St. Arnaud and T. K. Bose (2000). "Renewable energy systems based on hydrogen for remote applications." Journal of Power Sources 96 168–172.

Ahluwalia, R. K. and X. Wang (2005). "Direct hydrogen fuel cell systems for hybrid vehicles." Journal of Power Sources 139: 152–164.

Alazemi, J. and J. Andrews (2015). "Automotive hydrogen fuelling stations: an international review." Renewable and Sustainable Energy Reviews 48: 483–499.

Bartholmy, O. (2005). "Renewable hydrogen from wind in California." Proceedings of the National Hydrogen Associations.

Bernstein, I. M. (1970). "The role of hydrogen in the embrittlement of iron and steel." Materials Science and Engineering 6: 1–19.

Bilgili, M., A. Ozbek, B. Sahin and A. Kahraman (2015). "An overview of renewable electric power capacity and progress in new technologies in the world." Renewable and Sustainable Energy Reviews 49: 323–334.

Blarigan, P. V. (2000). Advanced internal combustion enginer research. Proceedings of the DOE Hydrogen Program Review. Golden, CO, National Renewable Energy Laboratory.

Boretti, A. (2013). "Renewable hydrogen to recycle CO2 to methanol." International Journal of Hydrogen Energy 38: 1806–1812.

Bowden, M., T. Autrey, I. Brown and M. Ryan (2008). "The thermal decomposition of ammonia borane: a potential hydrogen storage material." Current Applied Physics 8: 498–500.

Brigulgio, N., G. Brunaccini, S. Siracusano, N. Randazzo, G. Dispenza, M. Ferraro, R. Ornelas, A. S. Arico and V. Antonucci (2013). "Design and testing of a compact PEM electrolyzer system." International Journal of Hydrogen Energy 38: 11519–11529.

Burton, T., D. Sharpe, N. Jenkins and E. Bossanyi (2001). Wind Energy Handbook, Wiley.

Carta, J. A., P. Ramirez and C. Bueno (2008). "A joint probability density function of wind speed, and direction for wind energy analysis." Energy Conversion and Management 49(6): 1309–1320.

Chowdhury, S., J. Zhang, A. Messac and L. Castillo (2012). "Unrestricted wind farm layout optimization (UWFLO): Investigating key factors influencing the maximum power generation." Renewable Energy 38(1): 16–30.

Chowdhury, S., J. Zhang, A. Messac and L. Castillo (2013). "Optimizing the arrangement and the selection of turbines for wind farms subject to varying wind conditions." Renewable Energy 52: 273–282.

Chowdhury, S., J. Zhang, W. Y. Tong and A. Messac (2014). "Modeling the Influence of Land-Shape on the Energy Production Potential of a Wind Farm Site." Journal of Energy Resources Technology-Transactions of the Asme 136(1).

Contreras, A., S. Yigit, K. Ozay and T. N. Veziroglu (1997). "Hydrogen as aviation fuel: a comparison with hydrocarbon fuels." International Journal of Hydrogen Energy 22: 1050–1060.

Davis, W. and M. Martin (2014). "Optimal year-round operation for methane production from CO2 and water using wind energy." Energy 69: 497–505.

DeLuca, N. W. and Y. A. Elabd (2006). "Polymer electrolyte membranes for the direct methanol fuel cell: a review." Journal of Polymer Science B: Polymer Physics 44: 2201–2225.

Diaz-Gonzalez, F., A. Sumper, O. Gomis-Bellmunt and R. Villafafila-Robles (2012). "A review of energy storage technologies for wind power applications." Renewable and Sustainable Energy Reviews 16(4): 2154–2171.

Dieguez, P. M., A. Ursua, P. Sanchis, C. Sopena, E. Guelbenzu and L. M. Gandia (2008). "Thermal performance of a commercial alkaline water electrolyzer: experimental study and mathematical modeling." International Journal of Hydrogen Energy 33: 7338–7354.

Dong, Y., M. Yang, Z. Yang, H. Ke and H. Cheng (2015). "Catalytic hydrogenation and dehydrogenation of N-ethylindole as a new heteroaromatic liquid organic hydrogen carrier." International Journal of Hydrogen Energy 40: 10918–10922.

Draxl, C., A. Clifton, B. Hodge and J. McCaa (2015). "The Wind Integration National Dataset (WIND) Toolkit." Applied Energy 151: 355–366.

Drunsic, M. (2012). Accurate wind assessments: pinning down performance.

Du Pont, B. L. and J. Cagan (2012). "An extended pattern search approach to wind farm layout optimization." Journal of Mechanical Design 134(8): 081002.

Ela, E., M. Milligan and B. Kirby (2011). Operating reserves and variable generation. National Renewable Energy Laboratory.

Ellis, A., Y. Kazachkov, E. Muljadi, P. Pourbeik and J. J. Sanchez-Gasca (2011). Description and technical specifications for generic WTG models - A status report. Power Systems Conference and Exposition.

Erdem, E. and J. Shi (2011). "Comparison of bivariate distribution construction approaches for analysing wind speed and direction data." Wind Energy 14(1): 27–41.

EWEA (2014). Wind in power: 2013 European statistics, European Wind Energy Association.

Foley, A., P. Leahy, A. Marvuglia and E. McKeogh (2012). "Current methods and advances in forecasting of wind power generation." Renewable Energy 37(1): 1–8.

Foley, A. M., P. G. Leahy, A. Marvuglia and E. J. McKeogh (2012). "Current methods and advances in forecasting of wind power generation." Renewable Energy 37(1): 1–8.

Frandsen, S., R. Barthelmie, S. Pryor, O. Rathmann, S. Larsen, J. Højstrup and M. Thøgersen (2006). "Analytical modelling of wind speed deficit in large offshore wind farms." Wind Energy 9(1–2): 39–53.

Freedman, J. M., J. Manobianco, J. Schroeder, B. Ancell, K. Brewster, S. Basu, V. Banunarayanan, B.-M. Hodge and I. Flores (2014). The Wind Forecast Improvement Project (WFIP): A Public/Private Partnership for Improving Short Term Wind Energy Forecasts and Quantifying the Benefits of Utility Operations–the Southern Study Area, AWS Truepower, LLC.

González, J. S., A. G. G. Rodriguez, J. C. Mora, J. R. Santos and M. B. Payan (2010). "Optimization of wind farm turbines layout using an evolutive algorithm." Renewable Energy 35(8): 1671–1681.

Grady, S., M. Hussaini and M. M. Abdullah (2005). "Placement of wind turbines using genetic algorithms." Renewable energy 30(2): 259–270.

Greiner, C. J., M. Korpas and A. Holen (2007). "A Norwegian case study on the production of hydrogen from wind power." International Journal of Hydrogen Energy 32: 1500–1506.

Herman, R. G., K. Klier, G. W. Simmons, B. P. Finn, J. B. Bulko and T. P. Kobylinski (1979). "Catalytic synthesis of methanol from CO and H2: I. Phase composition, electronic properties, and activities of the Cu/ZnO/M2O3 catalysts." Journal of Catalysis 56: 407–429.

Hodge, B.-M., A. Florita, K. Orwig, D. Lew and M. Milligan (2012). A Comparison of Wind Power and Load Forecasting Error Distributions. World Renewable Energy Forum. Denver, CO.

Hurlbut, D., J. McLaren and R. Gelman (2013). Beyond Renewable Portfolio Standars: An Assessment of Regional Supply and Demand Conditions Affecting the Future of Renewable Energy in the West. Golden, CO, National Renewable Energy Laboratory.

IEA (2014). IEA Wind: 2013 Annual Report, International Energy Agency.

Ishihara, T., A. Yamaguchi and Y. Fujino (2004). "Development of a new wake model based on a wind tunnel experiment." Global wind power.

Jensen, N. (1983). "A note on wind turbine interaction." Risoe National Laboratory, Roskilde, Denmark, Technical Report No. M-2411.

Jorgensen, C. and S. Ropenus (2008). "Production price of hydrogen from grid connected electrolysis in a power market with high wind penetration." International Journal of Hydrogen Energy 33: 1665–1676.

Kanamitsu, M. (1989). "Description of the NMC global data assimilation and forecast system." Weather and Forecasting 4(3): 335–342.

Katic, I., J. Højstrup and N. Jensen (1986). A simple model for cluster efficiency. European Wind Energy Association Conference and Exhibition.

KEMA, D. and S. Wind (2012). Reducing uncertainty in wind project energy estimates: a cost-benefit analysis of additional measurement campaign methods.

Korpas, M. and C. J. Greiner (2008). "Opportunities for hydrogen production in connection with wind power in weak grids." Renewable Energy 33: 1199–208.

Kusiak, A. and Z. Song (2010). "Design of wind farm layout for maximum wind energy capture." Renewable Energy 35(3): 685–694.

Lantz, E., M. Hand and R. Wiser (2012). The Past and Future Cost of Wind Energy. World Renewable Energy Forum. Denver, CO USA.

Larsen, G. C., J. Højstrup and H. Aagaard Madsen (1996). Wind fields in wakes. 1996 European Wind Energy Conference and Exhibition.

Li, C. H., X. J. Zhu, G. Y. Cao, S. Sui and M. R. Hu (2009). "Dynamic modeling and sizing optimization of stand-alone photovoltaic power systems using hybrid energy storage technology." Renewable Energy 34: 815–826.

Linnemann, J. and R. Steinberger-Wilckens (2007). "Realistic costs of wind-hydrogen vehicle fuel production." International Journal of Hydrogen Energy 32: 1492–1499.

Malyshenko, S., S. Mitrokhin and I. Romanov (2015). "Effects of scaling in metal hydride materials for hydrogen storage and compression." Journal of Alloys and Compounds 645: S84–S88.

Manobianco, J., J. W. Zack and G. E. Taylor (1996). "Workstation-based real-time mesoscale modeling designed for weather support to operations at the Kennedy Space Center and Cape Canaveral Air Station." Bulletin of the American Meteorological Society 77(4): 653–672.

Manwell, J. F., J. G. McGowan and A. L. Rogers (2002). Wind Energy Explained: Theory, Design and Application, Wiley.

McGrath, K. M., G. K. S. Prakash and G. A. Olah (2004). "Direct methanol fuel cells." Journal of Industrial Engineering and Chemistry 10: 1063–1080.

Menanteau, P., M. M. Quemere, A. Le Duigou and S. Le Bastard (2011). "An economic analysis of the production of hydrogen from wind-generated electricity for use in transport applications." Energy Policy 39: 2957–2965.

Messac, A., S. Chowdhury and J. Zhang (2012). "Characterizing and mitigating the wind resource-based uncertainty in farm performance." Journal of Turbulence 13(13): 1–26.

Monteiro, C., R. Bessa, V. Miranda, A. Botterud, J. Wang and G. Conzelmann (2009). Wind power forecasting: state-of-the-art 2009, Argonne National Laboratory (ANL).

Moore, R. M., K. H. Hauer, S. Ramaswamy and J. M. Cunningham (2006). "Energy utilization and efficiency analysis for hydrogen fuel cell vehicles." Journal of Power Sources 159: 1214–1230.

Morgan, E., J. F. Manwell and J. G. McGowan (2014). "Wind-powered ammonia fuel production for remote islands: a case study." Renewable Energy 72: 51–61.

Niaz, S., T. Manzoor and A. H. Pandith (2015). "Hydrogen storage: Materials, methods and perspectives." Renewable and Sustainable Energy Reviews 50: 457–469.

NREL. (2015). 2015, from http://www.nrel.gov/.
NYSERDA (2010). Wind resource assessment handbook, New York state energy research and development authority.
Olateju, B. and A. Kumar (2011). "Hydrogen production from wind energy in Western Canada for upgrading bitumen from oil sands." Energy 36(11): 6326–6339.
Patil, A., L. Laumans and H. Vrijenhoef (2014). "Solar to ammonia - via Proton's NFuel units." Procedia Engineering 83: 322–326.
Perez-Navarro, A., D. Alfonso, C. Alvarez, F. Ibanez, C. Sanchez and I. Segura (2010). "Hybrid biomass-wind power plants for reliable energy generation." Renewable Energy 35: 1436–1443.
Putnam, P. C. (1948). Power From the Wind, Van Nostrand.
Reed, T. B. and R. M. Lerner (1973). "Methanol: a versatile fuel for immediate use." Science 182: 1299–1304.
Rodrigo, J. S. (2010). State-of-the-Art of Wind Resource Assessment CENER National Renewable Energy Centre.
Samolada, M., W. Baldauf and I. Vasalos (1998). "Production of a bio-gasoline by upgrading biomass flash pyrolysis liquids via hydrogen processing and catalytic cracking." Fuel 77: 1667–1675.
Sharifzadeh, M., C. Richard, K. Liu, K. Hellgardt, D. Chadwick and N. Shah (2015). "An integrated process for biomass pyrolysis oil upgrading: A synergistic approach." Biomass and Bioenergy 76: 108–116.
Sherif, S. A., F. Barbir and T. N. Veziroglu (2005). "Wind energy and the hydrogen economy - review of the technology." Solar Energy 78: 647–660.
Singh, M., E. Muljadi, J. Jonkman, V. Gevorgian, I. Girsang and J. Dhupia (2014). Simulation for Wind Turbine Generators - With FAST and MATLAB-Simulink Modules. Golden, CO, National Renewable Energy Laboratory.
Şişbot, S., Ö. Turgut, M. Tunç and Ü. Çamdalı (2010). "Optimal positioning of wind turbines on Gökçeada using multi-objective genetic algorithm." Wind Energy 13(4): 297–306.
Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang and J. G. Powers (2005). A description of the advanced research WRF version 2, DTIC Document.
Soltanieh, M., K. M. Azar and M. Saber (2012). "Development of a zero emission integrated system for co-production of electricity and methanol through renewable hydrogen and CO2 capture." International Journal of Greenhouse Gas Control 7: 145–152.
Stiles, A. V. (1987). "Methanol, past, present, and speculation on the future." Journal of the American Institute of Chemical Engineers 23: 362–375.
Tallaksen, J., F. Bauer, C. Hulteberg, M. Reese and S. Ahlgren (2015). "Nitrogen fertilizers manufactured using wind power: greenhouse gas and energy balance of community-scale ammonia production." Journal of Cleaner Production 107: 626–635.
Teichmann, D., W. Arlt, P. Wasserscheid and R. Freymann (2011). "A future energy supply based on Liquid Organic Hydrogen Carriers (LOHC)." Energy & Environmental Science 4: 2767–2773.
Teichmann, D., W. Arlt and P. Wasserscheid (2012). "Liquid Organic Hydrogen Carriers as an efficient vetor for the transport and storage of renewable energy." International Journal of Hydrogen Energy 37: 18118–18132.
Turner, J. A. (2004). "Sustainable Hydrogen Production." Science 305: 972–974.
Vega, R. (2008). Wind directionality: A reliability-based approach, PhD thesis, Texas Tech University, Lubbock, TX, August.
Vitse, F., M. Cooper and G. G. Botte (2005). "On the use of ammonia electrolysis for hydrogen production." Journal of Power Sources 142: 18–26.
Wiser, R., G. Barbose and E. Holt (2011). "Supporting solar power in renewables portfolio standards: Experience from the United States." Energy Policy 39(7): 3894–3905.
Wu, H., E. Ela, I. Krad, A. Florita, J. Zhang, B. Hodge, E. Ibanez and W. Gao (2015). "An Assessment of the Impact of Stochastic Day-Ahead SCUC on Economic and Reliability Metrics at Multiple Timescales."

Xue, M., K. K. Droegemeier and V. Wong (2000). "The Advanced Regional Prediction System (ARPS) - A multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: Model dynamics and verification." Meteorology and Atmospheric Physics 75(3–4): 161–193.

Xue, M., K. K. Droegemeier, V. Wong, A. Shapiro, K. Brewster, F. Carr, D. Weber, Y. Liu and D. Wang (2001). "The Advanced Regional Prediction System (ARPS) - A multi-scale nonhydrostatic atmospheric simulation and prediction tool. part II: Model physics and applications." Meteorology and Atmospheric Physics 76(3–4): 143–165.

Yamada, N. and M. N. A. Mohamad (2010). "Efficiency of hydrogen internal combustion engine combined with open steam Rankine cycle recovering water and waste heat." International Journal of Hydrogen Energy 35: 1430–1442.

Yilanci, A., I. Dincer and H. K. Ozturk (2009). "A review on solar-hydrogen/fuel cell hybrid energy systems for stationary applications." Progress in Energy Combustion and Science 35: 231–244.

Zhang, J., S. Chowdhury, A. Messac and L. Castillo (2013). "A Multivariate and Multimodal Wind Distribution model." Renewable Energy 51: 436–446.

Zhang, J., A. Florita, B.-M. Hodge and J. Freedman (2014). Ramp Forecasting Performance From Improved Short-Term Wind Power Forecasting. ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers.

Zhang, J., C. Draxl, T. Hopson, L. Delle Monache, E. Vanvyve and B.-M. Hodge (2015). "Comparison of numerical weather prediction based deterministic and probabilistic wind resource assessment methods." Applied Energy 156: 528–541.

Solar Energy as Source for Power and Chemicals

Lidia Martín, Borja Hernández and Mariano Martín

Abstract In this chapter we present the different technologies that are currently used to capture solar energy and their application to power and chemicals production as a way to store it into handy fuels. We describe solar photovoltaics (PV solar), solar thermal, and solar thermoelectric. Once solar energy is converted in the form of power or heat, we use it for the production of chemicals or evaluate the performance of CSP plants comparing different cooling technologies using a system approach by presenting the results of modeling and optimizing the operation of several processes.

1 Introduction: Solar Energy

The Sun is a G-type star whose core can reach temperatures near 15 million degrees Kelvin. This temperature is reached due to exothermic fusion reactions where, usually, hydrogen is transformed into helium. Its energy is emitted as electromagnetic radiation, where the power could be estimated using a surface temperature of 5770 K to obtain a value of 3.84×10^{26} W (NASA 2015a,b; Coley 2008). Only a small fraction of this energy reaches and is captured by the Earth. However, its potential is very important since solar energy is the main energy resource in our planet, while other sources have stored it such as biomass, wind, or waves.

Commonly, the energy received on the top of the atmosphere per square meter of perpendicular surface to rays, is assumed to be a constant value, 1366 kW/m^2, and is known as Solar Constant. Out of this energy, approximately 20 % is absorbed by stratosphere and troposphere, 30 % is reflected by clouds, the surface and other atmospheric layers and 50 % reaches the land and the oceans, see Fig. 1.

L. Martín · B. Hernández · M. Martín (✉)
Department of Chemical Engineering, University of Salamanca,
37007 Salamanca, Spain
e-mail: mariano.m3@usal.es

M. Martín (ed.), *Alternative Energy Sources and Technologies*,
DOI 10.1007/978-3-319-28752-2_7

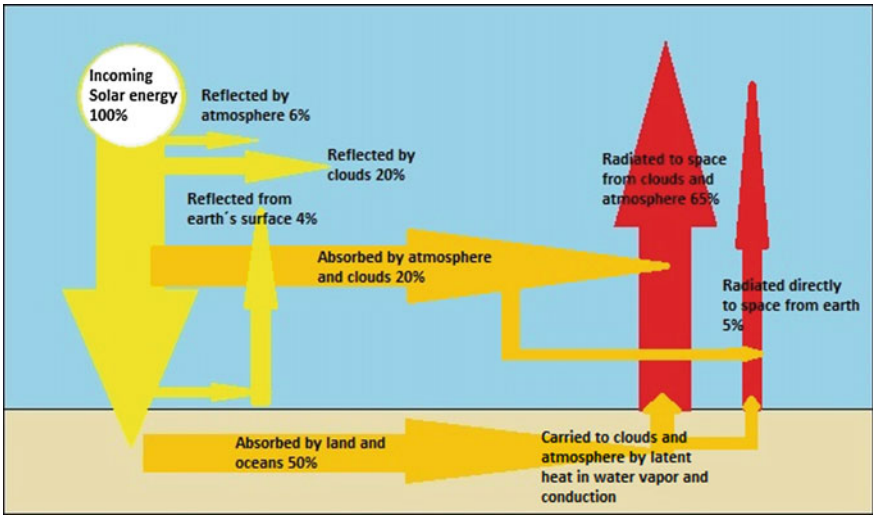


Fig. 1 Earth's energy budget (NASA 2015c)

The incident energy over the land has a heterogeneous distribution around the globe. Some factors are significant for that distribution such as latitude, climate conditions, landform, and season of the year. In addition, these factors will have a notable influence in design variables of capture systems such as orientation and inclination angle. In the map given by Fig. 2 the direct normal irradiation distribution around the planet is presented.

Incident radiation is commonly classified in two groups, direct and diffuse radiation. The first one is received by the receptor directly, whereas the second one is received after it has been reflected and/or turned aside by water, dust, air, etc. (Coley et al. 2008). The daily diffuse radiation is defined by Erbs correlations (Tham et al. 2010) as follows:

- For sunset hour angle ≤81.4°
  - If: $K_T < 0.715$

$$\frac{I_d}{I} = 1 - 0.2727\,K_T + 2.4495\,K_T^2 - 11.951\,K_T^3 + 9.3879\,K_T^4 \qquad (1)$$

  - If: $K_T \geq 0.715$

$$\frac{I_d}{I} = 0.143$$

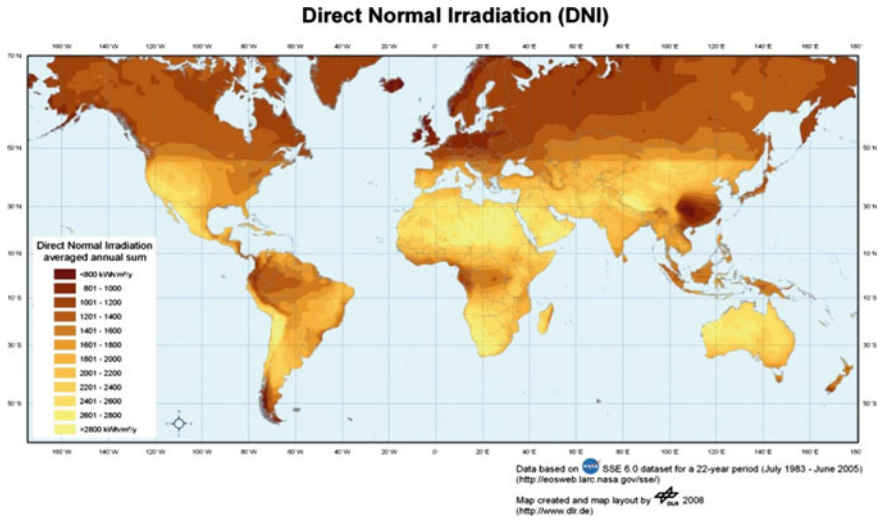


Fig. 2 Solar map (Trieb et al. 2008 with permission)

- For sunset hour angle >81.4°
  - If: $K_T < 0.715$

$$\frac{I_d}{I} = 1 + 0.2832\,K_T - 2.5557\,K_T^2 + 0.8448\,K_T^3 \tag{2}$$

  - If: $K_T \geq 0.715$

$$\frac{I_d}{I} = 0.175$$

where I is the daily irradiation that can be estimated using clearness index, K_T, following the studies of Liu and Jordan (1960), where K_T is also defined. However, K_T can be estimated using Hargreaves method too. This can only be used if the power plant is not in a wet climate. According to the method, the clearness index is defined as follows:

$$K_T = k_{\mathrm{rs}}(T_{\max} - T_{\min})^{0.5} \tag{3}$$

where k_{rs} is a coefficient with the value of 0.16 $\mathrm{K}^{-0.5}$ for interior locations and 0.19 $\mathrm{K}^{-0.5}$ for coastal locations.

In order to acuratelly determine the incident solar radiation in a specific place, experimental measure is necessary. The most commonly used devices to perform those measurements are:

- Pyranometer: It measures global incident radiation, direct plus diffuse.
- Pyroheliometer: It only measures direct incident radiation. It needs to be moved following the Sun.

While we can measure the energy received by the Earth's surface, when trying to use it as a source of energy, shadows may represent an important loss of efficiency. Therefore, the design of capture systems requires the evaluation of the effect of shadows on the energy received by the receptor during the period of a year. The Sun position is typically represented in diagrams that use orthographic, stereographic, gnomic, or cylindrical projections. The last ones are the most commonly used.

The shadows projected by an obstacle could be determined using geometric relationships. There are some simple expressions that allow calculating the minimum separation between photovoltaic panel as given by IDAE (2011), where an estimation of a minimum of 4 h of solar energy per day is captured by panels:

$$d = h/\tan(61^\circ - \text{latitude}) \tag{4}$$

However, a common systematic method used to determine available energy in a year is based on using cylindrical diagrams combined with a measure of captured energy in each position of the Sun during a day. The diagram is divided in sectors. Each of them has the amount of energy or the percentage of the annual energy received. If an obstacle exists in a certain position, this section does not produce energy. The total amount of energy is computed as the sum of the sections. As shown in Fig. 1, other loses of efficiency can exist and the total percentage will not be 100 %. There are some other diagrams to present this information such as those presented by Masters (2004) or IDAE (2011).

An example is shown in Table 1 and Fig. 3. The panel has an inclination of 35°, $\beta = 35$, and azimuth angle of 0°, $\alpha = 0°$. The percentage of energy losses due to shadows can be computed as follows:

$$\begin{aligned}\sum_{j=A}^{D}\sum_{i=1}^{14} \%\text{shadows losses} = {} & 0.5 \cdot A5 + 0.75 \cdot A6 + A8 + 0.25 \cdot A10 + 0.25 \\ & \cdot B4 + B6 + 0.5 \cdot B8 + 0.25 \cdot C6 \\ = {} & 6.16\,\%\end{aligned} \tag{5}$$

Furthermore, especially in PV cells, some authors have modeled the oscillations produced by cells under total or partial shadow conditions as it is showed in the experimental model given by Martinez-Moreno et al. (2010). The effect of these conditions in the main variables of a cell has also been studied and modeled to reduce undesired such as mismatch effect (Anani et al. 2013; Karatepe et al. 2007).

Table 1 Percentage of received energy for a latitude of 40° (IDAE 2011)

β = 35° α = 0°	A	B	C	D
13	0.00	0.00	0.00	0.03
11	0.00	0.01	0.12	0.44
9	0.13	0.41	0.62	1.49
7	1.00	0.95	1.27	2.76
5	1.84	1.50	1.83	3.87
3	2.70	1.88	2.21	4.67
1	3.15	2.12	2.43	5.04
2	3.17	2.12	2.33	4.99
4	2.70	1.89	2.01	4.46
6	1.79	1.51	1.65	3.63
8	0.98	0.99	1.08	2.55
10	0.11	0.42	0.52	1.33
12	0.00	0.02	0.10	0.40
14	0.00	0.00	0.00	0.02

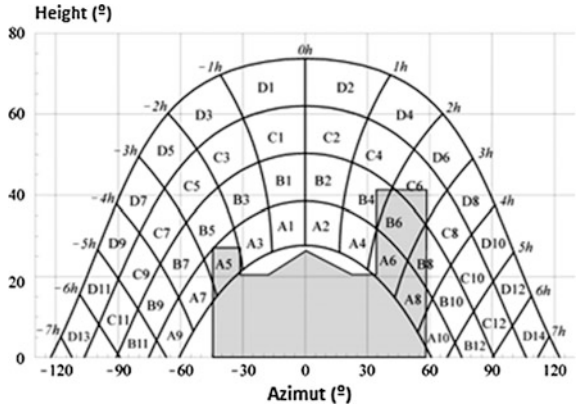


Fig. 3 Sections of captured energy for a latitude of 40° (reproduced with permission from IDAE 2011)

2 Types of Solar Capturing Technologies

2.1 PV Solar

2.1.1 Solar Gathering

PV solar is a type of solar energy based on photovoltaic (PV) effect. It happens when electrons, receiving the incident energy of photons, have the necessary energy potential to reach the conduction band, according to the valence–conduction band model. Commonly, a p–n connection is used in order to help the creation of an electric field in the cell.

Cells are usually characterized by some parameters such short circuit current (I_{SC}), open-circuit potential (V_{OC}), maximum power (P_{max}), form factor (F_F), and efficiency (η). Form factor is a constant, with a maximum value of 1 (Luque and Hegedus 2003). Thus, the power can be computed as follows:

$$P_{max} = I_{SC} \cdot V_{OC} \cdot F_F \tag{4}$$

where, the efficiency is defined as the ratio between the power produced ($P_{produced}$) and that received ($P_{received}$), where A is the area and $P_{incident}$ is the power per area unit:

$$\eta(\%) = \frac{P_{produced}}{P_{received}} = \frac{I_{SC} \cdot V_{OC} \cdot F_F}{A \cdot P_{incident}} \cdot 100 \tag{5}$$

In Fig. 4, a power curve diagram, these parameters are shown.

For electrical applications, PV cells are commonly represented as it is showed in Fig. 5, where I_{ph} is the current in the cell and, according to the power curve diagram, it has approximately the same value as I_{SC}. I_d is the reverse saturation current of diode. R_{sh} and R_s are the intrinsic shunt and series resistance of the cell respectively. Usually, the value of R_{sh} is very large if we compare it with R_s. Finally, V is the voltage produced by the cell which can give energy to feed a determinate resistance, R_d. Using this circuit and some expressions for each parameter, a modeling of the system can be done as it is given by Singh (2013).

On the one hand, PV cells can be classified according to the material that is used (Parida et al. 2011):

- Silicon
  - It is the most used material, and it can be classified in two main groups:

    Crystallized silicon:

    - Monocrystalline: It is the most common type in the market. It is the particular structure that provides a higher efficiency with silicon.
    - Polycrystalline: It is obtained using low-prize silicon. As a consequence, its efficiency is lower, but the price per kW is also lower.

    Amorphous silicon. It has a low production cost. However, it is sensitive to deterioration by radiation over time. On the other hand, it offers the possibility to create "thin films."
- Other materials. These ones can be divided into:
  - Materials used in current commercial cells such cells of cuprum-indium/gallium-di-selenide, cadmium tellurium, gallium arsenide, or titanium dioxide (O'Regan and Grätzel 1991).
  - New materials in development such as graphene (Liu et al. 2008), silicene (Ye et al. 2014), di-sulfide spices (Lee 2011) or perovskite (Gong et al. 2015).

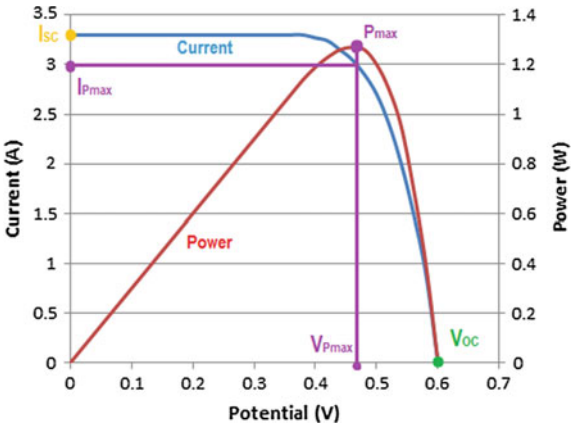


Fig. 4 Power and current curves of a regular PV cell

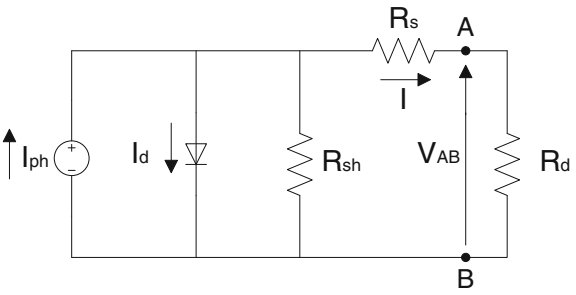


Fig. 5 Equivalent circuit of a cell

On the other hand, they can be classified according to the use required:

- PV cells in panels which are the most common ones to capture energy in houses or lands.
- "Thin film" cells for applications where the weight is important.
- PV concentration cells using Fresnel lenses.

2.1.2 Power Production

In order to obtain an important amount of energy using PV energy, an installation as presented in Fig. 6 is required.

The structure of the installation is composed of the following elements:

- PV cells that we have been described above.
- Controller. It helps to protect the battery in case of overload due to the different voltage emitted by the cell, when the charge required by the consumer is more

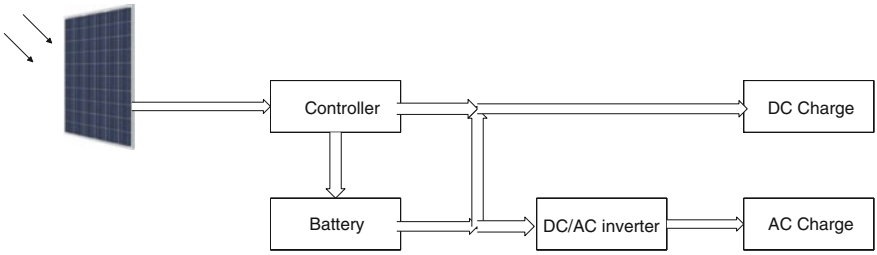


Fig. 6 PV installation diagram

than the system is able to provide or due to fails caused in the connection/disconnection of batteries.
- Batteries are used to store energy so as to it during a period when the PV cells do not work.
- DC/AC inverter. It maximizes the production of energy and transforms the produced energy from DC to AC.

Other optional equipment in PV installations is the transformers to adapt the AC to the requirements of the system or the electrical grid, and an electric meter that can measure the current in both senses, from the cells to the grid and reverse.

2.2 Solar Thermal

2.2.1 Solar Gathering

Solar thermal energy is based on using the sunlight to heat up a heat transfer fluid (HTF). The maximum temperature that the HTF reaches is lower than 373 K (López Cózar 2006). Thus, the technology is also known as low-concentration solar power. The system used to capture solar radiation consists of:

- Plane collectors or solar panels that capture sunlight with the same intensity as the one that reaches them not allowing its escape, greenhouse effect. They are made of, as seen in Fig. 7.
- Transparent cover.
- Absorbent surface or film.
- Hydraulic circuit for the HTF below the plate.
- Thermal isolation film.
- Protective case.

The system has to be resistant to extreme conditions, high and low temperatures, efficient in converting energy and they must last long.

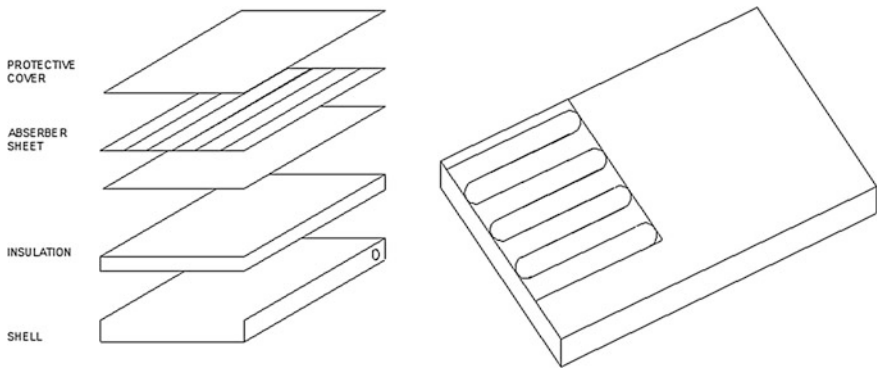


Fig. 7 Solar thermal capture system

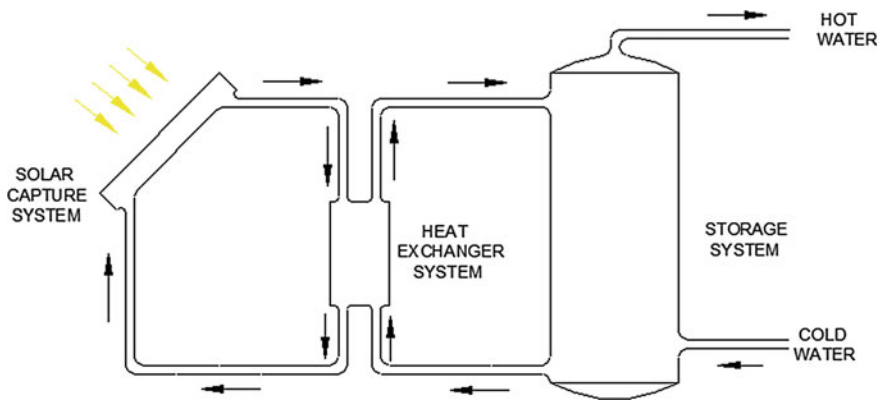


Fig. 8 Solar thermal scheme

2.2.2 Heat Production

Low temperature solar thermal technologies are used for water heating below its boiling point (Hudon et al. 2012; López Cózar 2006). Figure 8 shows a simple scheme of the system.

The systems consist of three elements:

- *Solar capture system*: Consisting of collectors as presented above. It is the primary circuit for the HTF.
- *Heat exchanger system*: Isolates the HFT from the water to be used or consumed.
- *Storage system*: It consists of stainless steel tanks that store the hot water to be used on demand. Secondary circuit.

2.3 *Solar Thermoelectric or Concentrated Solar Power (CSP)*

Thermoelectric solar energy is also known as concentrated solar power (CSP). Unlike solar thermal, it is used to produce power from concentrating solar radiation (Zhang et al. 2013; IEA 2010).

The basic principle consists of reflecting solar radiation toward a receiver where a heat transfer fluid is heated up (primary circuit). The high temperature of this fluid allows producing steam that will eventually produce power in a turbine (IEA 2010; López Cózar 2006; Ruíz Hernández 2009). Thus, the system consists of four sections, namely solar field and receiver, heat transfer train, power block, and cooling. Furthermore, the variability in solar availability typically suggests the addition of a storage system and/or the use of services or fossil fuels to provide energy in case there is no enough solar radiation. Solar thermoelectric systems can be divided into two, as a function of the temperature of the HTF (López Cózar 2006):

- Medium temperature CSP: 353 K < T < 673 K
- High temperature CSP: For temperatures above 673 K

Figure 9 shows a basic scheme of the process, although it will be further explained later in the chapter.

2.3.1 Solar Gathering

Solar energy is captured by means of a system that consists of three elements (IEA 2010; Py et al. 2013; Zhang et al. 2013):

- Collectors/Mirrors: They reflect solar radiation and concentrate it into a small surface to reach high temperatures.
- Receiver: It receives the solar energy from the previous element so that it is used to heat up a HTF.
- Heat transfer fluid (HTF): It is responsible for transferring the solar energy to produce steam.

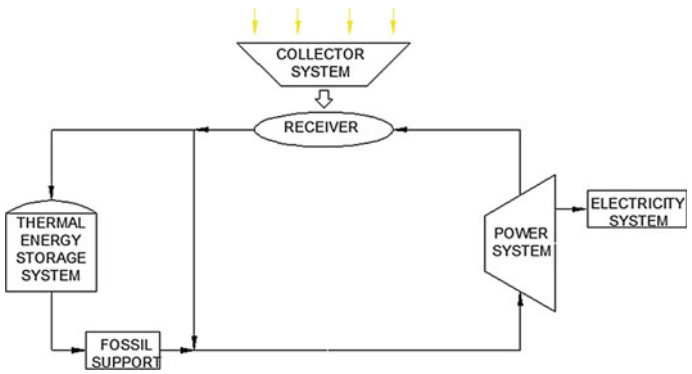


Fig. 9 Scheme of a CSP system

There are four different designs, namely (a) parabolic troughs, (b) solar towers, (c) parabolic dishes, and (d) linear Fresnel, whose differences are in the type of collector and the receiver. Figure 10 (IEA 2010) shows the alternative technologies.

(A) Parabolic Troughs Collectors

It is the more mature of the CSP technologies and it is used in most of the commercial size plants in operation. It provides from 30 to 320 MW per solar field. (López Cózar 2006). It consists of two elements (IEA 2010; Zhang et al. 2013).

- *Collector system (Mirrors)*: The solar field consists of curved parabolic mirrors that concentrate solar energy on the receiver allocated in the focal of the parabola.
- *Receiver*: It is a double pipe device. The HTF flows through the inner pipe. It is made of steel and it is covered by a film to allow high energy absorption and low emissivity. The outer pipe is made of glass to reduce thermal losses. Figure 11 shows a scheme of the receiver.

The collector–receiver system is designed in a way that it follows the sun from East to West. These systems concentrate solar energy up to 30 or even 80 times so that the temperature of the HTF reaches 673 K (Ruiz Hernández 2011). The most typical fluids are:

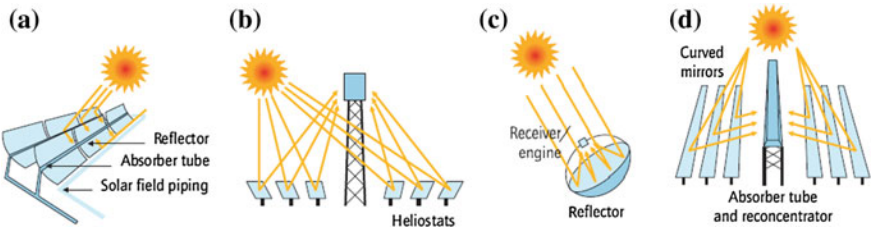


Fig. 10 CSP gathering technologies (IEA 2010). © OECD/IEA 2010 technology roadmap: concentrating solar power, IEA publishing. Licence: www.iea.org/t&c/termsandconditions

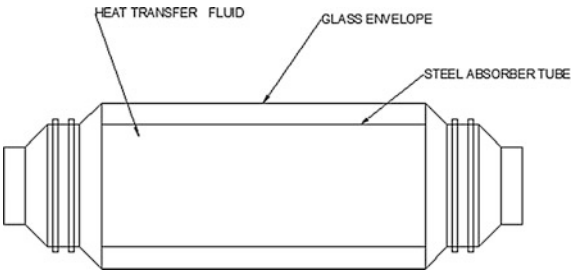


Fig. 11 Receiver for parabolic collectors

- **Synthetic oil**. It is the most used, being capable of reaching temperatures of 666 K. However, it cannot be used as energy storage fluid. On the other hand, it is compatible with molten salts-based systems for storage purposes for up to 7.5 h. As an example, La Florida CSP plant operates in this way (NREL 2011a).
- **Water**. It can operate at 673 K and 10 MPa but energy storage is still a challenge. Stillwater Geosolar Hybrid Plant operates using this system (NREL 2011b).
- **Molten salts**: It is the HTF that allows higher operating temperatures, 823 K, and they can be used as energy storage fluid as well for up to 8 h. Archimedes plant uses this concept (NREL 2011c).

(B) Solar power towers

This technology, also known as central tower, has recently been implemented at industrial scale for solar fields of 10 to 200 MW (IEA 2010; Stine and Geyer 2014; Tian and Zhao 2013; Zhang et al. 2013). The system consists of three elements:

- *Collector system*: Solar energy is collected by the heliostat field made of a number of mirrors. Each heliostat is typically of 40–150 m^2 with a reflecting surface of glass-metal. They have a two-axis tracking system to secure collecting solar radiation over a day period and redirecting it toward the upper part of a tower. There are two common configurations, north or circular as a function of the ground geography.
- *Tower*: It is the support for the receiver as well as it reduces shadows. It can be made of concrete or steel. Some facilities use a number of towers to provide the energy for the power block.
- *Receiver*: It is a fixed device that transforms concentrated solar energy into thermal energy. Two designs can be found, cylindrical, or cavity receivers. Figure 12 presents a scheme of both.

This technology is capable of concentrating solar energy from 500 to 1000 times so that the HTF reaches 838 K (Ruiz Hernández 2011). The most common HTF are (NREL 2011d–g):

- **Water**: It allows direct production of steam below 573 K. It is possible to store energy in the form of compressed water for about 1 h. Planta solar 10 uses this configuration
- **Molted salts (40 % KNO_3–60 % $NaNO_3$)**: They are heated up to 838 K and allow thermal energy storage for 15 h. An example of plant that uses this technology is Gemasolar Thermosolar plant.
- **Air**: It is possible to reach up to 953 K with energy storage times of 1.5 h such as the Jülich Solar Tower.
- **Liquid sodium**: We can reach temperatures up to 833 K and store solar energy for about 3 h. A HTF used by the Jemalong Solar Thermal Station.

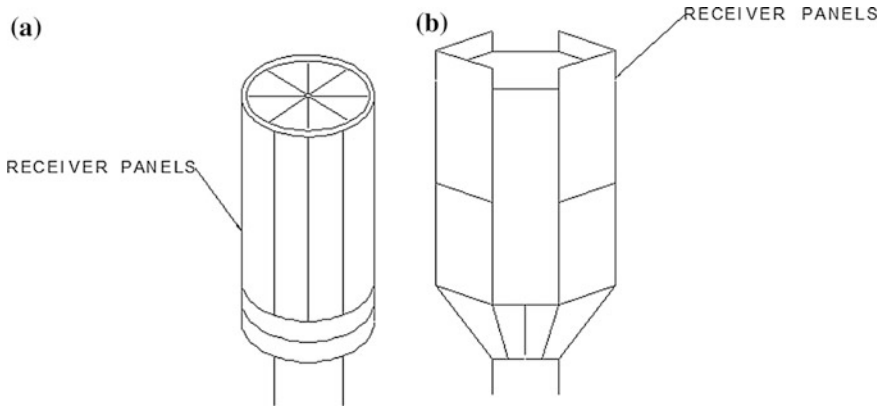


Fig. 12 Scheme of **a** cylindrical and **b** cavity receiver

(a) Parabolic dishes

Parabolic dish technology with Stirling engine is the less mature technology. It can provide 5–25 kW per dish. The device consists of three elements that track solar movement as a whole (IEA 2010; Tian and Zhao 2013; Zhang et al. 2013):

- *Collector*: They are individual glass mirrors of parabolic cross section with open diameters from 7 to 17 m. They focus sunlight on the focal point in the center of the dish. They are provided with a two-axis solar tracking system and polar tracking as well.
- *Receiver*: It is allocated in the focal point of the parabolic collector. It is responsible for transferring the energy to the HTF, typically high-pressure helium or hydrogen.
- *Independent engine*: It consists of a Stirling thermodynamic cycle and the power generator itself.

As a result of the geometry of the collector and the low focal distance to diameter ratio, it is possible to reach solar concentration ratios above 2000 which allow very high temperatures for the HFT, from 923 to 1023 K (Ruiz Hernandez 2011).

(b) Linear Fresnel Reflectors

This technology is at the same development status as that of the parabolic dish presented above. It can provide 1.5–5 MW per field (IEA 2010; Zhang et al. 2013): The system consists of two elements:

- *Collector*: It comprises parallel lines of a number of plane or slightly curved mirrors installed on the ground. The mirrors rotate to send the sunrays to a receiver.
- *Receiver*: It is a fixed tubular device allocated on an elevated place.

The system allows concentrated ratios of 25–40 times so that the HTF reaches 543 K (Ruiz Hernandez 2011). Typically, the facilities using this technology have water as HTF so that there is direct steam generation to feed the power block such as the plant Puerto Errado 1 (NREL 2011h).

Based on the previous discussion we can summarize the technologies in Table 2.

2.3.2 Thermal Energy Storage Energy (TES)

Solar energy is plentiful but its availability varies during a day and across seasons. Therefore, in order to buffer moments of lower or no solar radiation, thermal storage systems are being developed (Zhang et al. 2013). There are two major types, direct or indirect storage:

- Direct storage: The HTF is the same as that used for energy storage, i.e., molten salts.
- Indirect storage. In this case, the storage and the heat transfer fluids are different. The HTF transfers the energy to the storage fluid using a heat exchanger, i.e., synthetic oil—molten salts system.

As a result there are different configurations for the storage system (EASAC 2011):

- Indirect thermal storage in two tanks. A tank stores the cold fluid and a different one the hot fluid.
- Direct thermal storage in two tanks. One tank storages the HTF cold and in the other the hot HTF.

Table 2 Comparison of CSP capture technologies

	Parabolic troughs	Solar tower	Parabolic dishes	Linear fresnel reflectors
Size (MW)	30–320	10–200	5–25 KW	1.5–5 MW
Operating temperature (K)	666	838	1.023	543
Range of operating temperature (K)	<673	>673	>673	<673
Concentration ratio	30–80	500–1000	>2000	25–40
Peak efficiency (%)	20	23	29.4	14
Peak efficiency (conversion solar energy to power)	23–27	20–27	20–30	18–22
Annual efficiency (conversion Solar energy to power)	15–16	15–17	20–25	8–10
Thermodynamic efficiency	Low	High	High	Low
Tracking	One axis	Two axis	Two axis	One axis
Maturity of technology	Mature	Recent	Mature	Recent
Commercial status	Available	Available	Pilot plant	Pilot plant
Technological risk	Low	Medium	High	Low
Storage possibility	Yes	Yes	Batteries	Yes
Receptor	Mobile	Fixed	Mobile	Fixed

- One tank storage system: Thermocline or stratification is used so that both, the hot and the cold fluid, share the tank. The hot fluid is in the top part of the storage tanks while the cold one is stored in the lower part.
- One tank storage system with recycle: The cold fluid is recycled and the HTF is heated up by means of a gas furnace in absence of solar energy. Steam generation takes place in the same storage tank.

The actual storage capacity varies as a function of the load. We describe four operating conditions for the same solar field size and power production capacity but they differ in the time of the day the energy is produced (IEA 2010).

1. **Small size storage (Intermediate load)**. Power is produced during daytime when solar energy is available. See Fig. 13.
2. **Medium-size storage (Delayed intermediate load)**. It stores solar energy during daytime but it only produces, from noon on and after sunset. See Fig. 13.
3. **Large-scale storage**. It is used for the following load configurations:
   - Base load: To produce power 24 h a day.
   - Peak load. It only produces energy at peak hours.

Figure 14 shows the two operation modes.

The storage tanks must be properly isolated to avoid energy loses. We can classify the storage system into three types (Py et al. 2013):

- *Sensible heat-based storage system*. It is the technology with a higher degree of development. It uses species that store energy by heating up the storage fluid.

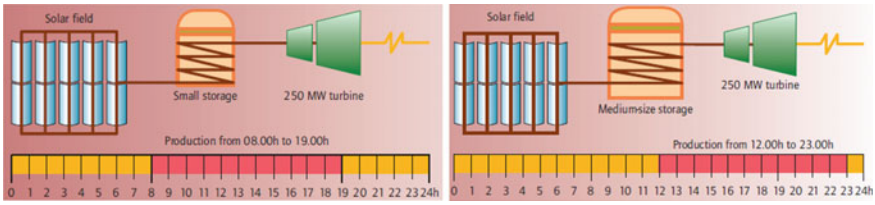


Fig. 13 Small and medium size storage system operation © OECD/IEA 2010 technology road map: concentrating solar power, IEA publishing. Licence: www.iea.org/t&c/termsandconditions

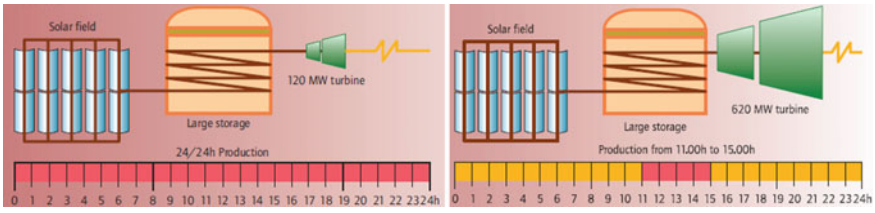


Fig. 14 Small and medium size storage system operation © OECD/IEA 2010 technology roadmap: concentrating solar power, IEA publishing. Licence: www.iea.org/t&c/termsandconditions

This fluid must meet certain characteristics such as thermal stability in the range of 473–1473 K, high heat capacity and thermal conductivity and low reactivity. They can be solids such as sand, concrete, refractory bricks, or liquids like synthetic oil or inorganic molten salts.

- *Latent heat-based storage systems*. They are based on phase changes of the chemical species which must meet the following properties such as high latent heat and heat conductivity and chemical stability, i.e., NaOH.
- *Chemical heat storage materials*. This technology is under development. It consists of using thermal energy to break and/or form bonds in reversible reactions that are endothermic/exothermic so that the material can absorb and generate energy. Typically, these species must follow simple reversible reactions with high enthalpy change. For instance,

$$CaCO_3 \leftrightarrow Ca + CO_2$$

Which is capable of operating from 1073 to 1173 K (Tian and Zhao 2013).

2.3.3 Power Production

Power production via CSP technology uses the solar collector system, a power block, and a cooling system as well as the possible storage system depending on the expected operation. Both, the solar gathering and the storage systems have been presented above. In this section we focus on the typical thermodynamic cycles and we leave to a particular example in the following section of the chapter to describe the operation of dry and wet cooling systems. Among the common thermodynamic cycles we distinguish (Kuravi et al. 2013):

- *Rankine cycle with steam turbine*. They are the most used cycles when the receptor temperature ranges from 573 to 823 K. In some cases a regenerative ranking cycle is implemented, steam extracted from the turbine at medium pressure is reheated up using solar energy to improve the efficiency.
- *Brayton cycle with gas turbine*. It becomes interesting for operating temperatures from 873 to 1473 K.
- *Stirling Engine*: Using helium or hydrogen as HTF, it is only used in the parabolic dish technology.
- *Combined cycle (Rankine + Brayton cycles)*: It is under development so that the wasted heat from the gas turbine (at 773 K) is used to produce steam and feed the steam turbine

Table 3 shows a brief comparison for the technologies and operation of typical CSP facilities.

Table 3 Sumarry of thermodynamic cycles for CSP technologies

	Parabolic trough	Solar tower	Linear fresnel reflectors	Dish parabolic
Cycle	Superheated steam Rankine	Superheated steam Rankine	Saturated steam Rankine	Stirling
Steam conditions (K/bar)	653–813/100	813/100–160	533/50	NA
Water requirement (m^3/MWh)	3 (wet cooling) 0.3 (dry cooling)	2–3 (wet cooling) 0.25 (dry cooling)	3 (wet cooling) 0.2 (dry cooling)	0.05–0.1 (mirror washing)

3 Chemicals and Power Production Process Evaluation

Solar energy cannot easily be used but it can be transformed into power, either using PV solar or concentrated solar power, or into chemicals. The production of power from solar energy is gaining attention but has several drawbacks related to the availability of solar incidence along the year as we have seen along the chapter. In this section we first discuss the production of power using CSP and later the use of solar power to produce chemicals in a way to store solar energy in a handy form.

3.1 CSP Power

A CSP facility consists of three sections, solar field, thermal cycle, and cooling. See Fig. 15.

In the previous section we have seen the options to model and evaluate the performance of the solar field as function of the solar availability and the actual allocation of the facility. Modeling studies on the field to establish the allocation of the mirrors and including the effect of the solar incidence on the efficiency of the solar field can be found in the works of Montes et al. (2009a, b), Ghobeity et al. (2011), Desai et al. (2014). The solar energy collected is used to heat up a heat transfer fluid, typically molten salts. These salts are used to store the energy so that the plant operates continuously overnight. Furthermore, the salts are heated up above 700 K to produce the steam that will be used in a Rankine cycle to produce electricity. A number of simulation and not so many optimization studies are available to evaluate the performance of the thermal cycle. Most of them use a Ranking cycle with regeneration. The extraction from the high pressure turbine is reheated up using the molten salts again to superheat the steam and increase the power generated in further expansions in the turbine. The simulation-based studies by Xu et al. (2011), Nezammahalleh et al. (2010), Morin et al. (2010), Halb et al. (2012) present different operating conditions for the operating pressures of the steam across the turbine. Palenzuela et al. (2011, 2013) performed a sensitivity analysis to determine the optimal exhaust temperature from the low pressure steam.

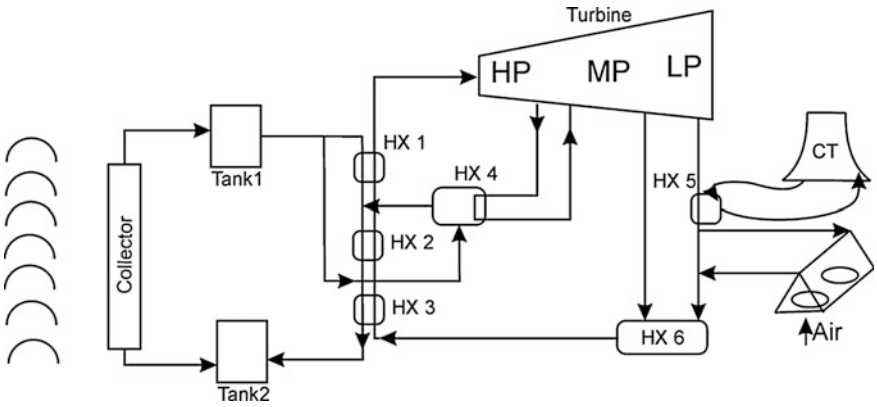


Fig. 15 Scheme of a CSP facility

Martín and Martín (2013) optimized the performance of such thermal cycle using an equation-based approach determining the optimal operating conditions of the cycle using a wet cooling system. The operation of the plant was subjected to the solar availability and for an average production of 18 MW, the actual production ranges from 9.5 MW during winter to 25 MW during summer. The average consumption of water is 2.1 L/kWh, due to evaporation losses in the cooling tower, which is competitive with thermoelectric plants. However, the performance of the cooling tower varies with the weather conditions. The investment of the plant is 260 M€ and the production cost 0.15 €/kWh, a little high compared to fossil fuel-based electricity. However, economies of scale are expected to reduce the production cost and the investment per kW generated by half when the production capacities reach those of current thermal power facilities.

Water consumption is an important issue in power generation. When it comes to solar facilities it becomes a real concern. Regions with high solar availability correspond with those with low freshwater access. Therefore, alternative cooling technologies must be evaluated. Over the last years there are a number of studies that compare dry and wet cooling systems for power plants, not only for CSP facilities. There is a trade-off: the savings in water consumption versus the energy consumption by the dry cooling, air cooling system. The first studies deal with fossil-based power plants. Turchi et al. (2010) showed that the use of dry cooling represents an increase in the production cost of electricity of 3–8 % due to the energy consumed by the fans of the air cooler. The investment of the dry cooled facility was 2 % higher compared the performance of wet induced draft cooling tower and dry A-frame cooling technologies for coal power plants. Their results showed that the dry cooling system was more than twice as expensive per kW of cooling, resulting in an investment cost 7.5 % higher when using the dry cooling system and the efficiency of the plant got reduced 2 % (Zhai and Rubin 2010). The water saved by dry cooling was that required by the wet cooling system, 2.46 t of water per MWh. Blanco-Marigorta et al. (2011) and Habl et al. (2012) compared a

forced draft wet cooling tower with an A-frame air cooler for the operation of a CSP facility, the Andasol 1 power plant. In spite of the fact that no water consumption values were reported, the electricity production cost were 15.27 ct/kWh when using wet cooling and 16.08 ct/kWh in case of dry cooling. Liqreina presented the operation of dry cooled-based plants, in particular Andasol in Spain and Ma'an in Jordan, and compared the use of dry cooling with wet cooling over a year long using the Greenius software as platform. The dry cooled facility showed 4–8 % lower energy production. The LCOE tuned out to be 0.1284 €/kWhe with an investment of 248 M€ while the dry cooled-based plant produced energy with a LCOE of 0.1491€/kWh and an investment of 289 M€ for a production facility of 50 MW. The water consumption of the wet cooled facility was 1.6 L/kWh, while the dry cooled was 0.095 L/kWh, for mirror cleaning. Palenzuela et al. (2013) used Engineering Equation Solver (EES) software to evaluate both cooling types of technologies. The efficiency of the dry cooling-based plant was 2 % lower than the evaporative water cooling system while the levelized energy cost of the dry cooling was 0.249 €/kWh versus the 0.241 €/kWh of the wet system. In parallel to simulated-based approaches, Martín (2015) optimized the operation of an A-frame within a CSP facility to evaluate the energy consumption due to the operation of the fan. The average production of energy is 17.2 MW but it ranges from 9.5 MW during winter to 25 MW during summer. 5 % of the energy produced is consumed by the fans of the A-frame cooling system. The operation of the A-frame varies with the air temperature and the number of fans in operation can vary along the day but mostly with the power generated by the solar field. The advantage is the reduced water consumption. The investment of the plant is 265 M€ and the production cost 0.16 €/kWh, slightly higher compared to the wet cooled facility (Martín and Martín 2013).

In Chap. 8 hybridation between biomass and CSP is presented for constant power production over a year.

3.2 Production of Methane, Methanol

Power storage is a challenging problem. Batteries, so far, do not have the capacity to deal with the energy generated through solar or wind energy. However, we can store them as chemical energy, in the form of H_2 or ready-to-use fuels such as methane or methanol. Typically wind and solar PV energy have been used to produce hydrogen via water electrolysis. (Saur and Ramsden 2011; Ursua et al. 2012) discussed the production of hydrogen from wind and PV solar for several configurations due to the challenges related to the availability of both. The use of hydrogen as fuel depends on the development of fuel cells and the evolution of its storage (Smolinka 2014). Hydrogen is a small molecule that can easily escape metals. Furthermore, lately the level of CO_2 in the atmosphere has surpassed

400 ppm. Therefore, it is possible to use electrolytic H_2 and CO_2 to produce methane and methanol among others, see Fig. 16.

The process for the production of methane was presented by Davis and Martín (2014a, b). It consists of the use of renewable energy to produce power that is used for water electrolysis. The oxygen is dried and compressed to be sold. On the other hand, the hydrogen follows a similar path, dehydration and compression, but in this case the traces of O_2 remaining in the gas stream must be removed using deoxygenation. At this point, it is mixed with CO_2.

Methane: The production of methane takes place in the range of 140–350 °C and from 100 to 3000 kPa (Gassner and Marechal 2009).

$$CO + 3H_2 \leftrightarrow CH_4 + H_2O$$
$$CO_{2(g)} + H_{2(g)} \leftrightarrow CO_{(g)} + H_2O_{(g)}$$

Next, the gas produced is purified by condensing the water and removing the CO, CO_2 to obtain a composition similar to natural gas.

Methanol: The reactions to produce methanol are similar (Chinchen et al. 1988). The methanol synthesis reactor operates at 200 °C and 50 bar using $Cu/ZnO/Al_2O_3$ as catalyst. Next, unreacted gases are recycled and the methanol is purified by separating it from the water produced (Martín and Grossmann 2015).

$$CO_2 + 3H_2 \leftrightarrow CH_3OH + H_2O$$
$$CO_2 + H_2 \leftrightarrow CO + H_2O$$

In both cases it is important for the operation of the facility to recycle the water generated in the reaction. That will reduce the water consumption of the facility since it is a raw material. Comparing with the products, see Table 4, we see that the

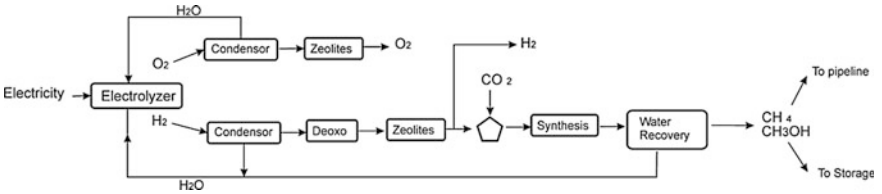


Fig. 16 Block diagram for electrolysis-based products

Table 4 Main raw materials and products in the production of chemicals from CO_2 and renewable H_2

Main product/other products and raw materials	CH_4	CH_3OH
CO_2 captured (kg/kg)	3.0	1.4
H_2O consumption (kg/kg)	2.8	1.2
O_2 production (kg/kg)	3.3	1.5

production of methane requires larger amounts of raw materials, but it also has a larger capacity to capture CO_2 compared to methanol. The larger the CO_2 captured the larger the consumption of water.

In a previous chapter, the use of wind to produce power has been analyzed. Therefore, hydrogen can also be produced using power generated from wind turbines or solar fields. The advantage of using solar field is currently its lower cost. The drawback is the large need for devoted area for the panels or the mirrors.

3.3 Solar Augmented Processes

CO_2 reduction to Syngas: The photosynthesis is a natural process that converts CO_2 and water into chemicals, carbohydrates, using solar energy—The Sunshine to Petrol Project from the US government (Miller et al. 2012) had the ambition to emulate that process to produce liquid fuels using the same three raw materials following a chemical reaction like.

Solar Energy + xCO_2 + $(x + 1)\ H_2O \rightarrow C_xH_{2x+2}$ (liquid fuel) + $(1.5x + 0.5)\ O_2$

However, to do so the CO_2 must be reduced to CO via the reenergizing reaction

$$2CO_2 + 4H_2O \rightarrow 2CO + 4H_2 + 3O_2$$

Once we obtain syngas, endless possibilities arise based on Fischer–Tropsch technologies. We have seen above the possibility of using electrolytic hydrogen to react with the CO_2. The current conversion from solar energy into electricity has a low efficiency, and therefore alternative technologies and methods are proposed. The thermolysis of water occurs spontaneously at 3000 °C, which is impractical from the technical point of view. There is, however, a way around. Thermal cycles can be used to obtain the same products by combining a series of chemical reactions. CSP can provide sufficient high temperatures for these reactions to happen. A typical cycle is that based on metal oxides

$$1/\delta\,MOx \rightarrow 1/\delta\,MO(x_\delta) + {}^1\!/_2O_2$$
$$1/\delta\,MO(x_\delta) + H_2O \rightarrow 1/\delta MOx + H_2$$
$$H_2O \rightarrow H_2 + {}^1\!/_2O_2$$

Among the metals that can be used in the system Zn/ZnO (Galvez et al. 2008), Ferrites ($MxFe_{3\_x}O_4$) (Kodama et al. 2005) or the recently developed, "hercynite cycle" in which a cobalt ferrite spinel ($CoFe_2O_4$) decomposes simultaneously with alumina (Al_2O_3) at 1460 °C to form aluminates (Arifin et al. 2012) are feasible alternatives. The scheme of the process can be seen in the figure below, Fig. 17. The

idea was to develop a reactor to carry out this task. Chueh and Haile (2010) developed a prototype reactor using a CeO_2 based cycle, while Diver et al. (2008) at Sandía National Lab modified a previous design for the CR5 reactor that used a mixture of cobalt ferrite and zirconia which is rotated through a solar-irradiated high-temperature receiver where it is reduced at 1400 °C and then a low-temperature section where the second step is carried out to produce liquid fuels out of CO_2.

Kim et al. (2011) designed a process and performed a technoeconomic analysis based on the CR5 reactor for the production of methanol from CO_2 using this technology. The block diagram of the process can be seen in Fig. 18. They determined a break-even point of $1.22/kg for this technology to be competitive.

Solar-based steam reforming: Steam reforming of hydrocarbons is a widely extended technology for the production of syngas. Syngas, as presented above and as it will be further discussed in Chap. 8, is a valuable mixture and building block for liquid fuels. Methane steam reforming process is endothermic that operates at high temperature (750–950 °C) to increase the conversion, since it is an equilibrium. Traditionally a fraction of the raw material is spent in order to provide the energy to maintain the temperature in the furnace. One interesting combination of solar energy with the typical petrochemical process is to use solar to provide that energy since we can have those high temperatures available in solar furnaces (De Maria et al. 1986; Böhmer et al. 1991).

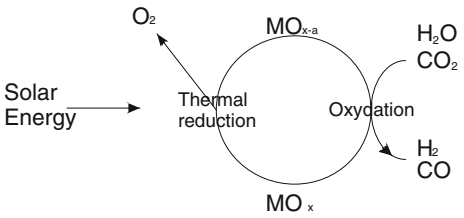


Fig. 17 CO_2 decomposition cycle

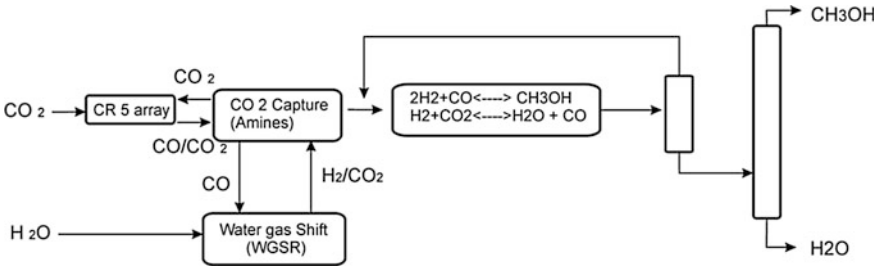


Fig. 18 Methanol production from water and CO_2

References

Anani, N., Shahid, O., Al-kharji, J.P., (2013) A CAD Package for Modeling and Simulation of PV Arrays under Partial Shading Conditions, Energy Procedia 42: 397–405.

Arifin, A., Aston, V.J., Liang, X.H., McDaniel, A. H., Weimer, A.W., (2012) CoFe2O4 on Porous Al2O3 Nanostructure for Solar Thermochemical CO2 Splitting, Energy Environ. Sci. 5: 9438–9443.

Blanco-Marigorta, AM, Sanchez-Henríquez, MV, Peña-Quintana, JA. (2011) Exergetic comparison of two different cooling technologies for the power cycle of a thermal power plant. Energy, 36: 1966–1972

Böhmer, M., Langnickel, U., Sanchez, M. (1991) Solar steam reforming of methane. Solar. Energy Materials 24: 441–447.

Chinchen GC, Denny RJ, Jennings JR, Spencer MS, Waugh KC. (1988) Synthesis of Methanol: part 1. Catalysts aid Kinetics Applied. Catalysis 36: 1–65

Chueh, W.C.; Haile, S. (2010) A Thermochemical Study of Ceria: Exploiting an Old Material for New Modes of Energy Conversion and CO_2 Mitigation Philos. Trans. R. Soc., A, 368: 3269–3294

Coley, D. (2008) Energy and Climate Change: Creating a Sustainable Future. Ed. Wiley. ISBN: 978-0-470-85312-2.

Davis, W., Martín, M (2014a) Optimal year-round operation for methane production from CO_2 and Water using wind energy. Energy 69: 497–505

Davis, W., Martín, M (2014b) Optimal year-round operation for methane production from CO_2 and Water using wind and/or Solar energy. J. Cleaner Prod. 80: 252–261

De María, G., Tiberio, C.A., D'Alessio, L., Piccirilli, M., Goffari, E., Polucci, M. (1986) Thermochemical conversión of solar energy by steam reforming of methane. Energy, 11 (8): 805–810

Desai, N.B., Kedare, S.B., Bandyopadhyay, S., (2014). Optimization of design radiation for concentrating solar thermal power plants without storage. Sol. Energy 107: 98–112.

Diver, R. B.; Miller, J. E., Allendorf, M. D., Siegel, N. P., Hogan, R. E. (2008) Solar Thermochemical Water-Splitting Ferrite-Cycle Heat Engines J. Sol. Energ-T ASME, 2008, 130, 041001.

European Academies Science Advisory Council, EASAC (2011), Concentrating solar power: its potential contribution to a sustainable energy future

Galvez, M.E:, Loutzenhiser, P. G., Hischier, I., Steinfeld, A., (2008) CO_2 splitting via two-step solar thermochemical cycles with Zn/ZnO and FeO/Fe3O4 redox reactions: thermodynamic analysis. Energy Fuels, 22: 3544–3550.

Gassner, M., Marechal, F., (2009). Thermo-economic process model for thermochemical production of Synthetic Natural Gas (SNG) from lignocellulosic biomass. Biomass Bioenergy 33: 1587–1604

Ghobeity, C.J. Noone, C.N. Papanicolas, A. Mitsos, (2011) Optimal time-invariant operation of a power and water cogeneration solar-thermal plant. Solar Energy 85: 2295–2320

Gong, J., Darlin, S., You, F. (2015) Perovskite Photovoltaics: Life cycle assessment of energy and environmental impacts. Energy Environ. Sci, 8: 1953–1968

Habl, P, Blanco-Marigorta, AM, Erlach, B. Exergoeconomic comparison of wet and dry cooling technologies for the Rankine cycle of a solar thermal power plant. Proceedings of ecos 2012 - the 25th

Halb, P., Blanco-Marigorta, A.M., Erlach, B. (2012) Exergoeconomic comparison of wet and dry cooling technologies for the Rankine cycle of a solar thermal power plant. Proceedings of ecos 2012 - the 25th international conference on efficiency, cost, optimization, simulation and environmental impact of energy systems 300–1, 300–14

Hudon, K., Merrigan, T., Burch, J. and Maguire, J., (2012) Low-Cost Solar Water Heating Research and Development Roadmap, National Renewable Energy Laboratory. NREL/TP-5500-54793.

IDAE (2011) Pliego de Condiciones Técnicas de Instalaciones Conectadas a Red. Available on: http://www.idae.es/uploads/documentos/documentos_5654_FV_pliego_condiciones_tecnicas_instalaciones_conectadas_a_red_C20_Julio_2011_3498eaaf.pdf

International Energy Agency, IEA (2010) Technology Roadmap Concentrating Solar Power France. Paris

Karatepe, E., Boztepe, M., Çolak, M., (2007) Development of suitable model for characterizing photovoltaic arrays with shaded solar cells, Solar Energy, 81, 977–992.

Kim, J., Henao, C.A., Johnson, T.A., Dedrick, D.E., Miller, J.E., Stechel, E.B., Maravelias, C. T. (2011) Methanol production from CO_2 using solar-thermal energy: process development and techno-economic analysis. Energy Environ. Sci., 4, 3122–3132

Kodama, T., Kondoh, Y., Yamamoto, R., Andou, H. and Satou, N. (2005) Thermochemical hydrogen production by a redox system of ZrO2-supported Co(II)-ferrite Sol. Energy, 78: 623–631.

Kuravi, S., Trahan, J., Goswami, Y, Rahman, M., Stefanakos, E., (2013) Thermal energy storage technologies and systems for concentrating solar power Plants, Prog.Energy and Comb. Sci., 39: 285–319.

Lee H. (2011) Metal species surface treatment of thin film photovoltaic cell and manufacturing method. US patent: US8026122

Liu, BYH, Jordan RC. (1960) The interrelationship and characteristic distribution of direct, diffuse and total solar radiation. Solar Energy, 4: 1–19.

Liu, Z. Liu, Q. Huang, Y. Ma, Y. Yin, S. Zhang, X. Sun, W. Chen, Y. (2008) Organic Photovoltaic Devices Based on a Novel Acceptor Material: Graphene. Advanced Materials. 20: 3924–3930.

López Cózar, J.M. (2006) Energía Solar térmica, Instituto para la Diversificación y Ahorro de la Energía (IDAE): 31–65.

Luque, A., Hegedus, S. (2003) Handbook of Photovoltaic Science and Engineering John Wiley & Sons, NY.

Martín, L, Martín M. (2013) Optimal year-round operation of a Concentrated Solar Energy Plant in the South of Europe Applied Thermal Engineering. 59: 627–633.

Martín, M., Grossmann, I.E: (2015) Optimal integration of a self sustained algae based facility with solar and/or wind energy. Submitted. Comp. Chem. Eng.

Martinez-Moreno, F., Muñoz, J., Lorenzo, E. (2010) Experimental model to estimate shading losses on PV arrays, Solar Energy Materials & Solar Cells, 94: 2298–2303.

Masters, G., (2004), Renewable and Efficient Electric Power Systems. Ed. Wiley-Interscience, New Jersey.

Miller, J.E., Allendorf, M.D., Ambrosini, A., Chen, K.S., Coker, E. N., Dedrick, D.E., Diver, R.B., Hogan, R.E., Ermanoski, I., Johnson, T.A., Kellogg, G.L., McDaniel, A.H., Siegel, N.P., Staiger, C.L., Stechel, E. B. (2012) Final Report Reimagining Liquid Transportation Fuels: Sunshine to Petrol SAND2012–0307 Sandia National Laboratories Albuquerque, New Mexico

Montes, M.J., Abánades, a., Martínez-Val, J.M., (2009) Performance of a direct steam generation solar thermal power plant for electricity production as a function of the solar multiple. Sol. Energy 83: 679–689.

Montes, M.J., Abánades, A., Martínez-Val, J.M., Valdés, M., (2009) Solar multiple optimization for a solar-only thermal power plant, using oil as heat transfer fluid in the parabolic trough collectors. Sol. Energy 83: 2165–2176.

Morin, P. Richter, P. Nitz. (2010) New method and software for multivariable technoeconomic design optimization of CSF plants. www.mathcces.rwth-aachen.de/_media/5people/richter/pascalrichter-2010-solarpaces.pdf (last accessed December 2012

Nasa (2015a) http://sunearthday.gsfc.nasa.gov/2009/TTT/65_surfacetemp.php

Nasa (2015b) https://solarsystem.nasa.gov/planets/profile.cfm?Object=Sun

Nasa (2015c) http://www.nasa.gov/images/content/57911main_Earth_Energy_Budget.jpg

Nezammahalleh, H., Farhadi, F., Tanhaemami, M. (2010) Conceptual design and techno-economic assessment of integrated solar combined cycle system with DSG technology. Solar Energy 84: 1696–1705

NREL (2011a) Concentrated solar power projects. La florida, http://www.nrel.gov/csp/solarpaces/project_detail.cfm/projectID=27
NREL (2011b) Concentrated solar power projects. Stillwater GeoSolar Hybrid Plant, http://www.nrel.gov/csp/solarpaces/project_detail.cfm/projectID=4279
NREL (2011c) Concentrated solar power projects. Archimede: http://www.nrel.gov/csp/solarpaces/project_detail.cfm/projectID=19
NREL (2011d) Concentrated solar power projects. Planta Solar 10 http://www.nrel.gov/csp/solarpaces/project_detail.cfm/projectID=38
NREL (2011e) Concentrated solar power projects. Gemasolar Thermosolar Plant http://www.nrel.gov/csp/solarpaces/project_detail.cfm/projectID=40
NREL (2011f) Concentrated solar power projects. Jülich Solar Tower http://www.nrel.gov/csp/solarpaces/project_detail.cfm/projectID=246
NREL (2011g) Concentrated solar power projects. Jemalong Solar Thermal Station http://www.nrel.gov/csp/solarpaces/project_detail.cfm/projectID=4284
NREL (2011h) Concentrated solar power projects. Puerto Errado 1 http://www.nrel.gov/csp/solarpaces/project_detail.cfm/projectID=46
O'Reagan, B., Gratzel, M (1991) A low-cost, high-efficiency solar cell based on dye-sensitized colloidal TiO2 films Nature 353, 737–740
Palenzuela, P., Zaragoza, G., Alarcón-Padilla, D.C., Guillén, E:, Ibarra, M., Blanco, J. (2011) Assessment of different configurations for combined parabolic-trough (PT) solar power and desalination plants in arid regions, Energy 36: 4950–4958
Palenzuela, P., Zaragoza, G., Alarcón-Padilla, D.C., Blanco, J. (2013) Evaluation of cooling technologies of concentrated solar power plants and their combination with desalination in the mediterranean area Applied Thermal Engineering 50: 1514–1521
Parida, B., Iniyan, S., Goic, R., (2011) A Review of solar phonovoltaic technologies. Renew. Sust. Energ. Revs. 15 (3): 1625–1636.
Py, X., Azoumah, Y., Olives, R. (2013) Concentrated solar power: Current technologies, major innovative issues and applicability to West African countries, Renewable and Sustainable Energy Reviews, 18: 306–3015
Ruiz Hernandez, V. (2009) La electricidad solar térmica, tan lejos, tan cerca, Fundación Gas Natural: 33–91
Ruiz Hernandez, V (2011) Evaluación del potencial de energía solar termoeléctrica. Estudio Técnico PER 2011–2020: 7–50
Saur, G., Ramsden, T (2011) Wind Electrolysis: Hydrogen Cost Optimization Technical Report NREL/TP-5600-50408 May 2011
Singh, G.K. (2013) Solar power generation by PV (photovoltaic) technology: A review. Energy 53: 1–13.
Smolinka, T (2014) Water Electrolysis: Status and Potential for Development Fraunhofer-Institut für Solare Energiesysteme ISE Water Electrolysis Day Brussels (BE), April 03, 2014 http://www.fch-ju.eu/sites/default/files/2%20Water%20Electrolysis%20Status%20and%20Potential%20for%20Development.pdf
Stine, W., and Geyer, M., (2014) Power from the sun. http://www.powerfromthesun.net/Book/chapter11/chapter11.html
Tham Y, Muneer T, Davidson B. (2010) Estimation of hourly averaged solar irradiation: evaluation of models. Building Service Engineering Research and Technology 31(1): 9–25.
Tian, Y., Zhao, C.Y., (2013) A review of solar collectors and thermal energy storage in solar thermal applications, Applied Energy 104: 538–553.
Trieb, F., Hoyer-Klick, C., Schillings, C., (2008), Institut für Technische Thermodynamik. Available on: http://www.dlr.de/tt/desktopdefault.aspx/tabid-2885/4422_read-16596/
Turchi, CS, Wagner, MJ, Kutscher CF. (2010) Water Use in Parabolic Trough Power Plants: Summary Results from WorleyParsons' Analyses. NREL/TP-5500–49468. December 2010
Ursua, A., Gandía, L.M., Sanchis, P (2012) Hydrogen Production From Water Electrolysis: Current Status and Future Trends Proceedings of the IEEE | Vol. 100, No. 2, 4410–426

Xu, C., Wang, Z., Li, X., Sun, F. (2011) Energy and exergy analysis of solar power tower plants Applied Thermal Engineering 31: 3904–3913

Ye, X. Shao, Z. Hongbo, Z. Yang, L. Wang, C. (2014) Electronic and optical properties of silicone nanomeshes. RSC Advances. 4: 37998–38003

Zhai, H, Rubin ES. (2010) Performance and cost of wet and dry cooling systems for pulverized coal power plants with and without carbon capture and storage Energy Policy. 38: 5653–5660.

Zhang, H.L., Baeyens, J., Degreve, J., Caceres, G., (2013) Concentrated solar power plants: Review and design methodology, Renewable and Sustainable Energy Reviews 22: 466–481

Biomass as Source for Chemicals, Power, and Fuels

Mariano Martín and Ignacio E. Grossmann

Abstract In this chapter we discuss on the possible routes to process biomass for the production of chemicals, fuels, and power. Once the individual processes are described, we present several alternatives for the use biomass to produce a number of products simultaneously taking advantage of synergies between processes, the possibility of producing intermediates out of the same raw materials, and process integration opportunities among processes and energy sources. The methodology used is based on mathematical optimization techniques to allow for solving tradeoffs and identifying the best integrated operation of multiproduct plants.

1 Introduction and Method

Biomass has been used as a source of energy and construction materials for centuries. However, the easy access to crude oil in the late 1970s and its cheap price reduced the use of biomass and limited its applications (BP 2012). The unstable situation in many producing countries and the expected depletion of the resources has encouraged the search for alternative fuels. Actually, it has been more a "back to the past" trend rather than a step forward. For instance, the Ford T, the first car produced in an assembly line, was meant to run on ethanol, while vegetable oil fed diesel engines in the beginning. Over the last decades, research has focused on the use of biomass to substitute fuels by sustainable ones and lately other chemicals have been included in the portfolio. In this chapter, we start with individual processes that transform various types of biomass into products. Next, we comment on

M. Martín (✉)
Department of Chemical Engineering, University of Salamanca,
37008 Salamanca, Spain
e-mail: mariano.m3@usal.es

I.E. Grossmann
Department of Chemical Engineering, Carneguie Mellon University,
Pittsburgh, PA 15213, USA
e-mail: grossmann@cmu.edu

M. Martín (ed.), *Alternative Energy Sources and Technologies*,
DOI 10.1007/978-3-319-28752-2_8

the properties and composition of the biomass that will allow evaluating the possibilities for integrated production processes. Finally, biomass, unlike solar or wind energy, can be stored for a certain period. This is an important property to have since it allows using biomass as back up renewable energy source in the operation of solar and/or wind-based facilities.

The design of individual or integrated processes for the transformation of biomass into different fuels can benefit from process system engineering approaches. Conceptual design based on superstructure optimization (Martín and Grossmann 2013a–d) is a powerful tool to select among technologies, i.e., biodiesel catalysts, gasification alternatives, and determine the optimal operating conditions. Basically, the approach consists of modeling the different options using simple but reliable models, formulating the superstructure of alternatives in the form of mixed integer nonlinear programming problems and solving it. Simultaneous optimization and heat integration can be included in the formulation (Duran and Grossmann 1986). Next, for a given process topology, we can design the heat exchanger network (HEN) to reduce the utilities consumption based on the model presented by Yee and Grossmann (1990). Water consumption issues can be addressed either at the design stage of the flowsheet and/or by designing the optimal water treatment and reuse network (Ahmetovic and Grossmann 2011).

Typically, individual processes do not operate independently but as a part of chemical complexes. In order to identify integration opportunities and evaluate non obvious tradeoffs, systematic analysis, and process design using mathematical programming techniques is again a powerful ally. Once the process or the integrated complex is established, rigorous simulation using commercial software such as ASPEN plus or CHEMCAD, can be used to simulate the performance of the process (i.e., Zhang et al. 2003 for biodiesel production, de la Cruz et al. 2014 for integrated production of diesel substitutes).

2 Individual Processes

2.1 Grain Based

Grain, corn, or wheat, consist mainly of starch which is a polymer of glucose units. Therefore, they both are used within the food chain. However, by proper breakage of the polymer, grain can be a source of glucose. This is the basis for first generation bioethanol.

In order to extract the starch out of the grain, it is processed through grinding and cooking with steam. Next, the biomass is subjected to liquefaction and saccharification at 90 and 65 °C, respectively, for 30 min each, so that the starch breaks into maltose and then into glucose, as presented in the following reactions. Apart from the proper temperature, enzymes, amylases, and glucoamylases, catalyze the process (Jacques et al. 1999).

$$2(C_6H_{10}O_5)_n + nH_2O \longrightarrow \alpha\text{-amylase} \longrightarrow nC_{12}H_{22}O_{11}$$
$$C_{12}H_{22}O_{11} + H_2O \longrightarrow \text{glucoamylase} \longrightarrow 2C_6H_{12}O_6$$

At this point, glucose is available as a raw material in a dilute stream with water. There are a number of chemicals that can be produced out of glucose fermentation such as acetone or butanol, but ethanol is the easiest one. The fermentation of glucose to ethanol takes place isothermally at 32–38 °C under anaerobic conditions. We need to ensure that the ethanol concentration in the mix is always below 15 %, so that the fermentation is not inhibited. Ethanol is toxic for the *saccharomyces cerevisiae*, the bacterium catalyst.

$$C_6H_{12}O_6 \longrightarrow \text{yeast} \longrightarrow 2C_2H_6O + 2\ CO_2$$

While ethanol is the main product, there are a number of secondary reactions whose conversions and stoichiometry are presented in Table 1.

The fermentation takes place for 24–72 h for the ethanol to be produced. After that, and before further treatment, Karuppiah et al. (2008) showed that solids separation is the best option. A flotation unit is used to recover the protein that will be dried to produce the so-called dried distiller grains and solubles that can be used as cattle feed. The liquid is now sent to a distillation column to concentrate the ethanol. One of the main disadvantages of the process is the energy intensive distillation in the beer column. In order to reduce the energy and cooling needs at that column, a multieffect distillation system can be used. It consists of splitting the feed into a number of columns so that the reboiler of the lower pressure column acts as condenser of the higher pressure column. A three effect system is used (Karrupiah et al. 2008) to reduce the energy needs with a limited increment in the investment cost. The bottoms of the column contain mainly water, but a number of organic species follow such as lactic acid, acetic acid, glycerol, etc. This stream is actually a good source of water to reduce the overall water consumption, (Ahmetovic et al. 2010) but it has to be properly treated to remove the organics. Ethanol fuel cannot contain more than 0.03 % water. Therefore, a dehydration step

Table 1 Main reactions in the production of corn-based ethanol

Reaction	Conversion
$\text{glucose} \longrightarrow \text{yeast} \longrightarrow 2\ \text{ethanol} + 2\ \text{carbon dioxide}$	0.92
$C_6H_{12}O_6 + 2\ H_2O \longrightarrow \text{yeast} \longrightarrow 2C_3H_8O_3 + O_2$	0.034
$C_6H_{12}O_6 + 2\ CO_2 \longrightarrow \text{yeast} \longrightarrow 2C_4H_6O_4 + O_2$	0.01
$\text{glucose} \longrightarrow \text{yeast} \longrightarrow 2\ \text{lactic acid}$	0.002
$\text{glucose} \longrightarrow \text{yeast} \longrightarrow 3\ \text{acetic acid}$	0.0024
$\text{glucose} + 1.2\ \text{ammonia} \longrightarrow \text{yeast} \longrightarrow 6\ \text{cell mass} + 2.4\ \text{water} + 0.3\ \text{oxygen}$	0.0316

Fig. 1 Block diagram for corn-based ethanol

follows the beer column. Rectification is the typical option (Jacques et al. 1999) but it is more energy intensive. Therefore, a combination between adsorption in corn grits and molecular sieves allows reducing the energy consumption while securing fuel quality. Energy integration in this process is crucial for a positive net energy balance and a competitive price of 1.24 $/gal (Karuppiah et al. 2008). The advantage of energy integration is twofold. First, it reduces the energy consumption in a challenging process since the operating temperatures are low, and reusing the energy is a difficult task. Furthermore, the fermentation is exothermic, but operates at low temperatures increasing the cooling needs. As a result, energy integration also reduces the cooling needs for a minimum of 1.5 gal of water per gal of ethanol produced (Ahmetovic et al. 2010). Figure 1 shows a scheme of the process.

2.2 Oil Based

The second main biofuel is biodiesel. First generation biodiesel has traditionally been produced from oil seeds. The oil is extracted from the seed using mechanical operations, providing heat, and typically using a solvent. There are a number of procedures, but most of them make use of any of these three options. Therefore, once the oil is extracted, it has to be separated from the solvent using flash distillation. The high difference in the boiling point between the oil and the solvent simplifies the process. As in the case of the bioethanol, seeds and grain are both food. The ethics behind their use as a source of fuels has pushed the industry toward substituting the source of oil to non edible seeds such as jatropha or cooking oil. In this last case, the impurities must be taken into account before processing the oil.

The production of biodiesel from oil is based on the transesterification reaction of the oil with alcohols. Basically the oil consists of three chains of hydrocarbons whose viscosity is difficult to process. Therefore, the idea is to break it down into three chains. By doing this, we do not only reduce the viscosity of the mixture, but also in this way the properties of the product match those of the crude-based diesel.

The transesterification is an equilibrium reaction between the oil and alcohols. For economic reasons, methanol has been used for a long time. From the technical point of view, it provides high yield to biodiesel, fatty acid methyl ester (FAME), and quick reaction times. Martín and Grossmann (2012) evaluated a number of catalysts, homogeneous acid and basic, heterogeneous, enzymatic, and non-catalyzed under supercritical conditions. In order to drive the equilibrium to

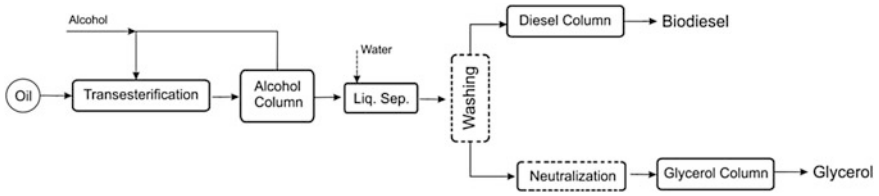


Fig. 2 Oil-based biodiesel

products, an excess of alcohols is typically used as well as the proper amount of catalysts and the operating pressure and temperature (Meher et al. 2006). The impurities in the oil, water, and free fatty acids, pose on the use of alkali catalyst an extra challenge since pretreatment must be used to esterify them using an acid catalyst. When comparing the different catalysts, the actual process flowsheet structure changes from one to another. After the transesterification, a distillation column is typically used to recover the excess of methanol. If homogeneous catalysts are used, a washing step is needed to separate the catalyst with the polar phase from the biodiesel. Next, the polar phase is neutralized before the glycerol is purified. The organic phase containing the biodiesel is distilled to eliminate the unconverted oil. However, the use of heterogeneous catalysis or supercritical conditions simplifies the purification steps since a solid-liquid separation removes the catalyst after the reaction, and after the polar—organic phase separation, the products have higher purity. Figure 2 shows the block diagram from purified oil. Units and streams in discontinuous boxes and lines are optional depending on the catalyst used.

Martín and Grossmann (2012) optimized the production of biodiesel using methanol providing the optimal operating conditions at the reactor and suggesting the use of alkali catalyst for oil with a limited amount of impurities and heterogeneous catalyst for cooking oil. The formulation they used in the optimization, Duran and Grossmann's Model (1986) for simultaneous optimization and heat integration, allowed determining tradeoffs related to the excess of methanol used in the transesterification versus the energy that is to be spent in the recovery stage. In this way, the optimal operating conditions at the reactor are computed. Note, that the results are different to the ones reported in the literature (Zhang et al. 2003). The values used by Zhang et al. (2003) are based on experimental evaluation of the reactor aiming at the highest conversion, but they do not account for the trade-off between an increase in the conversion and the cost of energy required to recover the excess of alcohol. The optimization of the energy consumption also reduces water consumption. The work shows promising production costs of $0.42/gal, and reduced energy and water consumption values of 1.94 MJ/gal and 0.6 gal/gal using the alkali catalyst when the oil is clean, and values of $0.66/gal, 1.94 MJ/gal and 0.33 gal/gal, respectively, using the heterogeneous catalyst, when we use cooking oil. Note that the use of heterogeneous catalyst reduces water consumption since there is no washing step. The higher cost is due to the cost of cooking oil according to literature values.

Not only is methanol competitive as alcohol, but also ethanol. Severson et al. (2013) evaluated the use of ethanol as transesterifying agent. The reason is the

availability of ethanol within biorefineries, and therefore the reduced dependency on fossil-based raw materials. The process flowsheet is similar to the one in Fig. 2 for oil sources with limited amount of impurities. It turned out that the alkali homogeneous catalysis is preferred yielding production costs of $0.51/gal, water consumption values of 0.47 gal/gal and energy consumption of 2.81 MJ/gal.

2.3 *Lignocellulosic Biomass*

2.3.1 The Biomass

In our race towards sustainability we look for alternative raw materials that provide the same building blocks as grain or seeds, but with no interference with the food supply chain. Lignocellulosic raw materials such as corn stover, miscanthus or switchgrass are an alternative source of sugars and biobased chemicals. The structure of all these resources is similar. They have a structure of lignin to support the plant, and within that structure, we can find cellulose and hemicellulose binding it. Cellulose and hemicellulose are sugar polymers. In order to get to them we need to break the structure. Its stability results in the need for energy intensive pre-treatments. Figure 3 shows the structure of the biomass.

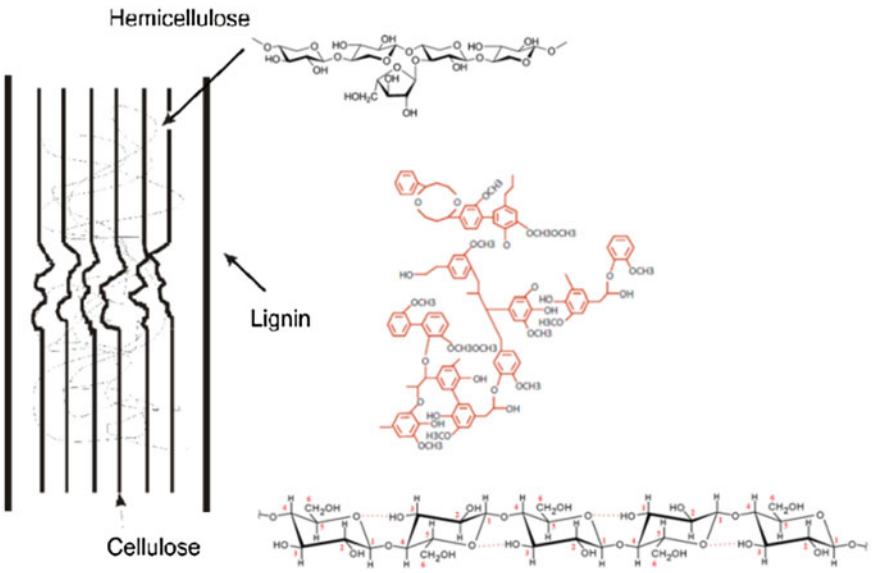


Fig. 3 Lignocellulosic structure

2.3.2 The Pretreatment

Pretreatments are developed depending on the target intermediates. We can aim at the production of liquids, using fast pyrolysis at medium temperature, sugars, for which we need to break the physical structure using moderate pressures and temperatures, or produce syngas at high temperatures, a very versatile building block. For this last option we use gasification. In this section we briefly discuss the path to sugars and to syngas. Pyrolysis is a complex process that results in a wide range of products difficult to tackle due to their corrosivity and viscosity, and hence further upgrading is needed (Brown et al. 2012).

To sugars: As we presented above, the production of sugars out of lignocellulosic raw materials is a challenging process since we just need to break the physical structure. We first grind the material to small pieces for the pretreatment to be effective (Mani et al. 2004). Next we can use a number of alternatives at moderate temperatures, 100–180 °C, and pressures, 12–20 bar, adding different chemicals such as sulfuric acid, ammonia, CO_2, Ozone, etc. With this stage, we can break the biomass into its polymers. There are a number of interesting reviews commenting on different pretreatments (Sung and Chen 2002; Keshwani and Chen 2009). Among them the use of dilute sulfuric acid, developed by the NREL (Aden and Foust 2009; Kazi et al. 2010), the Ammonia Fiber Explosion, AFEX (Alizadeh et al. 2005), or organosolv pretreatments (Zhao et al. 2009) are the most widely used. Organosolv uses a number of solvents, mainly alcohols, to fractionate the lignocellulosic biomass extracting the lignin. Different variants are available such as organic acid, organic peracid, peracetic acid, or high boiling point alcohols pretreatment (Zhao et al. 2009). AFEX consists of using a solution of ammonia and water at 20 bar and 90–180 °C so that in the expansion of the ammonia, the biomass breaks down (Alizadeh et al. 2005). Until recently, ammonia recovery was the most expensive part of the process. The dilute acid consists of the use of steam and a solution of 0.5–2 % of sulfuric acid to break the lignocellulosic structure operating at 12 bar and 140–180 °C. Sugar dehydration may occur generating inhibition species such as furans (Aden and Foust 2009; Kazi et al. 2010).

After any pretreatment, the polymers are exposed to the enzymes to break them into glucose, xylose, and other sugars. The hydrolysis takes place at 50 °C. It is characterized by endothermic reactions, Eqs. (4)–(5). The sugars are basic building blocks for a number of chemicals as we will see later.

$$(C_6H_{10}O_5)_n + nH_2O \longrightarrow nC_6H_{12}O_6 \quad \Delta H = 22.1n \text{ kJ mol}^{-1} \tag{4}$$

$$(C_5H_8O_4)_m + mH_2O \longrightarrow mC_5H_{10}O_5 \quad \Delta H = 79.0m \text{ kJ mol}^{-1} \tag{5}$$

To Syngas: This type of pretreatment is the extreme one in the sense that we use high temperature, 850–1000 °C, to obtain carbon monoxide and hydrogen as building blocks. See Fig. 4 for a scheme of the process and the alternatives. Actually, syngas is quite a versatile building block. Based on Fischer Tropsch (FT) type of catalysis, alcohols, and hydrocarbons can be obtained (Dry 2002). To produce

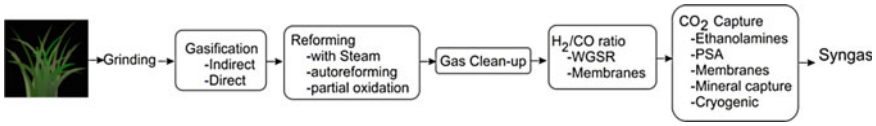


Fig. 4 Alternatives to process lignocellulosic biomass into syngas

syngas, biomass is gasified. A number of technologies are available based on the design of the gasifier and the operation (Brigdwater 1995). For biomass processing, we can highlight the use of direct gasification (Dutta and Phillips 2009) and indirect gasification (Phillips et al. 2007). Indirect gasification is based on the use of the Ferco Battelle Gasifier. The system consists of a fluidized bed gasifier, using steam as gasification agent, and sand, olivine, to provide the energy for the gasification. The char generated in the gasification and the sand is transferred to a combustor where, by burning the char, energy is obtained to reheat up the sand. Hot sand is sent back to the gasifier to provide the energy for gasification. The gasifier operates close to atmospheric pressure but yields high throughput per unit volume of the reactor. The gas generated is rich in small hydrocarbons though. The direct gasification uses the Renugas gasifier from the Gas Technology Institute (GTI). Steam and oxygen are used to gasify the biomass. The gas contains larger proportion of CO_2 but lower concentration of hydrocarbons. It typically operates at 20 bar.

Once the raw syngas is produced, the hydrocarbons are eliminated via reforming. Steam reforming, partial oxidation or autoreforming are available technologies. Steam reforming is endothermic but with higher yield to hydrogen. Partial oxidation, on the other hand, is exothermic but with lower yield to hydrogen. Autoreforming is adiabatic combining exothermic and endothermic reactions. After reforming the syngas, it must be purified from solids, using filters or scrubbers, and its composition may need to be adjusted depending on the final use. For instance, methanol production requires a H_2 to CO ratio of around 2, as well as the production of FT liquids. Ethanol and DME production, on the other hand, require a ratio of 1 (Martín and Grossmann 2013c). In all the cases, the CO_2 and the H_2S must be removed to avoid poisoning the catalyst.

CO_2 capture is a topic of increasing interest and a number of technologies are available (USDOE 2015). Absorption in ethanol amines (with chemical reaction), absorption in physical solvents, adsorption using pressure swing adsorption systems, membranes, mineral capture, are among the most widely used. Chemical absorption requires a large amount of energy to recover the solvent. In terms of physical absorption, there are a large number of solvents such as methanol, dimethyl ether of propylene glycol, etc., that operate at high pressures and low temperatures to improve the capture process. Adsorption beds have lower capacities than chemical absorption. The use of membranes for CO_2 capture requires a carrier, such as ethanol amines that must be later regenerated. Cryogenic recovery typically results in high compression and cooling costs, while the use of mineral capture is attracting attention due to the reversible reaction involving CaO, CO_2, and $CaCO_3$.

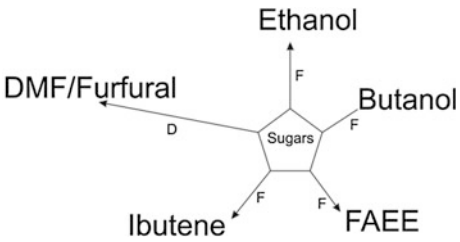


Fig. 5 Main products from sugars. *D* Dehydration; *F* Fermentation

2.3.3 Sugar-Based Products

Sugars are fermented by a number of microorganisms, or dehydrated to obtain higher value products. Figure 5 shows the main products from sugars in the biofuels industry. We can think of the production of wine or beer. In this work we focus on biofuels, and thus our aim is to comment on alternative fuels based on sugars, but lately, the tight benefits have lead to extending the scope. Actually, we could have described some of these products from grain as sugar source. However, the move from first generation to second generation to avoid competition with the food chain has stopped the development of processes from grain-based glucose.

Bioethanol The production of ethanol from lignocellulosic-based sugars, mainly glucose and xylose, takes place at 32–38 °C anaerobically and 1.2 bar, to avoid air from entering the fermentor. It is an exothermic reaction catalyzed by *Zymomonas Mobilis*. The ethanol generated inhibits the process and concentrations from 6–8 % of ethanol in water are typically obtained, although values up to 12 % are expected (Dimian 2008). Apart from ethanol, a number of other chemicals are produced. Table 2 shows the reactions and the conversions to each one.

Once the ethanol−water mixture is obtained, the dehydration step is similar to first generation ethanol, see Fig. 1. It consists of a multieffect distillation column, followed by molecular sieves to produce fuel grade ethanol. The lignin is not degraded in the process, and it can be used as energy source for the process. Martín and Grossmann (2012) optimized this process resulting in a net positive energy balance, using the lignin to provide energy. The flowsheet involved dilute acid pretreatment followed by hydrolysis, sugar fermentation, and ethanol dewatering. A production cost of 0.8 $/gal was reported together with $169 MM of investment for 60 MMgal/year of ethanol. The water consumption of this process turned out to be 1.7 gal/gal (Martín et al. 2011), due to the energy intense pretreatment and dehydration step together with an exothermic reaction operating at low temperature (fermentation).

FAEE Recently Paap et al. (2013) presented interesting results for the production of biodiesel, FAEE, from aerobic fermentation of sugars at 32–38 °C. The main reactions shown in Table 3 are based on the evaluation of the fermentation products and their yield (Martín and Grossmann 2015). The advantage of the production of biodiesel is the inexpensive separation of the product from water

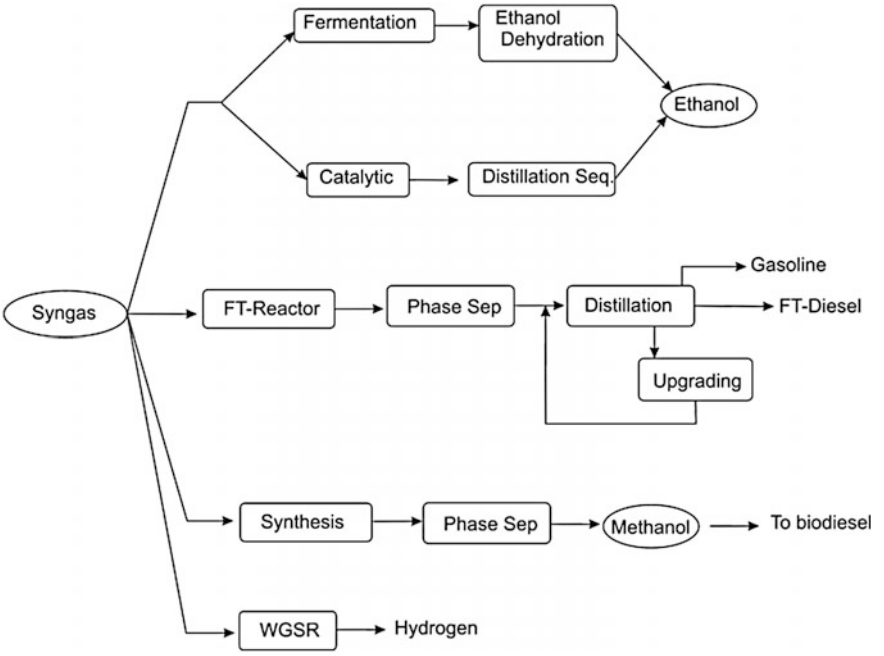


Fig. 6 Uses of syngas

Table 2 Main reactions in the production of second generation ethanol

Reaction	Conversion
Glucose → 2 Ethanol + 2 CO_2	Glucose 0.92
Glucose + 1.2 NH_3 → 6 Z. mobilis + 2.4 H_2O + 0.3 O_2	Glucose 0.04
Glucose + 2 H_2O → Glycerol + O_2	Glucose 0.002
Glucose + 2 CO_2 → 2 Succinic Acid + O_2	Glucose 0.008
Glucose → 3 Acetic Acid	Glucose 0.022
Glucose → 2 Lactic Acid	Glucose 0.013
3 Xylose → 5 Ethanol + 5 CO_2	Xylose 0.8
Xylose + NH_3 → 5 Z. mobilis + 2 H_2O + 0.25 O_2	Xylose 0.03
3 Xylose + 5 H_2O → 5 Glycerol + 2.5 O_2	Xylose 0.02
3 Xylose + 5 CO_2 → 5 Succinic Acid + 2.5 O_2	Xylose 0.03
2 Xylose → 5 Acetic Acid	Xylose 0.01
3 Xylose → 5 Lactic Acid	Xylose 0.01

using a centrifuge. On the other hand, currently the conversions to biodiesel are low, and the yield to other products such as ethanol or glycerol results in a loss in raw material since it is not economically interesting to recover them. The

Table 3 Main reactions in the production of FAEE

Reaction	Conversion	DH_r (kJ/mol)
9 Glucose + 2 O_2 → 2 $C_{18}H_{36}O_2$ + 18 CO_2 + 18 H_2O	Glucose 0.3	−405
Glucose → 2 Ethanol + 2 CO_2	Glucose 0.3	−84,394
Glucose + 1.2 NH_3 → 6 Z. mobilis + 2.4 H_2O + 0.3 O_2	Glucose 0.1	NA
Glucose + 2 H_2O → 2 Glycerol + O_2	Glucose 0.3	504
27 Xylose + 5 O_2 → 5 $C_{18}H_{36}O_2$ + 45 CO_2 + 45 H_2O	Xylose 0.2	−338
3 Xylose → 5 Ethanol + 5 CO_2	Xylose 0.2	−74,986
Xylose + NH_3 → 5 Z. mobilis + 2 H_2O + 0.25 O_2	Xylose 0.2	NA
3 Xylose + 5 H_2O → 5 Glycerol + 2.5 O_2	Xylose 0.2	418

production cost of biodiesel is 3.6 $/gal with an investment of $178 MM for current yields for 10 MMgal/year of FAEE.

Ibutene The interest in the production of ibutene is not that much in the biofuel industry but as intermediate for polymerization or for diesel substitutes production. Typically obtained from the C4 fraction of crude, ibutene has recently been used in the production of glycerol ethers (Cheng et al. 2011) In order to avoid the use of raw materials from fossil resources in the production of alternative fuels, Van Leeuwen et al. (2012) presented the production of ibutene from glucose fermentation. The reaction is as follows:

$$C_6H_{12}O_6 \longrightarrow \text{yeast} \longrightarrow C_4H_8 + 2CO_2 + 2H_2O + 2ATP$$
$$\Delta H = -41 - 9\ \text{kJ mol}^{-1}$$

It takes place at 32–38 °C at atmospheric pressure using *S. Cerevisiae*. The reaction time is around 24 h. The advantage of this chemical, compared to ethanol or even FAEE, is that the product is a gas. Therefore, the separation from the reaction medium is straightforward. Next, ibutene and CO_2 are separated using PSA or membranes. The economics of the process depends on the possibility of using glucose and xylose to obtain this chemical. There is no experimental evidence of the fermentation of xylose to ibutene. Switchgrass contains cellulose and hemicellulose that can be used to obtain a source for glucose and xylose. If we can use only glucose to produce ibutene, the cost per kg of ibutene is $0.75. If we can use both sugars, the production cost drops by almost half to $0.45/kg. Another possibility is to use the xylose for the production of ethanol, and therefore in this case we have a multiproduct facility. The production cost of ibutene can be as low as $0.39/kg considering the credit from ethanol (Martín and Grossmann 2014a) The investment cost of the three options discussed are $122 M, $143 M and $188 M, respectively, using the same flowrate of raw material than that used for 60 MMgal/year of ethanol via hydrolysis.

Butanol Sugars can be fermented to produce a mixture of butanol, acetic acid, and ethanol. It is the so-called ABE fermentation. Clostridium is used as the microorganism to produce such a mixture via two steps acidogenesis and

solventogenesis (Patakova et al. 2013). The mixture produced out of the fermentation is complex and the separation involves liquid–liquid separation and several distillation stages (Kraemer et al. 2011).

HMF and DMF In the attempt to look for higher added value products, advanced fuels, and intermediates, Dumesic's group has extensively worked on the dehydration of sugars, glucose, and xylose, to produce hydroxymethyl furfural (HMF) and dimethyl furfural (DMF) from glucose and furfural from xylose (Roman-Leshkov et al. 2007; Roman-Leshkov and Dumesic 2009). The dehydration of C6 sugars is easy when fructose is the raw material. However, glucose is the C6 naturally obtained. Isomerization using catalysts such as $CrCl_3$ and $CrCl_2$ is a feasible alternative. The advantage of these catalysts is that it is possible to use them in the dehydration of xylose too. Typically, the dehydration takes place at around 160–180 °C using different acids as catalyst and a two phase system to help separate the product. Various organic solvents such as butanol, THF can be used. Once furfural and/or HMF are produced, a complex purification sequence is followed to recover the solvent and the products involving liquid–liquid separation, and several distillation steps. Subsequently, DMF can be produced from HMF by hydrogenation. The reaction takes place using an organic solvent and an excess of hydrogen operating at 120 °C and 16 bar (Kazi et al. 2010). Martín and Grossmann (2015b) evaluated the production of DMF and furfural from algae and switchgrass as source for sugars. Using switchgrass as raw material the production of cost of DMF and furfural is $3/kg ($570 MM of investment cost), while for algae we present the results later in the chapter.

2.3.4 Syngas-Based Products

Syngas is a very versatile raw material or building block to produce chemicals. In this section, we present the major uses of the lignocellulosic-based syngas. In Fig. 6, the block diagram is presented. We also include the production of methanol that will be described later from glycerol.

Bioethanol The mechanism to produce sugars is similar to that presented for Fischer–Tropsch hydrocarbons. Actually, the idea is to grow the carbon chain by constantly adding CO to the previous piece on the surface of the catalyst. The production of a particular chemical is based on controlling the growth and the termination of the chain. Thus, the production of ethanol from syngas is known as mixed alcohols synthesis. The main reactions can be seen below. The reactor operates at 300 °C and 68 bar. Higher temperature favors the production of shorter chemicals. The reactions and conversions of the mixed alcohols synthesis can be seen below (Phillips et al. 2007). As it can be seen, a part from ethanol, methanol, propanol, and butanol are produced. This fact reduces the yield since, while the unreacted syngas can be recycled and reused, the alcohols must be separated using a distillation column sequence and sold to the market as a mixture. Dutta et al. (2009) and Phillips et al. (2007) simulated the production of ethanol via this path, but using two different gasification schemes (direct and indirect respectively) showing

promising productions costs for ethanol around \$1/gal. Later Martín and Grossmann (2011) optimized the production of bioethanol from switchgrass and the best option found was the use of direct gasification followed by steam reforming of the hydrocarbons for a production cost of \$0.41/gal based on the credit that the production of hydrogen could provide. The investment cost turned out to be \$335 MM for a 60 MMgal/year of ethanol. The facility showed a net positive energy balance of 8.5 MJ/gal.

$$\begin{aligned}
&CO + H_2O \longrightarrow H_2 + CO_2; \quad Con_CO_2 = 0.219\\
&CO + 2H_2 \longrightarrow CH_3OH; \quad Con_MetOH = 0.034\\
&CO + 3H_2 \longrightarrow CH_4 + H_2O; \quad Con_CH_4 = 0.003\\
&2CO + 4H_2 \longrightarrow C_2H_5OH + H_2O; \quad Con_EtOH = 0.282\\
&2CO + 5H_2 \longrightarrow C_2H_6 + 2H_2O; \quad Con_C_2H_6 = 0.003\\
&3CO + 6H_2 \longrightarrow C_3H_7OH + 2H_2O; \quad Con_PropOH = 0.046\\
&4CO + 8H_2 \longrightarrow C_4H_9OH + 3H_2O; \quad Con_ButOH = 0.006\\
&5CO + 10H_2 \longrightarrow C_5H_{11}OH + 4H_2O; \quad Con_PentOH = 0.001
\end{aligned}$$

However, syngas can also be fermented (Klasson et al. 1991). The fermentation, as any other fermentation for the production of ethanol presented above, takes place at 32–38 °C following the reaction below

$$3CO + 3H_2 \longrightarrow C_2H_5OH + CO_2$$

The H_2 to CO ratio needed for both synthetic paths is 1 and, as a result, there is an excess of hydrogen in the syngas that, if recovered, can provide a promising credit. The concentration of ethanol in the water mixture is 5 % at the most, which is an important burden in terms of energy consumption for the process. After the fermentation, the ethanol must be dehydrated. The same scheme described above is used, a multieffect distillation column followed by molecular sieves. The low concentration of ethanol in the water is a drawback for this option. The net energy balance of this process is negative, 27 MJ/gal, and the production cost \$0.81/gal using direct gasification and steam reforming with an investment required of \$260 MM for a 60 MMgal/year of ethanol (Martín and Grossmann 2011).

FT Fuels The production of FT fuels is an interesting alternative for the use of syngas. The composition and the operating conditions at the reactor determine the main product, from gasolines, to diesel or heavier oil. The Anderson–Schulz–Flury model is a simple approach to predict the products distribution (Schulz et al. 1999). For instance, the production of diesel fraction requires lower temperatures, 200 °C, 30 bar, a H_2 to CO ratio of 1.7, and the proper catalysts based on cobalt or iron. Gasoline products are typically obtained at higher temperatures and using iron-based catalysts. Apart from hydrocarbons, water and flue gas are produced. Liquid–liquid separation withdraws water from the mixture before hydrocarbon fractionation into gasoline and diesel. To increase the yield of the process to lighter

fractions, the heavy products can be upgraded using catalytic cracking, hydrocracking, or similar (Dry 2002). The net energy balance of the process is positive based on the production of flue gas. The production cost of fuels is higher than that of ethanol, \$0.72/gal, but the investment cost is lower, \$216 MM, for the same biofuels production capacity.

Hydrogen In the previous examples, we have presented the production of syngas as a mixture of CO and hydrogen; we can use the WGS reaction to drive the mixture of gases to H_2 and CO_2. Using a membrane reactor (Ji et al. 2009) and taking into account that the small size of the hydrogen plays an important role, it is possible to obtain a fairly pure stream of hydrogen with a reduced number of purification stages. Therefore, it is possible to produce hydrogen at \$0.68/kg and with a reduced investment, \$148 MM for 60 MMkg/year (Martín and Grossmann 2013c).

2.4 Algae

Algae are a rich raw material with a high yield compared to other biomass sources (Chisti 2007). Its composition as biomass consists basically of carbohydrates (starch), lipids, and proteins. The biodiesel industry has pushed the development of methods to increase the oil accumulated by modifying the algae growth. However, algae can also be used to produce ethanol due to the starch content. As a result, algae can be a good source for ethanol. During the 1970s, the aquatic program of the US government studied the production of algae as a source for fuels. Their conclusions were simple; the low cost of crude oil and the easy access to it resulted in closing the program (Sheehan et al. 1998). The main problem is therefore the cost for growing and harvesting the algae. There are a number of options to grow the algae. We can distinguish between raceway ponds (circular, tanks, paddlewheel raceways) or photoreactors (airlift, tubular, bag cultures). The first ones are simple civil engineering structures with small depth that are filled with water where the algae grow using CO_2 or other carbon source and nutrients (Sazdanoff 2006). The algae growth depends on the solar incidence and the carbon intake. They have important evaporation losses, poor control, and contamination. Furthermore, the concentration of algae in the pond is quite low, 0.1–0.5 h per L, and the energy required to harvest the algae is so high that the production cost per gallon is prohibitive. Another option is the use of photoreactors. The design of this equipment is more complex to allow solar energy to reach the algae. They are transparent pipes or bags of small diameter allocated in structures to provide large volumes. The advantage of this system is the controllability. They are closed units not subjected to direct atmospheric contamination as the ponds are. The main drawback is the higher cost and more complex design for high-volume production, and the fact that the growth of algae in the walls blocks the light. In both cases, algae harvesting is carried out using flotation tanks and a drier to reduce the water content to 5–10 % so that we can extract the oil from the biomass (alfalaval.com, Brennan

and Owende 2010; Mata et al. 2010). The higher concentration of algae in the case of using photoreactors reduces the processing costs.

Algae can be used to produce syngas via gasification. From that, we can use it as described above in the case of lignocellulosic raw materials. Another option is the thermochemical liquefaction producing liquid fuels directly at 300–350 °C and 5–20 MPa. Pyrolysis is also an option to produce biooil form algae operating at 350–700 °C, but has the same drawbacks related to the wide range of products and the complex composition. A part from these thermal processing, we can follow biochemical conversion such as anaerobic digestion, alcoholic fermentation, photobiological hydrogen production, and biodiesel production.

The Anaerobic fermentation (AD) of any biomass generates the so-called biogas consisting of methane and CO_2. The AD is a process consisting of three steps: hydrolysis, fermentation, and methanogenesis, so that the biomass is broken into sugars, then fermented into alcohols, acetic acid, and gas, mainly H_2 and CO_2 and finally the mixture is metabolized to produced methane and CO_2. Typically, 50–70 % of methane by volume can be obtained. The presence of protein in the algal biomass reduces the yield, but it can be mitigated by adding another biomass with a high C to N ratio.

Algae, because of its content of starch, can be fermented into ethanol. The ethanol water mixture, typically between 10–15 % ethanol, must follow the same dehydration process discussed for any other biomass presented in the chapter. The solid residue can be use as cattle fed. Thus, algae species with high starch content are good for this option (Brennan and Owende 2010). We have seen in the case of lignocellulosic that the sugars can be used to produce a number of other chemicals such as FAEE, ibutene for its further use (de la Cruz et al. 2014), DMF at the cost of $1.98/gal ($693 MM of investment) (Martín and Grossmann 2015b).

We can also produce hydrogen. During photosynthesis, microalgae produce protons and oxygen from water. Hydrogenase enzymes convert the protons into hydrogen (Brennan and Owende 2010).

Apart from all these options, algae are typically devoted to the production of oil for biodiesel production. Oil has to be extracted from the algae biomass. The extraction is similar to that employed for seeds consisting of mechanical action and the use of solvents. As a result, the price per gallon of algae oil is high (Klise et al. 2011); however, it is expected that values of $0.07/lb could be reached (Pokoo et al. 2010). Actually, the use of capillarity-based harvesting, proposed by Univenture Inc can help achieve values of $0.06/lb (Martín and Grossmann 2012). Once the oil is available, we can produce biodiesel via transesterification as described above. The production cost of biodiesel from algae actually depends heavily on the cost of harvesting. The use of advanced harvesting methods can reduce the production cost to $0.42/gal with investment cost dominated by algae growing around 110 MM$ per 72 Mgal/year of biodiesel (Martín and Grossmann 2012).

2.5 Wastes: Biogas

Anaerobic digestion is one of the most energy efficient and environmental friendly processes for the production of energy from biomass. Anaerobic digestion is a biological process performed by many classes of bacteria on a large number of biomass types from algae, as indicated above, to wastes. It consists of four steps: hydrolysis, acidogenesis, acetogenesis, and methanogenesis. In particular, it is an interesting technology to obtain further energetic value from water treatment sludge, cattle manure, and others. The process can be mesophilic, running at 20–40 C for more than a month, or thermophilic, at 50–65 C. It consists of hydrolysis, fermentation, and methanogenesis. The entire process lasts for several days. The actual biogas composition, from 50–70 % in methane and 30–40 % CO_2, with small amounts of N_2, O_2, H_2S, and NH_3 depends on the raw material, as well as the yield from biomass to biogas, Gunaseelan (1997), Steffen et al. (1998).

3 Integrated Processes

For many years, process design was focused on the transformation of a biomass type or source into one main product, i.e., bioethanol, as some byproducts from the same process. The tight margins and the competitive fuels and energy markets have altered that trend. Actually, chemical complexes consist of a large number of processes operating in a symbiotic way to take advantage of the excess of energy, the possible use of the byproducts and the integration of technologies. Thus, the biofuels industry has followed the same trend, and with the help of process system engineering techniques a number of integrated processes have been developed. In this section, we discuss some of those processes paying special attention to the integration of energy, raw materials, and technologies.

3.1 First and Second Generation Bioethanol

Bioethanol has been produced from corn in the US for the last decades (Jacques et al. 1999). The concerns related to the use of food as a raw material for the production of fuels have led to the development of technologies to process biomass that do not interfere the food market such as lignocellulosic energy crops. Among them, let us focus on corn stover. In the production of corn, grain represents 45 % of the total biomass produce. The remaining 55 % is stover, a lignocellulosic type of biomass that can be used as lignocellulosic raw material for sugars. In order to maintain minerals in the ground, it is not recommended to harvest more than 70 % of this mass. Therefore, around 85 % of the biomass grown in the production of corn can be used in an integrated facility (Atchinson 2003; Nielsen 2009). The

integrated process that uses the entire corn plant can benefit from the symbiosis of first and second generation production processes. In the first place, it will increase the production capacity by almost 100 %. Typically, second generation bioethanol presents a positive net energy balance. The excess of energy, either due to the exothermic reactions at high temperature from the gasification or the synthetic reactor, or the use of lignin as source for energy, can provide the energy for the dehydration of the ethanol produced in the fermentation of glucose from corn grain. In case the bio-path is followed from the lignocellulosic part of the biomass, there are a number of steps, the dehydration involving beer column and molecular sieves, that are common, and thus economies of scale improve the economics of the process. Cucek et al. (2011), integrated first and second generation technologies to process the grain and the stover to evaluate the integrated biorefinery. The most economical process used the thermochemical route for processing the stover so that the excess of energy could be used to feed the beer column used in dehydration of the grain-based ethanol. Ethanol was produced at $1.22/gal. Water consumption was also limited to 1.56 gal/gal.

3.2 *Algae-Based Fuels*

In the previous section the rich composition of the algae, typically starch, up to 40 %, lipids, up to 75 %, and protein (Mata et al. 2010). There is a large number of species that can be grown in marine or freshwater. This composition allows for integration of the production of several products within the same biorefinery complex so that no fossil-based chemical or raw material is needed. The basic process is the production of ethanol and biodiesel from algae (Martín and Grossmann 2013a). Figure 7 shows a scheme of the processes described below.

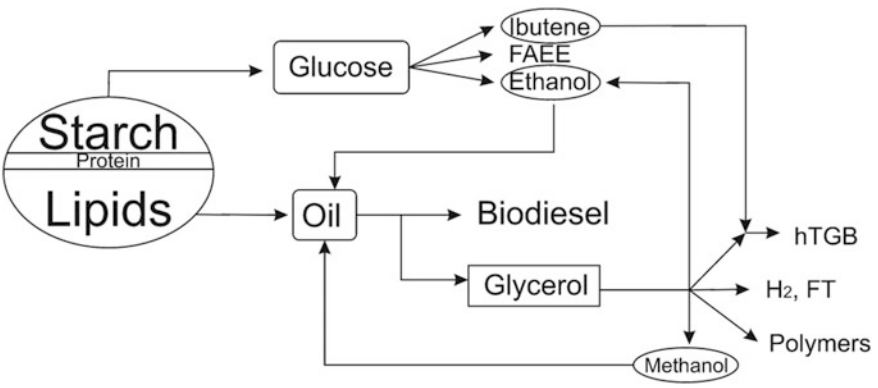


Fig. 7 Biorefinery based on algae

3.2.1 Ethanol and Biodiesel

Severson et al. (2013) showed that the production of biodiesel, Fatty acid ethyl ester, or FAEE, was not only technically feasible, but also competitive for certain prices of ethanol as transesterifying agent. Algae provide the opportunity to produce ethanol from the starch, and oil from the lipids. Therefore, we have all the ingredients to obtain FAEE from algae with no need to buy fossil-based methanol. One thing remains open, what is the optimal composition of the algae in terms of starch and lipids, for the operation of such an integrated facility. Again, the PSE approach provided the tools to come up with a model that allowed computing the optimal algae composition for the simultaneous production of ethanol and biodiesel. The equilibrium that governs the transesterification of the oil suggests including simultaneous optimization and heat integration within the formulation (Duran and Grossmann 1986). Thus, the biorefinery consists of algae growing. Next, the oil is extracted using mechanical action and hexane. The starch is sent to be processed following first generation bioethanol technologies, namely, liquefaction followed by saccharification and glucose fermentation to ethanol. Once the ethanol is produced, proteins are removed from the liquid stream to be sold and the ethanol is dehydrated using a multieffect column and molecular sieves. Part of the ethanol is sold, if there is an excess, and the rest is used in the transesterification of the oil extracted from the algae. The high energy required in ethanol dehydration, and the high yield from oil to biofuels suggested a composition of 60 % lipids, 30 % starch and 10 % protein dry biomass basis. Biodiesel is the main product of the integrated facility. The production cost of biodiesel was $0.35/gal using 4 MJ/gal for a production capacity of 90 Mgal/year using enzymatic catalysts. Ethanol represented 9 % of the total biofuels sold since part of it was internally used. The investment of such a plant added up to 180 MM$. The use of alkali catalyst was cheaper by 3c $/gal, but the energy consumption was 50 % higher. (Martín and Grossmann 2013a) This process has two main byproducts, protein and glycerol. High added value products can be obtained from protein, the glycerol represents an interesting source of carbon for the production of chemicals.

3.2.2 Use of Glycerol

Production of methanol One of the main drawbacks of the production of biodiesel in current industrial processes is the use of methanol as transesterifying agent. Methanol is typically produced from coal or natural gas via syngas production. However, the bioproduct of the process, glycerol can be reformed to produce syngas, see Fig. 6. Therefore, we can reduce the dependency on fossil-based ethanol using the glycerol as raw material for the production of methanol. The process consists of glycerol reforming with steam, with a mixture of pure oxygen and steam (autoreforming) or aqueous phase glycerol reforming with water in liquid phase to produce syngas. Next, the syngas must be cleaned up for hydrocarbons removal, its composition adjusted to achieve a H_2 to CO ratio of 2, and the CO_2 is

partially removed, since the catalyst works better with concentrations of CO_2 from 3–8 % in volume. Methanol synthesis is governed by two main reactions in equilibrium. The production can be carried out in gas phase at low pressure, or liquid phase, at high pressure.

$$CO + H_2 \leftrightarrow CH_3OH$$
$$CO_2 + H_2 \leftrightarrow CO + H_2O$$

Finally, the unreacted gas is recycled and the methanol–water mixture separated. The optimal process allowed saving 50 % of fossil-based methanol using autoreforming. However, the production cost of biodiesel increased with respect to its production alone by 0.15$/gal and the energy consumption went up to 3.55 MJ/gal and the investment cost by less than 10% (Martín and Grossmann 2013d).

Hydrogen and FT fuels The syngas from glycerol can also be used to increase the yield to fuels by producing hydrogen and FT fuels. The processing technologies have already been described above. However, this use is not competitive with the use of switchgrass for the production of the same fuels (Martín and Grossmann 2014b).

Production of ethanol Glycerol can be fermented anaerobically at 38 °C to ethanol using E. Coli for 60 h. Ethanol can reach 10 % in water (Dharmadi et al. 2006)

$$C_3H_8O_3 \rightarrow C_2H_6O + H_2 + CO_2 \ (98\,\%)$$
$$CO_2 + C_3H_8O_3 \rightarrow \text{Succinic Acid} + H_2O \ (1\,\%)$$
$$C_3H_8O_3 \rightarrow 3.0075(\text{Biomass}) + H_2O \ (1\,\%)$$

The water–ethanol mixture is, in some sense, similar to the production of ethanol from sugars. Therefore, we can develop an integrated process that produces ethanol from the algae starch as well as from glycerol. The dilute ethanol mixture from both sources is dewatered using a beer column and molecular sieves. Economies of scale are an asset for this process. On the other hand, the high energy consumption required for ethanol dehydration is the drawback of this option. Part of this ethanol is used internally to transesterify the oil from the algae lipids and the rest is sold. The optimal algae composition for the operation of this facility consists of 60 % lipids, 30 % starch and 10 % protein, dry biomass basis. Ethanol production increases by 50 % using the glycerol, compared to the option that does not further uses it, while the cost of biofuels adds up to $0.45/gal. The energy consumption increases up to 4.2 MJ/gal and the investment goes up to 211 MM$ (Martín and Grossmann 2014c).

Production of high glycerol ethers To increase the yield from oil to fuels, it is possible to use the glycerol to obtain glycerol ethers. The reaction is an equilibrium that typically uses ibutene to etherify the crude glycerol (Behr and Obendorf 2001).

$$\text{Glycerol} + \text{iButene} \underset{k_{-1}}{\overset{k_1}{\longleftrightarrow}} \text{MTBG}$$

$$\text{MTBG} + \text{iButene} \underset{k_{-2}}{\overset{k_2}{\longleftrightarrow}} \text{DTBG}$$

$$\text{DTBG} + \text{iButene} \underset{k_{-3}}{\overset{k_3}{\longleftrightarrow}} \text{TTBG}$$

Several papers (Cheng et al. 2011; Vlad et al. 2010; Martín and Grossmann 2014d) have proposed processes to transform glycerol into di- and tri-tertbutyl glycerol (hTBG), since both can be used as diesel substitutes. The process actually uses glycerol as separation agent of the mixture of ethers, mono, di, and tri. The glycerol and monoether are recycled back to the reactor, while the ibutene with the di and tri ethers are further separated recycling the ibutene and purifying the hTBG. The initial drawback of this process is the use of ibutene, a typical C_4 chemical. However, as presented above, we can produce it from sugars. That possibility opens the opportunity to synthesize a process based on algae with no need for fossil-based chemicals. We grow the algae in such a way that we produce starch to be used to obtain glucose. 50 % of the glucose will be used for ethanol production and the rest to produce ibutene, while the ethanol once dehydrated as before, is used to trans-esterify the oil, the ibutene is later used to etherify the glycerol. In this way, we do not need any fossil-based chemical for the production of diesel substitutes, biodiesel (FAEE) and hTBG's. Furthermore, the integrated facility is not only environmental friendly, but it also reduces the production cost from $1/gal (Martín and Grossmann 2014d) to $0.46/gal (de la Cruz et al. 2014). On the other hand, the investment cost of the facility increases from $167 MM (Martín and Grossmann 2014d) to $205 MM (de la Cruz et al. 2014) if ibutene is internally produced.

Polymers In the last examples, we have seen that the use of glycerol for enhancing the production of fuels is an interesting option but has no economic incentive. There are other possibilities such as the use of glycerol for the production of high added value products out of glycerol. For instance, we can produce polyesters from glycerol using adipic acid. Out of glycerol and adipic acid liquid or solid polymers can be produced for paintings or as adsorption materials (Bueno et al. 2015). This alternative is more profitable; the production cost is around 1.7 €/kg and the selling price adds up to 5 €/kg Apart from this PHB can also be produced via fermentation (Ibrahim and Steinbüchel 2009). Economic evaluations of this option are still in progress.

3.3 Multiproduct Processes from Lignocellulosic Biomass

Actually this topic is not new. FT production processes discussed above are capable of producing a wide range of products, and among them, substitutes for crude-based gasoline and diesel. However, we can extend this study to those other processes that

generate ethanol, FAEE and ibutene. In a second step, since syngas is a building block that can be produced from several renewable and nonrenewable sources alike. In this section, we discuss the use of shale gas and biomass to produce FT fuels.

Simultaneous production of ibutene, FAEE, and ethanol The production of ethanol from lignocellulosic biomass is the base case for second generation biorefineries. However, such a process is not flexible to meet the demand for fuels in the biorefinery since with ethanol we only substitute gasoline. Furthermore, the economics of biofuels is not still competitive with crude-based fuels. Based on the versatile use of sugars, as presented in the previous section, we can find in the literature two integrated biorefineries of interest. On the one hand, we can simultaneously produce FAEE and ethanol from switchgrass (Martín and Grossmann 2015a). In this sense, the integrated facility can produce substitutes for gasoline and diesel at the same time. The trade-off involved in this facility is simple. Ethanol dehydration is energy intensive, but the yield from biomass to ethanol is moderate. On the other hand, the separation of FAEE from an aqueous mixture is easy, but the yield to FAEE is low. Therefore, the analysis presented was to evaluate what was the optimal production of both fuels. It turns out that the yield from biomass had larger impact on the economics than the energy savings in product purification. Only if the demand requires the production of both, it is technically feasible to do so, but it is economically not attractive. The production of ethanol and glycerol together with FAEE reduces the yield, but those components were very diluted in water, and therefore recovering them was not economically interesting.

Since the economics of the ethanol is tight and the production of ibutene from xylose has not yet been experimentally validated, it is an interesting possibility to simultaneously produce ibutene and ethanol from switchgrass. The glucose in the biomass is used for ibutene production, while the xylose is devoted to ethanol. In this way, the facility is not only able to produce ethanol at a reasonable price and amount, but the ibutene, whose price from crude oil is close to $2/kg, is possible to be produced below $0.4/kg using renewable raw materials.

Renewable and non renewable source for syngas In this chapter, we have presented the production of syngas from biomass, lignocellulosic, and glycerol. In 2009, there was a large increase in the production of natural gas from shale gas. The availability of this nonconventional natural gas changes the energy market affecting the electricity price. Since shale gas is widely available in the US, and in the same region switchgrass grows natively; it is interesting to evaluate the possibilities of using one and/or the other in order to produce FT liquids. Martín and Grossmann (2013b) used a mathematical programming approach to evaluate the limit in the price of biomass and shale gas for economically producing FT liquids from either feedstocks. A large sensitivity analysis was performed using a base production capacity of 60 Mgal/year. Using as target price for the liquid fuels $1/gal, the study showed that the price of biomass below $100/t is competitive, but above that shale gas is preferred to meet the liquid fuels demand. On the other hand, the price of the shale gas must be below $11.5/MMBTU for it to be an attractive source. Using biomass and shale gas, the dependency of US on foreign crude could be reduced, while all the effort into second generation biofuels is still useful.

3.4 Integrated Solar, Wind, and Biomass

So far we have focused mainly on the production of fuels or chemicals; however, biomass can be used to produce energy too. Here it is where the properties of biomass can be of great advantage compared to solar and wind and thus complement each other. Biomass can be stored. Energy crops are typically perennial, and thus harvesting can be done when biomass is needed. On the other hand, solar and wind energy are available only at certain times of the day or the year, as it has been discussed in other chapters of this book. Therefore, it is straightforward to integrate biomass with solar and wind in order to secure the production of power, and with it chemicals and fuels.

CSP-Biomass The variability of power production from solar energy means that during the year favorable months produce even more than twice the power, than in less favorable seasons. While the demand for power does not follow the production profile, and the processes operate better at a constant flow, the integration between a biomass-based polygeneration system and a concentrated solar power facility allows constant production of power over a year. The excess of energy can be chemically stored. For instance, hydrogen can be produced when solar energy is capable of meeting the demand and the production capacity of hydrogen will vary over the year. By doing this, we reduce the idle sections of the process. The biomass-based polygeneration system consists on the generation of syngas from biomass. The same technologies for biomass processing as those used for the production of second generation bioethanol apply in this case. After syngas cleanup, we use a Brayton cycle to generate power. Furthermore, there is an excess of energy at different parts of the process. For instance, the gases from the combustor or the gasifier, the partial oxidation can provide part of the energy to reheat up the salts within the CPS facility. Figure 8 shows a scheme of the integrated process.

Vidal and Martín (2015) used a mathematical optimization approach to evaluate the operation of such integrated facility. The biomass is processed using indirect gasification followed by steam reforming. The reason is that the syngas produced in this way has more energy within for its use in the Brayton cycle. The streams from the combustor are used to provide heat to the salts, taking advantage of the high temperature of the gases. On the other hand, the CSP facility consists of the solar field, a regenerative Rankine cycle and a wet cooling tower. For the constant production of 340 MW, the production cost of electricity is €0.073/kWh, obtaining together with the electricity, 97 kt/year of hydrogen. The investment for the integrated plant adds up to 2305 M€. Furthermore, in order to maintain electricity costs below 0.1 €/kWh, we need to assure biomass costs below 100 €/t, and/or hydrogen prices above 0.58 €/kg. Similar facilities to the one evaluated are being currently built worldwide.

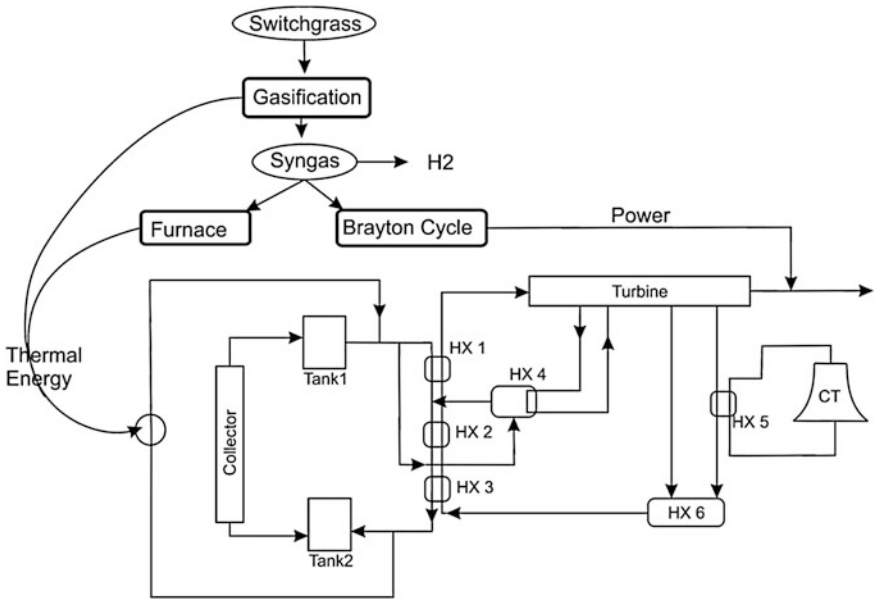


Fig. 8 Integrated CSP-biomass facility for constant power production

PV-Wind—Biomass Solar, wind, and biomass can also be integrated for the production of chemicals using CO_2 as carbon source. One of the alternatives to use CO_2 to produce methane, methanol, or others is to reduce it with hydrogen, For the process to be environmental friendly, CO_2 is not only captured but also reused, the hydrogen must be renewable. We can produce renewable hydrogen from biomass, as was seen in this chapter before, but we can also produce it from water splitting. In this case, the power must be renewable for the hydrogen to be it too. Therefore, Solar photovoltaics, PV solar, wind energy, and biomass can be integrated for the production of hydrogen so that it can be further used to obtain for instance methane. The advantage of producing methane is that we are storing solar and wind energy in the form of a chemical that it is easy to handle and ready to use as fuel. By integrating the three sources, it is possible to maintain constant production of methane independently of the solar or wind availability. See Fig. 9 for a scheme of the integrated process. The excess of energy can be sold into the market. Martín and Davis (2016) evaluated this integrated facility, again, using mathematical programming techniques. The study allowed determining the optimal combination of the three as a function of the biomass price and the cost for the solar and wind energy. Biomass was processed via indirect gasification, and steam reforming to produce syngas. Instead of using it to obtain hydrogen, the model suggested its use to produce power through a Brayton cycle. In parallel, solar was recommended to complement the biomass as long as the biomass price is below 50 €/t, and the

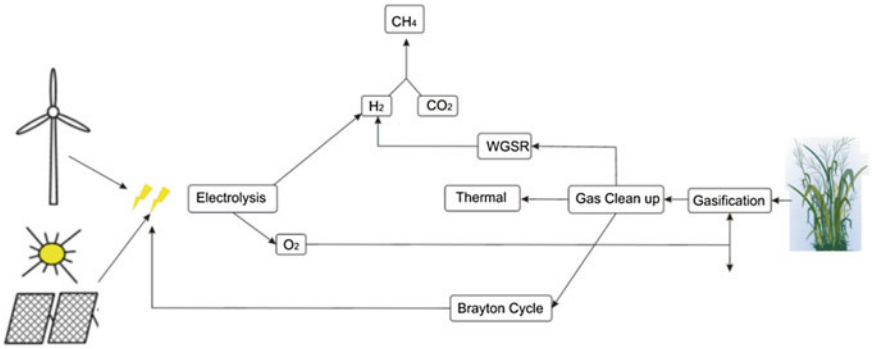


Fig. 9 Integrated system for constant methane production. Reproduced with permission

investment in the biomass section is below 1500 €/kW, and the solar incidence is above 1200 kWh/m^2 year. The use of wind is restricted to low solar incidence, wind velocities above 9 m/s and high prices for biomass.

4 Conclusions

Biomass, due to its rich composition, is a versatile raw material for the production of chemicals that can be used as biofuels. There are a number of alternatives to substitute crude-based gasoline and diesel, but most importantly, the technologies are becoming available and industrially feasible. Furthermore, the production costs are becoming competitive. Although the first attempt was to produce the fuels on their own, the development of integrated processes presents several advantages that exploit the synergies of several processes. Although the price of crude is still low for biomass-based fuels to be used, there is no doubt that sustainability requires the use of alternative sources. Among them, biomass can be stored for extended periods of time and be processed to meet significant demand of fuels.

References

Aden A, Foust T. (2009) Technoeconomic analysis of the dilute sulfuric acid and enzymatic hydrolysis process for the conversion of corn stover to ethanol. Cellulose. (16):535–545.

Ahmetovic E, Grossmann IE. (2011) Global superstructure optimization for the design of integrated process water networks. AIChE J. 57 (2): 434–457.

Ahmetovic E, Martín M, Grossmann IE. (2010) Optimization of Water Consumption in Process industry: Corn – based ethanol case study Ind. Eng. Chem Res. *49* (17): 7972–7982.

Alizadeh, H., Teymouri, F., Gilbert, T.I., Dale, B.E. (2005) Pretreatment of switchgrass by ammonia fiber explosion (AFEX). Appl Biochem Biotechnol. 121–124:1133–1141.

Atchison J, Hettenhaus J. (2003). Innovative methods for corn stover collecting, handling, storage, and transporting. Report No. ACO-1-31042-01, National Renewable Energy Laboratory, Golden, CO.

"BP Statistical Review of World Energy June 2012 http://www.bp.com/content/dam/bp/pdf/Statistical-Review-2012/statistical_review_of_world_energy_2012.pdf."

Behr A, Obendorf L. (2001) Process development for acid-catalyst etherification of glycerol with isobutene to form glycerol tertiary butyl ethers. Chem. Ing. Tech., 73: 1463–1467.

Bridgwater AV. (1995) The technical and economic feasibility of biomass gasification for power generation. Fuel 14(5): 631–653.

Brennan L, Owende P. (2010) Biofuels from microalgae. A review of technologies for production, processing, and extractions of biofuels and co-products. Renewable Sustainable Energy Rev, 14:557–577.

Brown TR, Zhang Y, Hu G, Brown RC. (2012) Technoeconomic analysis of biobased chemicals production via integrated catalytic processing. Biofuels. Bioprod. Biorefin. 6 (1): 73–87.

Bueno L, Toro C, Martín M. (2015) Technoeconomic evaluation of the production of polyesters form glycerol and adipic acid. Chem. Eng. Res. Des. 93, 432–440.

Cheng JK, Lee CL, Jhuang YT, Ward JD, Chien L (2011) Design and control of the glycerol tertiary butyl ethers process for the utilization of a renewable resource. Ind. Eng. Chem. Res. 50: 12706–12716.

Chisti Y (2007) Biodiesel from microalgae Biotechnology Advances 25: 294–306.

Dimian AC, Sorin C. (2008) Chemical process design. Computer e aided case studies. Weinheim: Wiley-VCH. New York.

Cucek L, Martín M, Grossmann IE, Kravanja Z (2011) Energy, water and process technologies integration for the simultaneous production of ethanol and food from the entire corn plant Comp. Chem. Eng. *35:1547–1557.*

De la Cruz V, Hernández S, Martín M., Grossmann I.E. (2014) Integrated synthesis of Biodiesel, Bioethanol, Ibutene and glycerol ethers from algae. Ind. Eng. Chem Res. *53* (28): 11371–11383.

Dharmadi Y, Murarka A, Gonzalez R. (2006) Anaerobic Fermentation of Glycerol by *Escherichia Coli*: A New Platform for Metabolic Engineering. Wiley InterScience; New York. 20 May 2006.

Dry ME (2002) The Fischer-Tropsch process: 1950-2000. Catal. Today 71: 227–241.

Duran M, Grossmann IE (1986). Simultaneous-optimization and heat integration of chemical processes. AIChE J. 32 (1): 123–138.

Dutta A, Phillips SD (2009), Thermochemical Ethanol via Direct Gasification and Mixed alcohol Synthesis of Lignocellulosic Biomass; National Renewable Energy Laboratory: Golden, CO, 2009 No. NREL/TP-510-45913.

Gunaseelan VN (1997) Anaerobic digestion of biomass fro methane production: A review. Biomass and Bioenergy. 13 (12):83–114.

Ibrahim MHA, Steinbïuchel A (2009) Poly(3-Hydroxybutyrate) Production from Glycerol by Zobellella denitrificans MW1 via High-Cell-Density Fed-Batch Fermentation and Simplified Solvent Extraction Appl. Environ. Microbiol. 75 (19): 6222–6231.

Jacques K, Lyons TP, Kelsall DR (1999) The Alcohol Textbook, 3rd ed., Nottingham University Press., United Kingdom.

Ji P, Feng W, Chen B (2009) Production of ultrapure hydrogen from biomass gasification with air. Chem. Eng. Sci. 64: 582–592.

Karuppiah R, Peschel A, Grossmann IE, Martín M, Martinson W, Zullo L (2008) Energy optimization of an Ethanol Plant AICHE J. 54 (6): 1499–1525.

Kazi FK, Fortman JA, Anex RP, Hsu DD, Aden A, Dutta A, Kothandaraman G (2010) Technoeconomic comparison of process technologies for biochemical ethanol production from corn stover. Fuel, 89 (1): S20–S28.

Keshwani DR, Cheng JJ (2009) Switchgrass for bioethanol and other value-added applications: A review Bioresour. Technology, 100: 1515–1523

Klasson KT, Ackerson CMD, Clausen EC, Gaddy JL (1991) Bioreactor Design for Synthesis Gas Fermentations. Fuel, 70 (9): 605–614

Klise GT, Roach, JD, Passell HD (2011) A Study of Algal Biomass Potential in Selected Canadian Regions; Sandia National Laboratories: Albuquerque, NM, 2011; http://prod.sandia.gov/techlib/access-control.cgi/2011/118528.pdf.

Kraemer K, Harwardt A, Bronnenberg R, Marquardt W (2011) Separation of butanol form acetone-butanol-ethanol fermentation by hybrid extraction-distillation process. Comp. Chem. Eng. 35:949–963.

Mani S, Tabil LG, Sokhansanj S (2004) Grinding performance and physical properties of wheat and barley straws, corn stover and switchgrass. *Biomass Bioenerg*. 2004, 27: 339–352.

Martín M, Davis W. (2016) Optimal integration of biomass, wind and solar for the production of synthetic natural gas Comp. Chem. Eng 84: 314–325.

Martín, M., Grossmann, I.E. (2011) "Energy Optimization of Bioethanol Production via Gasification of Switchgrass" AIChE J. 57, 12, 3408, 3428

Martín M, Grossmann IE (2012b) Simultaneous optimization and heat integration for biodiesel production from cooking oil and algae. Ind. Eng. Chem Res. 51 (23): 7998–8014.

Martín M, Grossmann IE (2013a) Optimal engineered algae composition for the integrated simultaneous production of bioethanol and biodiesel AIChE J. 59 (8): 2872–2883.

Martín M, Grossmann IE. (2013b) Optimal use of hybrid feedstock, switchgrass and shale gas for the simultaneous production of hydrogen and liquid fuels. Energy, 55: 378–391.

Martín M, Grossmann IE (2013c) On the systematic synthesis of sustainable biorefineries Ind. Eng. Chem. Res. 52 (9): 3044–3064.

Martín M, Grossmann I.E. (2013d) ASI: Toward the Optimal Integrated Production of Biodiesel with Internal Recycling of Methanol Produced from Glycerol. Environ. Prog. Sust. Energ. 32 (4): 791–801.

Martín M, Grossmann IE (2014a) Optimal simultaneous production of i-butene and ethanol from switchgrass Biomass Bioenerg., 61: 93–103.

Martín M, Grossmann IE (2014b) Optimal Simultaneous Production of Hydrogen and Liquid Fuels from Glycerol: Integrating the Use of Biodiesel Byproducts. Ind. Eng Chem Res. *53* (18): 7730–7745.

Martín M, Grossmann IE (2014c) Design of an optimal process for enhanced production of bioethanol and biodiesel from algae oil via glycerol fermentation App. Energ. 135: 108–114.

Martín M, Grossmann IE (2014d) Simultaneous dynamic optimization and heat integration for the co-production of diesel substitutes: Biodiesel (FAME & FAEE) and glycerol ethers from algae oil. Ind. Eng. Chem. Res. 53; 11371–11383.

Martín, M, Grossmann, I.E. (2015a) Optimal Simultaneous Production of Biodiesel (FAEE) and Bioethanol from Switchgrass Ind. Eng. Chem. Res. *54* (16), 4337–4346.

Martin, M, Grossmann, I.E. (2015b) Optimal production of Furfural and DMF from algae and swichgrass. Ind. Eng. Chem. Res. 10.1021/acs.iecr.5b03038

Martín, M, Ahmetovic, E, Grossmann, I.E. (2011) Optimization of Water Consumption in Second Generation Bioethanol Plants Ind. Eng. Chem. Res. *50*, 3705–3721.

Mata, TM, Martins AA, Nidia S, Caetano NS (2010) Microalgae for biodiesel production and other applications: A review. Renew. Sust. Energy Rev. 14: 217–232.

Meher LC, Sagar DV, Naik SN (2006) Technical aspects of biodiesel production by transesterification- a review. Renew. Sust. Energy Revs. 10: 248–268.

Nielsen (2009). http://www.agry.purdue.edu/ext/corn/pubs/agry9508.htm Accessed May 2010.

Paap SM, West TH, Manley DK, Steen EJ, Beller HR, Keasling JD, Dibble DC, Chang S, Simmons BA (2013) Biochemical production of ethanol and fatty acid ethyl esters from switchgrass: A comparative analysis of environmental and economic performance. Biomass Bioenerg. 49: 49–62.

Patakova P, Linhiva M, Rychtera M, Paulova L, Mezoch K. (2013) Novel and neglected issues of acetone–butanol–ethanol (ABE) fermentation by clostridia: *Clostridium* metabolic diversity, tools for process mapping and continuous fermentation systems. Biotech. Adv. 31 (1): 58–67.

Phillips S, Aden A, Jechura J, Dayton D, Eggeman T (2007) Thermochemical Ethanol via Indirect Gasif ication and Mixed Alcohol Synthesis of Lignocellulosic Biomass; National Renewable Energy Laboratory: Golden, CO, April 2007; No. NREL/TP-510-41168.

Pokoo-Aikins G, Nadim A, El-Halwagi MM, Mahalec V (2010) Design and analysis of biodiesel production from algae grown through carbon sequestration. Clean, Technol. Environ. Policy 2010, 12 (3): 239–254.

Roman-Leshkov Y, Dumersic A. (2009) Solvent effects on fructose dehydration to 5 – Hydroxymethufurfural in biphasic systems saturated with inorganic salts. Top Catal. 52: 297–303.

Roman-Leshkov Y, Barrett CJ, Liu ZY, Dumesic JA (2007) Production of dimethylfuran for liquid fuels from biomass-derived carbohydrates. Nature. 447: 982–986.

Sazdanoff N (2006) Modeling and Simulation of the Algae to Biodiesel Fuel Cycle. Undergraduate Thesis. The Ohio State University, Columbus, OH.

Schulz H, Schaub G, Claeys M, Riedel T (1999) Transient initial kinetic regimes of Fischer-Tropsch synthesis. Appl. Catal., A 186: 215–227.

Severson K, Martín M, Grossmann IE (2013) Process optimization bioDiesel production using bioethanol. AICHE J. 59(3): 834–844.

Sheehan J, Dunahay T, Benemann J, Roessler P (1998) A Look Back at the U.S. Department of Energy's Aquatic Species Program: Biodiesel from Algae NREL/TP-580-24190.

Steffen R, Szolar O, Braun R. (1998) Feedstocks for anaerobic digestion. *Q:\RODL\PROJEKTE \AD-NETT\FEEDNEW.DOC*.

Sun Y, Cheng J (2002) Hydrolysis of lignocellulosic materials for ethanol production: a review. Bioresour Technol. 83:1–11.

Univenture Inc. 2009; http://algaevs.com/ (Accessed June 2012).

USDOE, Chen ZY, Review of CO_2 capture technologies and some improvement opportunities.

Van Leeuwen BNM, van der Wupl AM, Duijnstee I, van Maris AJA, Straathof AJJ (2012) Fermentative production of isobutene. Appl Microbiol Biotechnol 93:1377–1387.

Vlad E, Bildea, CS, Bozga, G (2010) Integrated design and control of glycerol etherification processes. Bull. Inst. Polym. IASI 2010, 4, 139–148.

Vidal M, Martín M. (2015) Optimal coupling of biomass and solar energy for the production of electricity and chemicals. Comp. Chem. Eng, 72: 273–283.

Yee T, Grossmann IE (1990) Simultaneous-optimization models for heat integration. 2. Heat-exchanger network synthesis. Comput. Chem. Eng. 14 (10), 1165–1184.

Zhao X, Cheng K, Liu D. (2009) Organosolv pretreatment of lignocellulosic biomass for enzymatic hydrolysis. Appl. Microbiol. Biotechnol. 82:815–827.

Zhang Y, Dube MA, McLean DD, Kates (2003) M. Biodiesel production from waste cooking oil: 1. Process design and technological assessment. Bioresour. Technol. 89, 1–16.

CO_2 Carbon Capture, Storage, and Uses

Miguel Ángel Delgado and Fabrice Del Corso

Abstract CO_2 has become one of the recent concerns in industry. The extensive use of fossil-based carbon sources has increased the concentration of atmospheric CO_2 over 400 ppm. Carbon capture technologies are being developed to avoid its release; however, apart from capturing and storing the CO_2, it can become a valuable raw material for fuels and chemicals via hydrogenation or decompositions or as a valuable solvent, or even an enhancer in the production and recovery of oil or methane. In this chapter, we describe the reasons why CO_2 is an interesting raw material rather than a waste or an environmental problem and, the current efforts in carbon capture and storage for it to be available for further use.

1 CO_2 Properties and Associated Uses

Carbon dioxide is the media star as a greenhouse gas. By contrast, there are hardly any reports about its special and very useful properties. These properties make this molecule very appropriate for a lot of uses and applications in the industrial or medical fields.

Some applications exist for decades and new promising ones are appearing at laboratory, demonstration, and also industrial scales.

To supply CO_2 for these applications, its emissions are recovered and recycled. The total quantity of CO_2 which can be recycled via all its applications remains limited regarding the current CO_2 emissions. In a very optimistic scenario where all the new potential applications would reach the industrial scale, a small percentage

M.Á. Delgado (✉)
Fundación Ciudad de la Energia, Av. Presidente Rodríguez Zapatero s/n, 24492 Cubillos del Sil (León), Spain
e-mail: ma.delgado@ciuden.es

F. Del Corso
Air Liquide, 1 chemin de la Porte des Loges, BP 126, 78354 Jouy en Josas Cedex, France
e-mail: fabrice.delcorso@airliquide.com

M. Martín (ed.), *Alternative Energy Sources and Technologies*,
DOI 10.1007/978-3-319-28752-2_9

of worldwide CO_2 emissions could be recycled. Therefore, CO_2 uses are not an alternative but a complement to other CO_2 avoidance actions. As they are associated to an existing business model, they are enablers of CO_2 capture and transportation technologies at the industrial scale.

1.1 Thermodynamic Properties of CO_2

CO_2 can be solidified easily from liquid CO_2, as its triple point is at 5.28 bar. For example, any discharge of liquid CO_2 at atmospheric pressure through a valve will produce solid CO_2. The latent heat of solid CO_2 ("dry ice") is very high as shown in the chart below [encyclopedia airliquide].

Therefore, solid CO_2 can provide a lot of cold with a small volume of ice. Furthermore, solid CO_2 is at −78 °C, providing a cold source at a very low temperature and thermal transfer rates toward products to cool are very high (Table 1).

The latent heat of vaporization is about 2 times lower than the latent heat of sublimation but remains very high. Liquid CO_2 may be preferred to solid CO_2 as cooling media mainly because transportation and handling are easier (Fig. 1).

The critical point of CO_2 is 73.8 bar and 31.1 °C. Beyond this point, CO_2 is under supercritical state as shown in the diagram below.

The critical temperature of CO_2 is rather low. Therefore, supercritical CO_2 can be easily applied in various processes, especially extraction processes, to play a "solvent role." By increasing its specific gravity (for example by increasing its pressure), CO_2 will extract molecules more easily. In a following step where CO_2 specific gravity is decreased (for example by decreasing pressure) said molecules will be separated from their solvent. Therefore, supercritical CO_2 has a "tunable" solvent power.

- **Use of CO_2 as a cooling media**

Either liquid or solid CO_2 can be used as cooling media. The nontoxicity (CO_2 is not toxic; however, it has to be carefully handled especially in closed areas where high concentrations can lead to asphyxia) and stability of CO_2 makes it particularly interesting for open loop refrigeration processes but it has also a lot of advantages for closed loop refrigeration cycles.

Table 1 Latent heat for different substances

Compound	Latent heat (kcal/kg) at atm pressure
Water	540 (vaporization) 80 (solidification)
Nitrogen	47 (vaporization)
CO_2	137 (Sublimation)

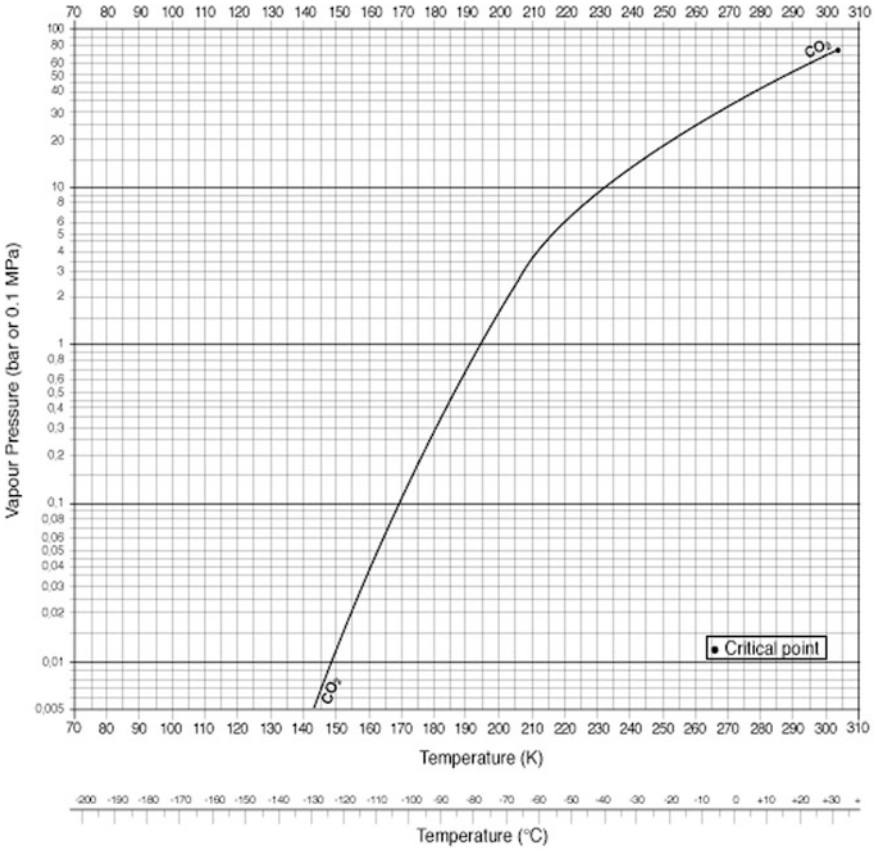


Fig. 1 Phase diagram of CO_2 (*Source* AIR LIQUIDE)

- **Use of CO_2 for closed loop refrigeration cycles**

Most of the refrigeration processes (like domestic of industrial fridges) use closed loop cycles driven by a mechanical energy, compressing and expanding a working fluid. CO_2 can be one of these fluids, see Fig. 2.

CO_2 was very much used in the first half of the twentieth century as refrigeration fluid. Then, the synthetic fluids (CFCs, HCFCs, HFCs,…) progressively replaced it. Since the 1990s, CO_2 is coming back strongly for this application as the synthetic fluids are getting progressively removed for environmental reasons (ozone layer protection but also global warming potential, GWP).

The most common HFCs, are R404a, which has a global warming potential (hereinafter, GWP) of 3260 kg CO_2 equivalent or R134a, with a GWP of 1300 kg CO_2 equivalent. Furthermore, CO_2 has no impact on the ozone layer depletion. With a GWP of 1 (vs. 1300 or 3260), using CO_2 instead of HFCs can guarantee the

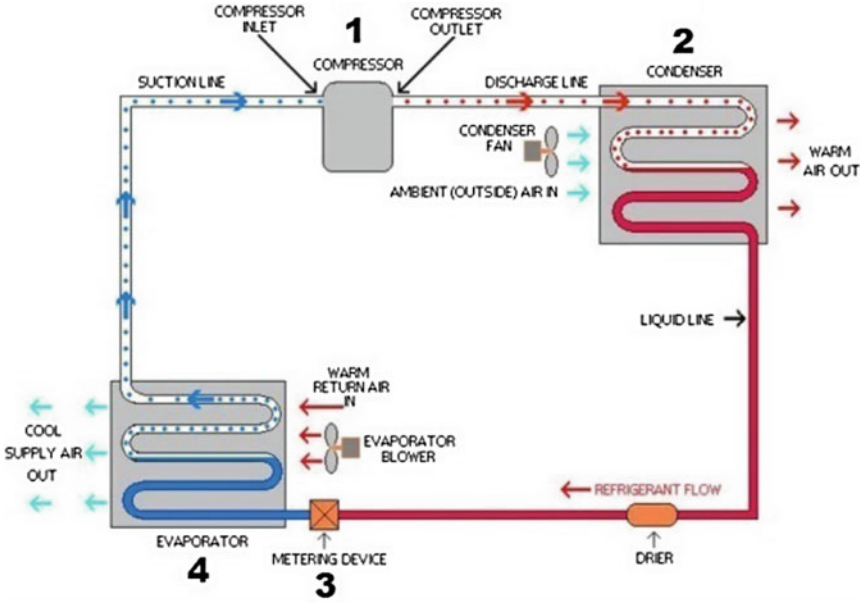


Fig. 2 Principle of a closed loop refrigeration cycle

refrigeration installations to comply with regulations on the long term and make leak management much easier, but it is not the only advantage of CO_2.

CO_2 is also an abundant fluid and heat produced (step 2 described in the figure above) by the refrigeration cycle with CO_2 can be valorized (for example for building heating of for food processing) because it can usually reach higher temperature (>60 °C) than with synthetic fluids cycles. In this case, the energetic efficiency of the whole system is greatly enhanced.

Furthermore, a refrigeration cycle with CO_2 can produce several level of cold with a single fluid.

- **Use of CO_2 with open loop refrigeration (direct contact)**

The closed loop refrigeration cycles can only transfer cold from the refrigerant to products to be cooled by indirect contact and convection heat transfer mode (see Fig. 3).

Another way to cool products with CO_2 is open loop refrigeration with direct contact of liquid/solid CO_2 with the products to be cooled. Technologically, for food freezing this operation can be done in refrigeration tunnels, see Fig. 4.

The direct contact cooling greatly enhances the heat transfer and products are frozen much faster than with indirect contact cooling. CO_2 is compatible with a lot of products to be cooled and especially food, see Fig. 5.

As a consequence, open refrigeration loops with CO_2 provide the following advantages:

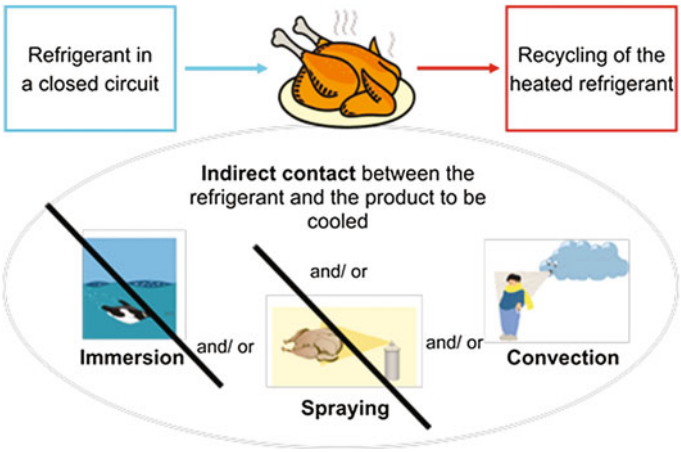


Fig. 3 Heat transfer mode for closed loop refrigerant cycles (*Source* AIR LIQUIDE)

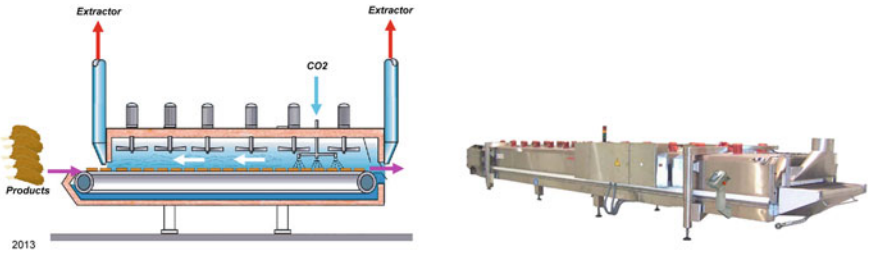


Fig. 4 CO_2 refrigeration tunnels (*Source* AIR LIQUIDE)

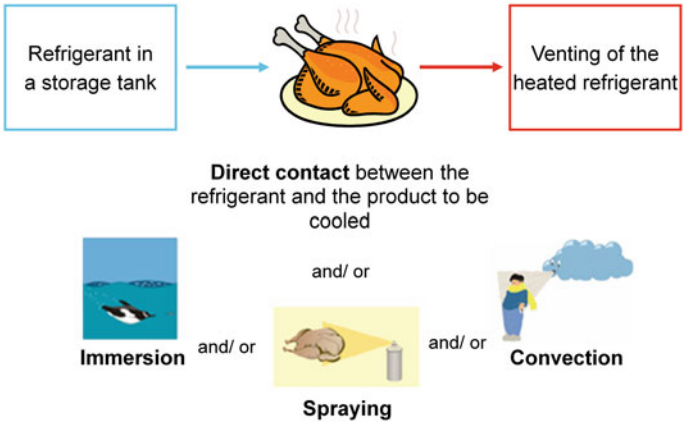


Fig. 5 Heat transfer mode for open loop refrigeration cycles (*Source* AIR LIQUIDE)

- Higher freezing speed is enhancing the product quality
- Higher flexibility
- Lower initial investment
- Quicker business start up
- Lower maintenance cost
- Lower floor space requirement
- Customized solutions
- Reduced water loss in the product

The main advantage of mechanical refrigeration versus open loop refrigeration cycles is the lower operating cost (cost of electricity vs. cost of CO_2).

Thanks to the advantages listed before, refrigeration with open loops using cryogenic liquids such as CO_2 has a typical 10–15 % share of food freezing market; the rest is covered by mechanical refrigeration units.

Food freezing is the most common application using CO_2 as a cooling media but there are other applications like cryogrinding (for example, for tires recycling).

NB: as CO_2 can replace oxygen, it can be asphyxiant at high concentration; this risk has to be carefully managed when using CO_2, in particular by ensuring an appropriate ventilation and oxygen control.

- **Use of CO_2 as a solvent**

Supercritical CO_2 offers many qualities and ecological advantages as a solvent. It is able to dissolve a wide range of chemical substances including organic substrates, catalysts, and light gases. It is a readily available molecule, cheap, recyclable, and nonflammable.

Under its supercritical state, carbon dioxide exhibits physical properties, which are intermediate between those of gases and liquids. These conditions are achievable using commercially available equipment. The main interest of supercritical fluids is related to their "tunable" properties, which can be changed easily by monitoring pressure and temperature: good solvent power at high densities (temperature near critical temperature and pressure much over critical pressure) to very low solvent power at low densities.

As a consequence, this solvent can be easily turned into a gas by simply releasing the pressure, leaving no solvent residues and requiring no energy for evaporation or separation.

As such it is widely used as an alternative to conventional solvents with numerous applications such as:

- extraction and purification of specialty chemicals and useful natural products,
- an alternative solvent in coatings industry, (replacing 40–90 % of volatile solvents), degreasing, and dry cleaning applications.
- it is also increasingly applied as a solvent in synthetic industrial processes.

With its low critical temperature, CO_2 is a key solvent to extract molecules sensitive to high temperature.

The great classics are extraction of caffeine from coffee, and also the extraction of essential oils for fine perfumery. In these cases, an alternative of supercritical CO_2 extraction would be distillation. This operation would require heating of the media containing the molecules and could therefore destroy the molecules and decreasing their aroma.

1.2 Solubility in Liquids and Lean Acid Properties of CO_2

CO_2 has higher solubility in a lot of liquids compared to other gases like nitrogen, oxygen, and air. CO_2 solubility in water, oil, or ethanol is 20–100 times larger, see Fig. 6.

CO_2 is a weak acid. When CO_2 dissolves in water, it forms carbonic acid which undergoes reactions typical of such an acid. The following equilibria are set up:

$$CO_{2\,aq} \rightleftarrows H_2CO_3 \rightleftarrows H^+ + HCO_3^- \rightleftarrows 2H^+ + CO_3^{2-}$$

At atmospheric pressure, the pH of a carbon dioxide solution is 3.7. As a dibasic acid, carbonic acid can react to form hydrogenocarbonates (like $NaHCO_3$) and carbonates (like Na_2CO_3).

- **Use of CO_2 for water environmental treatment**

For this use, CO_2 is playing an increasingly important and environment friendly role. Unlike other acids like sulfuric or chloridric acids, it does not leave any problematic salts like sulfates, chlorides, nitrates, etc. As a weak acid, it slightly reduces water pH (see figure below). This is an advantage as it allows a very precise

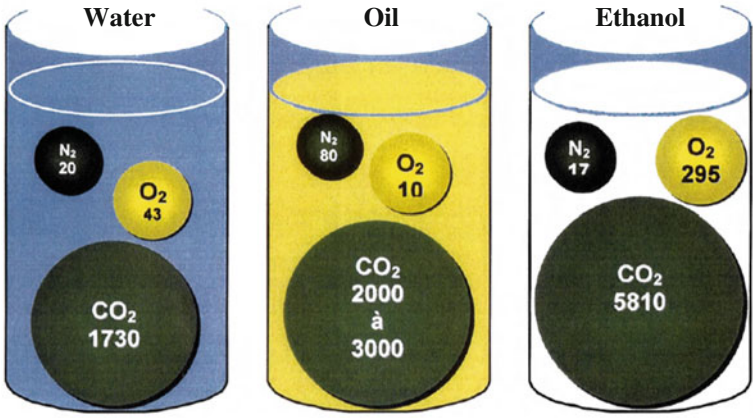


Fig. 6 Solubility of CO_2 in liquids compared to other gases (in mg/l at 20 °C and atmospheric pressure) (*Source* AIR LIQUIDE)

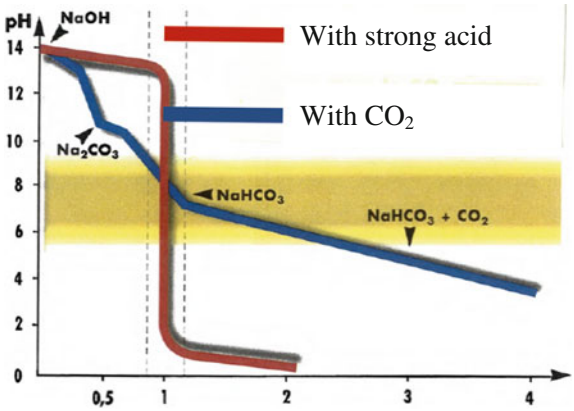


Fig. 7 Neutralization of basic solution via CO_2 compared to strong acid (*Source* AIR LIQUIDE)

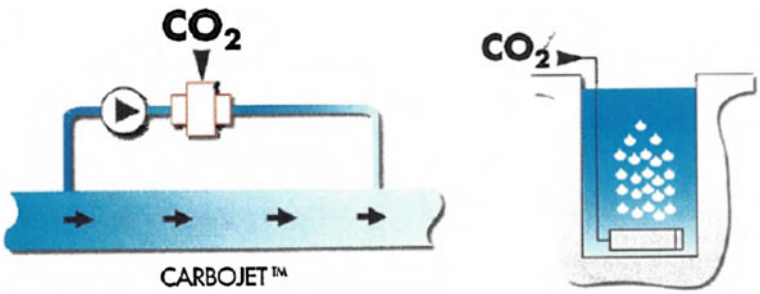


Fig. 8 CO_2 injection in water technologies (*Source* AIR LIQUIDE)

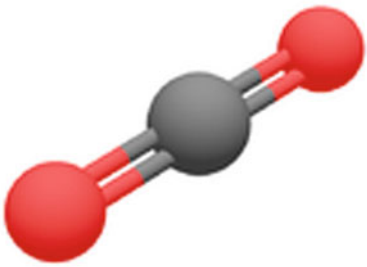
pH control and regulation reducing the risk to acidify too much the wastewater, see Fig. 7.

Its high solubility and easy diffusion in water allows a quick and homogenous mixing with liquids without the need of any mechanical agitators (see technological CO_2 injection in liquids technologies in Fig. 8).

1.3 Chemical Properties and Associated Uses of CO_2

Within the CO_2 molecule, the carbon atom is bounded with the two oxygen atoms via two double covalent links, see Fig. 9. These links are strong; therefore, the CO_2 molecule is very stable and depending on the targeted use which can be an advantage or a drawback.

Fig. 9 CO_2 molecule



Therefore, the reduction of CO_2 requires a lot of energy and requires high temperature, a catalyst, and a strong reducing molecule (like hydrogen).

The most common reduction is the reverse water–gas shift reaction and occurs by the reaction with hydrogen:

$$CO_2 + H_2 \rightleftarrows CO + H_2O$$

CO_2 can be also reduced to produce hydrocarbons.

On the other hand, thanks to its chemical stability, CO_2 can be used in contact of sensitive materials as a protective atmosphere. In particular, as CO_2 has a high compatibility with food and living tissues, it is used for food conservation but also preservation of organs for medical purposes.

- **Use for food conservation**

  Fresh food can be damaged by several mechanisms:

  – Physical: heat, cold, and water content
  – Enzymatic processes: enzymes are organic molecules made by development of microorganisms in food, this process can cause for example salad withering
  – Biochemical mechanisms: go rancid
  – Microbial degradation: by development of bacteria, algae, virus, mushrooms, mold, etc.

In many cases, an appropriate gaseous atmosphere can slow down the two last mechanisms (microbial and biochemical). In any case, this cannot replace the cold for food conservation. Ensuring the cold chain is still mandatory.

CO_2 is used for these gaseous atmospheres. Its reference as a food additive in Europe is E 290. The graph below shows the effect on an enriched CO_2 atmosphere on the bacteria development of ham at +3 °C.

So CO_2 is a neutral gas and it protects food from oxidation (from oxygen contained in air). On the other hand, at lower concentrations (5 %) CO_2 may enhance development of some microbes.

The formula of the protecting atmosphere has to be carefully designed. Indeed, for products containing a lot of water or grease, CO_2 will dissolve, thanks to its high solubility in these liquids, and may modify the tastes. This effect can be positive or negative for the taste (Fig. 10).

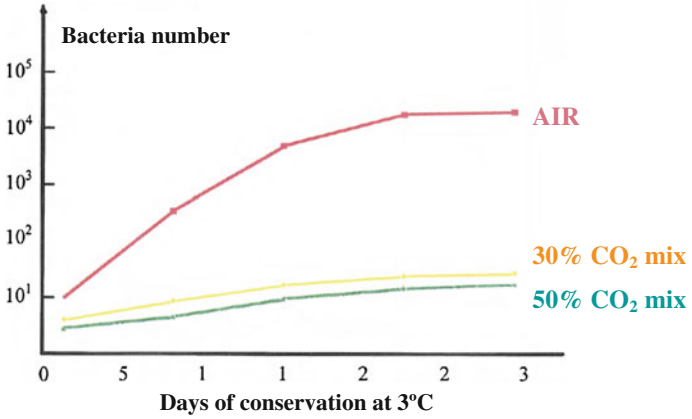


Fig. 10 Impact of CO_2 atmosphere on bacteria development (*Source* AIR LIQUIDE)

For beverages like beers or sodas, this effect is positive. Indeed, in these cases, a CO_2 atmosphere is playing its role of protection against oxygen and in addition, CO_2 allows to maintain the dissolved CO_2 concentration and to push it into delivery pipes.

- **Use for fire prevention: fire extinguisher**

CO_2 can be used for both fire prevention or as fire extinguisher, thanks to its high chemical stability.

Regarding other gases (like nitrogen), CO_2 is better suited for fire extinguisher than for fire prevention. Indeed, for fire prevention, CO_2 has to be in continuous contact with the products to protect. In this case, its high solubility in liquids like grease or water regarding nitrogen is a strong disadvantage as it will dissolve in most products and may change their properties, like for food (see above chapter).

On the contrary, CO_2 is often preferred as fire extinguisher. Liquid CO_2 has effects on 2 (heat and oxygen) out of 3 sides of the fire triangle.

1.4 CO_2 Oxidant Properties

Despite its high chemical stability, CO_2 can be reduced and therefore become an oxidant. Its main use in this case is within the natural carbon cycle, where it plays the key role of carbon provider to vegetables via the photosynthesis process.

By using energy of the sunlight and the accelerating influence of the chlorophyll, it reacts with water to form glucose.

$$6CO_2 + 6H_2O \Rightarrow C_6H_{12}O_6 + 6O_2 \quad (\Delta H = 2803\,\text{kJ})$$

Fig. 11 Greenhouse using enriched CO_2 atmosphere (*Source* AIR LIQUIDE)

- **Use for photosynthesis boosting in greenhouses and for algae production**

By increasing the partial pressure of CO_2, the thermodynamic equilibrium of the reaction $6CO_2 + 6H_2O \Rightarrow C_6H_{12}O_6 + 6O_2$ can be moved to increase glucose production. Enrichment of CO_2 concentration in greenhouses is today a fast-growing application. The concentration of CO_2 is typically increased from 350 to 800–2000 ppm. An increase in yields of 15 % is achievable (Hand 1982). A too high CO_2 concentration can have a toxic effect on the plants; therefore, it must be carefully controlled, see Fig. 11.

Photosynthetic conversion of CO_2 into biomass can be applied for cultures to make products that have energetic value, including biomass for gasification, methanation, and power.

Algae growth can be done in a bioreactor or raceway open ponds. The advantage of this process is that it uses directly solar energy and low purity CO_2. On the other hand, it requires large quantities of water and show low productivity (need of large bioreactors or ponds).

Lipid is extracted from algae, for example, mechanically and/or by a solvent. It is an intermediate product; additional steps like hydrogenation or esterification are required to transform this lipid into an additive for fuels.

A comparative cost analysis, (Comparative cost analysis of algal oil production for biofuels 2011) based on a preliminary audit of 12 sources for cost estimates,

calculated a production cost for a typical lipid (Triacylglyceride) at a mean value of 3.05 $/l.

Fuel production via algae is still at the research level today. Cost reduction is expected via algae selection and processes optimizations (LORNE 2015).

- **Chemical hydrocarbon production with recycled CO_2**

The chemical hydrocarbon production with recycled CO_2 typically starts by reducing CO_2 via a strong reductor like hydrogen (with catalyst) to produce a "C1" hydrocarbon; for example, methanol or methane:

$$CO_2 + 3H_2 \rightleftarrows CH_3OH + H_2O$$
$$CO_2 + 4H_2 \rightleftarrows CH_4 + 2\ H_2O$$

An alternative is to dissociate the CO_2 molecule by electrolysis or photoelectrolysis or a catalytic process ($CO_2 \rightleftarrows CO + 1/2\ O_2$).

To make sense and have a positive impact on CO_2 emissions, the energy for electrolysis or for the reductor production must be provided with renewable or decarbonized source, for example:

- electrolysis of CO_2 has to be performed with solar or wind mill electricity,
- if CO_2 reduction is made by H_2, it has to be produced via electrolysis of water fed by renewable electricity, etc.).

By these ways and associating other molecules, it is possible to produce final or intermediate chemical hydrocarbon products such as methane, methanol, polycarbonate, formic acid, diethyl ether, baking soda, etc.

These processes have the advantage to offer alternative routes to petrochemical products as well as substitutes, produced from renewable energy sources. However, these processes present challenges as they often require high-purity CO_2, (catalysts are intolerant to impurities) and the CO_2 needs to be combined with expensive products (e.g., hydrogen). They are not yet competitive compared to traditional fossil-based products. Therefore, most of them are at conceptual or lab-scale stage.

Methanol production from CO_2 is one of the more advanced processes on the industrialization path. It is currently at the demonstration stage. A typical process scheme for methanol production integrating renewable energy sources is shown below (Figs. 12 and 13):

The main industrial implementation today is on the site of Carbon Recycling International with a nominal production capacity of 5 million liters of methanol per year. The plant started in 2011.

Local favorable conditions increase the interest of this production: cheap electricity and cheap CO_2 provided by geothermic. Without such particular conditions, the methanol produced with recycled CO_2 is more expensive than the traditional production route with hydrocarbons. From the financial point of view, a minimum value of CO_2 of 170 €/t has to be reached to justify the development of a recycled CO_2-based methanol industry (ADEME 2014). For another chemical product like

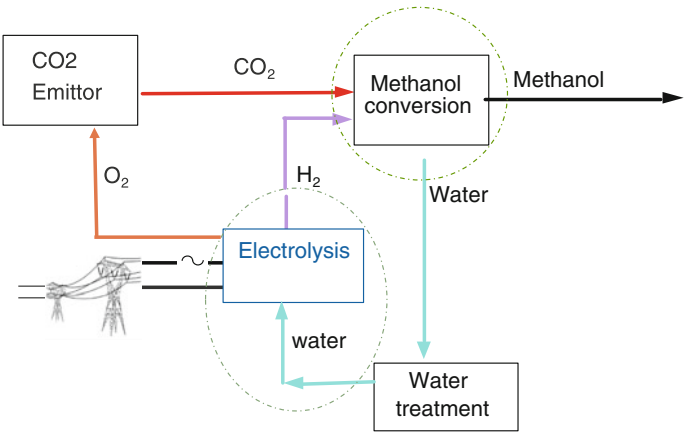


Fig. 12 Methanol production process (*Source* Project VItESSE2 French National research Agency)

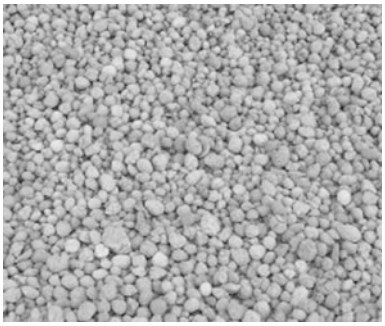


Fig. 13 Carbon negative aggregate (*Source* CARBON 8)

formic acid, the recycled CO_2-based product could be already competitive versus current hydrocarbon-based product (ADEME 2014).

- **Mineral carbonation**

Mineral carbonation consists of CO_2 absorption into a mineral-loaded alkaline brine, the resulting product is precipitated out as an inert mineral carbonate.

For example, the reaction can be: CO_2 + Oxide (MO) → Carbonate (MCO_3) + heat.

The mineral carbonates can be used in construction materials (e.g., aggregate equivalents and concrete). This process presents the advantage of valorizing industrial mining residues, replacing a portion of Portland cements in concrete and creating new, higher performing building materials.

Based on current knowledge, there seems to be no need for gases with high CO_2 concentration or a high level of gas purity. The challenges of mineral carbonation are its low chemical kinetics as well as new carbonate product uses and acceptance.

1.5 Use of CO_2 for Enhanced Oil Recovery (E.O.R.)

Enhanced Oil Recovery (E.O.R.) use is linked to the "Thermodynamic properties" of CO_2 and is qualitatively very close to the "solvent" role described in the former chapter. As written previously, the quantity of CO_2 recycled via the former uses remains very low versus the current CO_2 emissions. Indeed, the required technologies to produce and apply CO_2 for these uses are deployed at a small scale regarding the needs of CCS chain.

On the contrary, as it is consuming huge quantities of CO_2, EOR. is clearly a vector for diffusion of technologies at industrial scale for the CCS Chain.

- **What is EOR?**

After the exploration and drilling steps, oils fields' production can be divided into three phases:

– Primary recovery where the reservoir pressure is sufficient to ensure the production. Typically 30 % of oil in the reservoir can be extracted at this phase
– Secondary recovery where water or sometimes gas injection into the reservoir is applied to maintain its pressure and displace oil to the well
– Tertiary or enhanced oil recovery is used when water is not efficient enough to displace oil to the well and other fluids like gases, steam, or chemicals (polymers, detergent like surfactants,…) must be used, injected in the well. The physical action of theses fluids is to drive more oil to the well by mechanically pushing it and/or decreasing its viscosity or surface tension. Therefore, EOR processes are typically divided into three categories:

  - Thermal EOR (using mainly steam)
  - Gas EOR (using gas fluids and mainly CO_2 and nitrogen)
  - Chemical EOR

In 2013, the market shares were estimated at approximately 2,000,000 barrels per day for thermal EOR, 500,000 barrels per day for gas EOR, and 380,000 barrels per day for chemical EOR (Visiongain 2013).

- **CO_2 EOR**

Depending on the reservoir temperature, pressure, and oil composition, oil and CO_2 miscibility can vary a lot. A variable part of the CO_2 injected is going out of the well with the recovered oil. It is separated and reinjected into the well.

Indeed, CO_2 EOR has been optimized for oil production and not CO_2 Storage. However, CO_2 EOR can result in a very effective storage and nearly 100 % of the purchased CO_2 can be stored at the end (Godec 2011). CO_2 EOR began in the 1970s in the USA and has been developed first with natural CO_2 (from wells).

Considerable evolution has occurred in the implementation of this process. US CO_2 pipeline networks have been extended and anthropogenic CO_2 sources are accounting for increasing share of this CO_2 reaching 12 million metric ton/year supply in 2010.

USA is the country where CO_2 EOR is the most developed with about 80 % of oil production via this process in 2013. Other countries are developing it like Canada with the Weyburn Field where 23 million metric tons of CO_2 were planned to be injected, and also Oman, Mexico, and China.

2 CO_2 Capture

2.1 Introduction

The only way to limit global warming to less than 2° (i.e. 2DS or two degrees scenario) is to combine energy efficiency with renewable energy apart from a large expansion in the use of carbon capture and storage (CCS) (Christian 2015). CCS is recognized as one of the key existing technologies to reduce the global emissions of CO_2 into the atmosphere. Carbon dioxide is one of the major products of the combustion of fossil fuels such as coal, oil, or natural gas. Considering that coal-fired power plants are one of the most significant existing sources, these offer the single best target for applying carbon capture technologies to reduce global emissions (low number of plants and very important effect in reduction of CO_2 emissions).

Focused on technical aspects, although there are some important uncertainties to be solved, the technology has a green light; nevertheless, public acceptance and a stable regulatory framework are in a state of limbo being these issues mandatory to speed up the technology deployment. Apart from that, the main weakness of the technology is the high cost of CCS in a world without limits on carbon emissions.

The capital cost of the capture equipments, CO_2 pipelines, and storage reservoirs add extra costs in comparison with unabated gas and coal plants. In addition, capture process consumes about a quarter of the energy output of the power plant, imposing an efficiency penalty of 8 points according to the state of the art (e.g. Sankey diagram for oxycombustion, see Fig. 14). Together, over the lifetime of a gas- or coal-fired power plant, these costs would add at least 8 % to the cost of generating electricity.

Europe, which was expected to drive forward the technology with a series of early demonstration plants have failed to do so because of financial constraints within government and industry; at the time of writing, USA is taking the lead being the greatest need for CCS expected within developing nations such as China and India (Table 2).

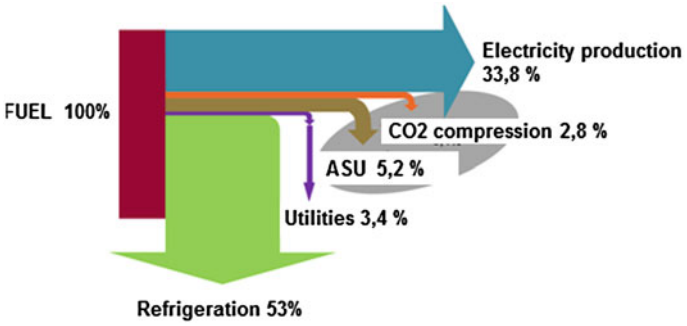


Fig. 14 Sankey diagram for oxycombustion

2.2 CO_2 Capture Technologies

Taking into account those cases in which the technology has achieved the demo or the commercial scale at the time of writing, there are three CO_2 capture technologies with a high potential (Carbon capture & storage association 2015a, b):

- Post-combustion capture: CO_2 can be captured from the exhaust of a combustion process by absorbing it in a suitable solvent. The absorbed CO_2 is liberated from the solvent and is compressed for transportation and storage. This process has been carried out industrially in Saskpower Boundary Dam. Post-combustion process offers the best method of retrofitting capture to existing plants; since the new installations are added at the end of the existing process, only the industrial land would be necessary.
- Precombustion capture: This technology is used in combination with gasification of coal. Gasification of fuel produces syngas which mainly contains CO and H_2. The carbon monoxide is reacted with water to form more H_2 and CO_2. Hydrogen is either used to power turbines and generate electricity or it is stored for later use. CO_2 is then captured from the gas stream and is ready for transport and storage.
- Oxy-combustion: In this process, coal is combusted with a mixture of oxygen and recirculated flue gases, instead of air. The flue gas, composed almost exclusively of CO_2 and water vapor, is easily separated. This process requires oxygen production. Although less-established than the other two strategies, oxy-combustion has made significant progress in the last decade, with a number of large pilots commissioned and three demonstration-scale projects in various stages of planning.

Finally, apart from the technologies previously explained, there are others emerging such as chemical looping combustion: oxygen carried by solid oxygen carriers reacts with fuel to produce a high concentration CO_2 stream in the flue gas, oxygen carriers are then regenerated to uptake oxygen from air in a second reactor, see Fig. 15.

Table 2 Large scale power plant CCS projects worldwide (MIT 2015; Marshall 2015)

ID	Name	Country	Continent	Thermal power (MWth)	Electric power (MWth)	Capture process	Status	CO_2 fate
1	Schwarze Pumpe	Germany	Europe	30	–	Oxy	Closed	Depleted natural gas field
2	CIUDEN	Spain	Europe	30	–	Oxy	In operation	Onshore Saline aquifer
3	Callide	Australia	Oceania	100	–	Oxy	In operation	Onshore saline aquifer
4	FutureGen 2.0	USA	America	–	229	Oxy	Cancelled	Onshore saline aquifer
5	White Rose	UK	Europe	–	426	Oxy	Planning	Offshore saline aquifer
6	Lacq	France	Europe	35	–	Oxy	In operation	Depleted hidrocarbon field
7	Mongstad	Norway	Europe	–	280[a]	Post	In operation	Saline aquifer
8	Peterhead	UK	Europe	–	385	Post	Planning	Offshore depleted gas reservoir
9	ROAD	The Netherlands	Europe	–	250	Post	Planning	Offshore depleted gas field
10	Porto Tolle	Italy	Europe	–	250	Post	Cancelled	Offshore saline formation
11	Getica	Romania	Europe	–	330	Post	Cancelled	Onshore saline formation
12	Belchatow	Poland	Europe	–	260	Post	Cancelled	Onshore Saline aquifer

(continued)

Table 2 (continued)

ID	Name	Country	Continent	Thermal power (MWth)	Electric power (MWth)	Capture process	Status	CO_2 fate
13	Don Valley	UK	Europe	–	650	Pre	Planning	Offshore saline formation
14	Kemper project	Mississippi	America	–	582	Pre	Planning	–
15	Boundary dam	Canada	America	–	139	Pre	In operation	EOR
16	Yanchang integrated CCS proyect	China	Asia	N/A	N/A	N/A	Planning	EOR

[a]Size for the power plant, not for the capture unit

Oxy Oxyfuel Combustion Capture; *Pre* Pre Combustion Capture; *Post* Post Combustion Capture; *EOR* Enhanced Oil Recovery

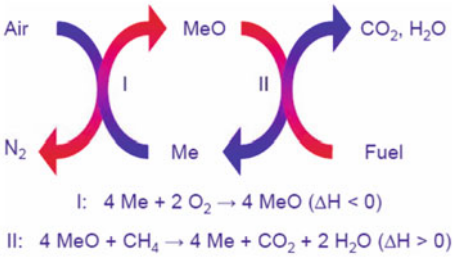


Fig. 15 Chemical looping general scheme

2.3 Postcombustion Capture

CO_2 removal by absorption and stripping with aqueous amine is a well-understood and widely used technology. The basic process, patented in 1930, is one in which CO_2 is absorbed from flue gas near ambient temperature into an aqueous solution of amine with low volatility, see Fig. 16. The amine is regenerated by stripping with steam (around 3 barg, according to the following figure), leaving pure CO_2 from the top of the stripping tower (Rochelle 2009).

In order to present the state-of-the-art about this process, the following projects are described:

- CO_2 Technology Centre Mongstad (hereinafter, TCM).
- SaskPower Boundary Dam: the world´s first commercial postcombustion CCS project.

CO_2 Technology Centre Mongstad

The amine plant was designed to be flexible in order to allow testing of different configurations, see Fig. 17. The major systems include (Hamborg et al. 2014):

- An induced draft (ID) blower to overcome pressure drops and blow the flue gas through the plant.
- A direct-contact cooler (DCC) system to initially quench and lower the temperature and saturate the incoming flue gas by a countercurrent flow water in order to improve the efficiency of the absorption process and provide prescrubbing on the flue gas.
- An absorber to remove CO_2 from the flue gas using solvent. The lower regions of the tower, where the amine solution contacts the flue gas, consist of three sections with structured packing whereas the upper region of the tower includes a water-wash system to scrub and clean the flue gas particularly of any solvent.
- Stripper columns (there are two although in the figure only one is represented) to recover the captured CO_2 and to recirculate the amine to the absorber column.

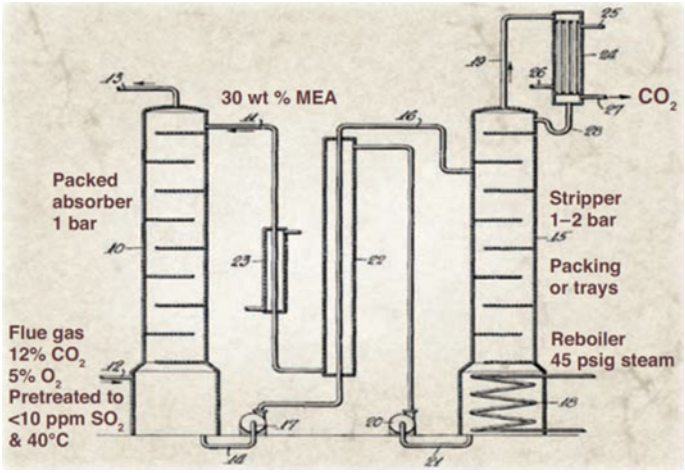


Fig. 16 The amine scrubbing process invented by Bottoms in 1930

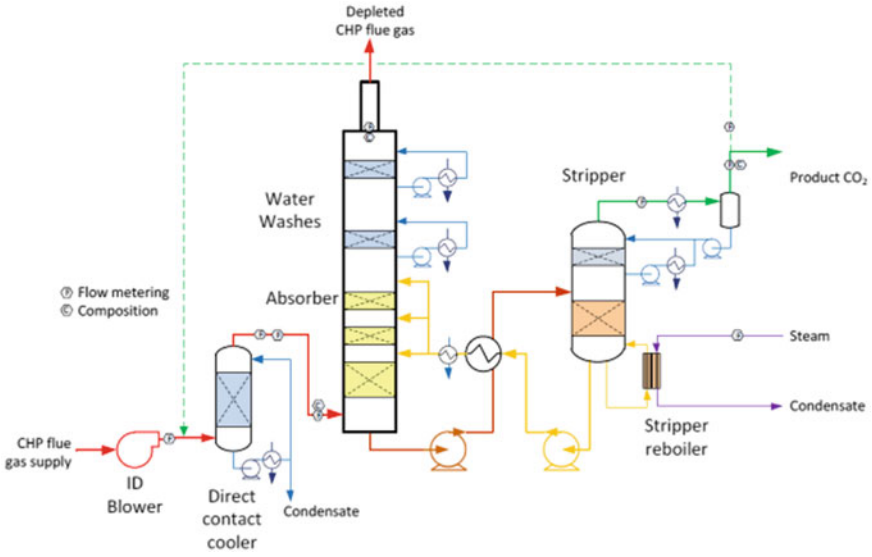
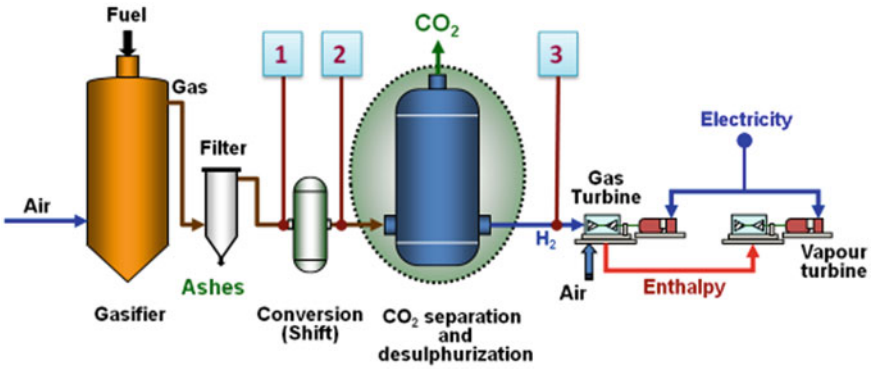


Fig. 17 Simplified flow schematic for TCM CO_2 capture

SaskPower Boundary Dam

The world´s first commercial scale post-combustion coal fired carbon capture and storage project was started in September 2014 at the SaskPower Boundary Dam Power Satation in Estevan (Stéphenne 2014).

At full capacity, the SaskPower Project captures over one million metric tons of CO_2 per year (5000 h/year basis), reflecting a 90 % CO_2 capture rate for the



% vol, aprox.	H_2
1	22
2	50
3	77

Fig. 18 Schematic process flow diagram in a IGCC plant

139 MW_{elec} coal-fired unit. The technology uses regenerable amines to capture both SO_2 and CO_2, which means that no direct waste by-products are generated.

2.4 Precombustion Capture

Integrated gasification combined cycle (hereinafter, IGCC) converts coal (or a solid fuel) into a gas fuel known as syngas containing hydrogen, carbon monoxide, and carbon dioxide at high temperature and pressure, see Fig. 18. IGCC uses a gas turbine where the syngas is burnt and a steam turbine to generate electricity. The main advantage of this option is that the CO_2 is relatively concentrated before separation and at a high pressure, so more efficient separation methods can be applied. Currently, solvent absorption is the method for removing the CO_2 from syngas; s,o liquid chemicals are used to absorb the CO_2 (see previous section).

2.5 Oxy-Combustion

The oxyfuel started in the early 1980s, when it was conceived as a way of producing pure CO_2 for enhanced oil recovery. Initially, oxycombustion was situated "behind" postcombustion technology which was well-established as an industrial process in natural gas purification. However, with growing interest in CCS in the

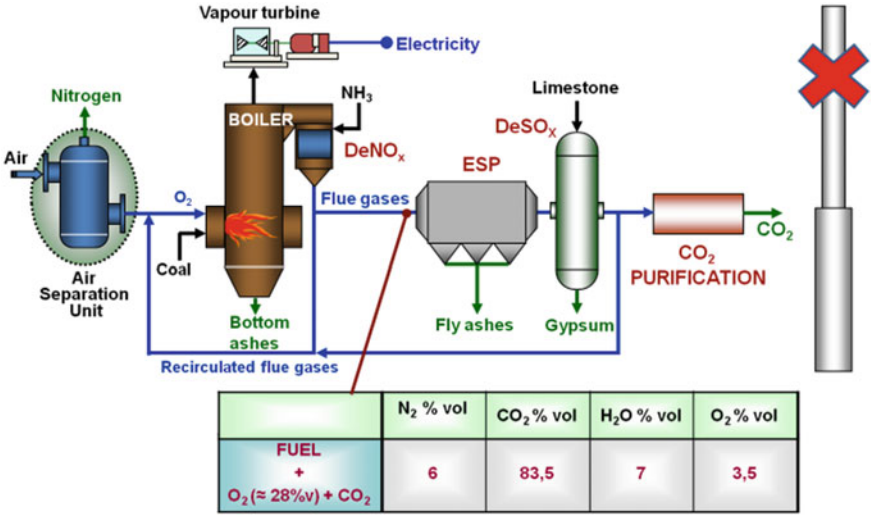


Fig. 19 Schematic process for a coal power plant in oxy-mode

1990s, early pilot-scale studies were performed, leading to increased interest from boiler manufacturers (Lockwood 2014).

Conventional coal power plants use air for combustion in which the nitrogen from the air (approximately 79 % by volume) dilutes the CO_2 concentration in the flue gases. During oxyfuel combustion, a combination of oxygen (typically of greater than 95 % purity) and recirculated flue gases are used for the combustion process. Consequently, a gas mainly conformed by CO_2 is obtained, being almost ready for sequestration. Worth mentioning is the fact that the recirculated flue gases are used to control the flame temperature and "make up" the volume of the missing N_2 to ensure there is enough gas to carry the heat through the boiler, see Fig. 19.

Oxyfuel combustion has been found to be different from air combustion in several ways, including reduced flame temperature, delayed flame ignition, reduced NOx and SOx emissions (represented as mg/MWe). In order to understand the reasons of these different effects, it is worthy to compare the gas properties between CO_2 and N_2 (main diluting gases in oxyfuel and air-mode, respectively) (Lockwood 2014), see Fig. 20.

The most relevant difference between the air and the CO_2 is the density (1.7 times higher for CO_2) and the molar heat capacity (1.6 times higher for CO_2) but also important differences are its behavior in the infrared spectrum and the low mass diffusivity for other gases such as oxygen (the oxygen diffusion ratio in CO_2 is 0.8 times that in N_2). CO_2 is also much more chemically active than the inert nitrogen and is able to participate in reactions such as the gasification for solid carbon (the Boudouard reaction). Once again, the ability to control the ratio of oxygen to recirculated flue gases provides a convenient additional degree of freedom with which to optimize the combustion process. However, the additional influence of this

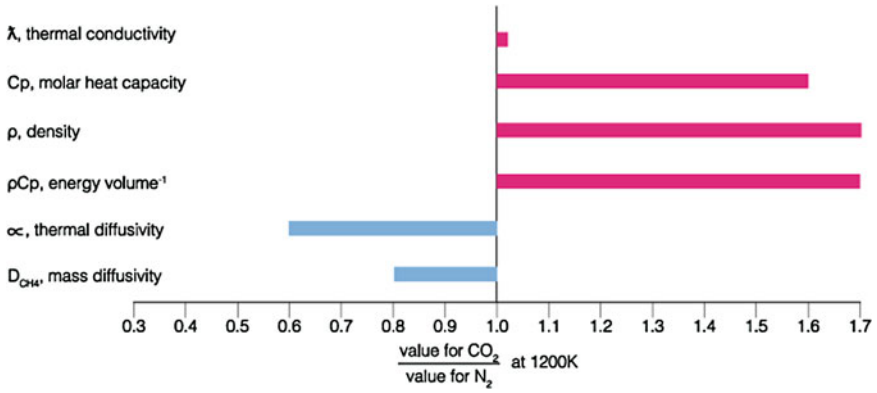


Fig. 20 Ratios of combustion relevant properties of CO_2 and N_2

parameter and the altered gas density on mass flows and burner aerodynamics must also be taken into account.

Focused on the CO_2 compression and purification unit (hereinafter, CPU) for flue gases produced during oxycombustion, a lot of schemes have been developed; since it is not possible to explain all of them, AIR LIQUIDE technology installed at CIUDEN site (Spain) has been selected due to the advantages that have become apparent (Lockwood et al. 2014):

- Cost-effective CO_2 recovery above 95 %.
- CO_2 purity greater than 99.99 %.
- Low water usage.
- Possible reduction in loss of ignition.

The AIR LIQUIDE's CPU includes following functional process blocks (Delgado et al. 2014), see Fig. 21:

- Flue gas quench and acid gas scrubbing (at close to atmospheric pressure).
- High-performance dust filter.
- Low-pressure flue gas drying system.
- Flue gas compression.
- Cryogenic CO_2 separation.

The following section describes the main CPU subunits:

- Flue gas quench and acid gas scrubbing: the flue gases, containing up to 84 % v d/b of CO_2, are passed into a quench where it is cooled to at least the saturation temperature by direct contact with sodium-based reagent. After this cooling stage, the gas will be scrubbed in a tower with the same solution in order to remove acidic compounds such as SOx, HCl, and HF.
- High performance dust filter: After the quench and scrubbing tower, the flue gas is compressed to around 1.4 barg in the blower C111 in order to overcome pressure drops. It is then filtered down to a very low level of particulate matter

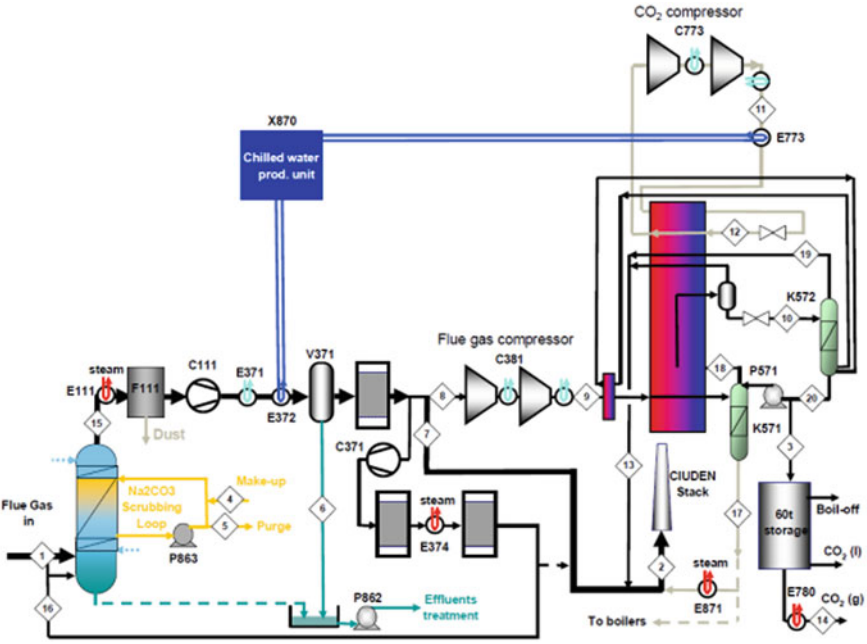


Fig. 21 Process flow diagram for a capture unit fed with flue gases from oxycombustion

load with a very high efficiency filtration system (F111). This very low dust load is required for reliable and efficient operation of the downstream process equipment.

- Low pressure flue gas drying system: a temperature swing adsortion (TSA) unit located downstream of the high-performance dust filtration system is responsible for drying the flue gas stream. The acid gas scrubbing upstream will protect the low pressure drying unit from potential problems caused by SOx. The first part of the drying system is a heat exchanger where water is condensed out by cooling first against cooling water and subsequently against chilled water. The second part of the drying system consist of three vessels (R371, R372, and R373) containing a specially selected adsorbent.
- Flue gas compressor: once the flue gas has been scrubbed, filtered, and dried, it is compressed to approximately 20 barg.
- Cryogenic CO_2 separation and CO_2-based refrigeration cycle: once the flue gas has been compressed, it is cooled to approximately −52 °C to liquefy the CO_2. After this cooling, two distillation towers will remove NOx and incondensable gases (K571 and K572, respectively); in the case of the first column, pure CO_2 from the product pressure vessel will be used to produce the liquid stream. The necessary cold for cooling the compressed flue gas to −52 °C will be provided by a refrigeration cycle using pure CO_2 as the working fluid.

3 CO_2 Transportation

When CO_2 is transported in large quantities (several million of metric tons per year), economies of scale make it economic to transport it over a few hundreds of km via pipelines. Therefore, CO_2 transportation by pipe is the major mean foreseen to implement a CCUS chain.

It is transported under its gaseous or supercritical state. CO_2 pipelines are similar to natural gas pipelines, generally made of steel. In its Lacq Demonstration Project, TOTAL S.A. reemployed an existing 27 km long natural gas pipeline to transport it at 30 bar prior its injection in a depleted natural gas reservoir (TOTAL 2007). Most of CO_2 impurities and especially water have to be removed to avoid pipelines corrosion.

Precise cost per km depends on a number of factors such as flow rate, location, terrain, composition operating pressure, rights of way, labour cost, etc. As a rule of thumb, 1 million $/km can be taken as order of magnitude of investment cost for several million tons of CO_2 per year pipeline onshore. For example, a 100-km pipeline depreciated on 20–30 years, 2–10 million tons/year of CO_2 transportation cost can be estimated between 1 and 6 $/ton of CO_2 (without the cost of initial compression) (IEA 2008).

Most of the existing CO_2 pipelines are located in the North America to supply CO_2 for EOR (see former chapter). A few thousand of km of pipelines are today under operation (IEA 2008).

Other European studies [COCATE project funded by European commission under FP7 energy (reference 241381)] showed the same order of magnitude for the estimation of the transportation cost results 6–8 €/ton, for 13 $MtCO_2/y$ transported on 700 km, project lifetime = 30 years. Cost of transportation by large ships is very close (less than 10 % more than pipelines transportation costs).

As investment costs and construction time for a pipelines network can be very high, transportation by ships is an interesting alternative to transport large quantities of CO_2. This is already done for smaller quantities for CO_2 industrial applications. In these cases, studies considered CO_2 will be transported under liquid form typically between −50 °C, 7 bar and −20 °C, 20 bar in large ships of several 10,000 tons of CO_2 capacities.

For smaller quantities (several hundreds or thousands of metric tons per year), truck or train transportation are the standard means to transport CO_2 via refrigerated or insulated tanks and liquid state typically at –20 °C, 20 bar.

4 CO_2 Storage

4.1 Introduction

Geological storage of CO_2 is one of the internationally accepted solutions to reduce CO_2 emissions into the atmosphere. It consists of injecting CO_2 to a depth of at

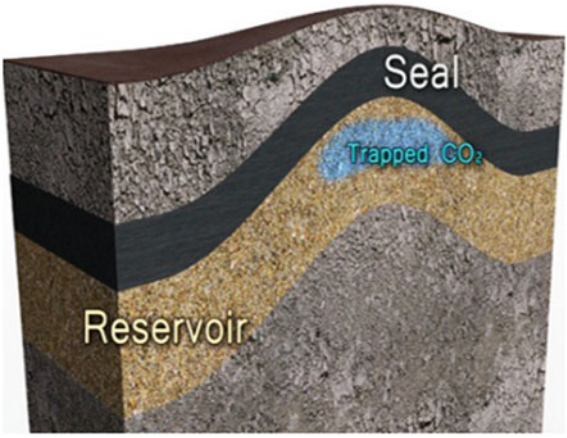


Fig. 22 CO_2 geological storage scheme

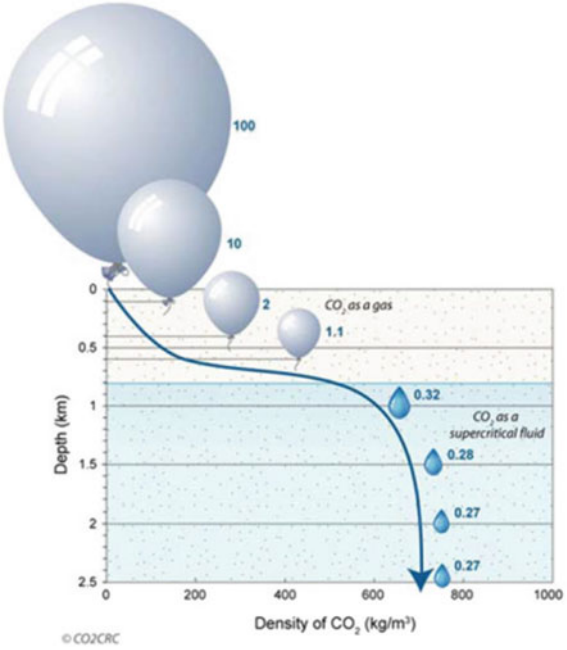


Fig. 23 CO_2 density as a function of the depth

least 800 m into suitable geological formations such as a highly porous and permeable rock, called reservoir rock, overlain by a cap rock that will stop the CO_2 from migrating to the surface (see Figs. 22 and 23).

The reason for storing carbon dioxide at such depths—close to 1000 m or deeper—is the behavior of CO_2 density with respect to pressure. We could say that CO_2 density increases sharply with the prevailing pressure at a depth of 800 m, allowing us to store

the same amount of CO_2 in a lower volume at depth in comparison with the volume that would be needed on the surface.

4.2 Trapping Mechanism

When CO_2 is injected in a geological reservoir, it can be physically trapped below a seal formation (structural trapping) or as residual gas relative because of permeability hysteresis. Geochemically speaking, CO_2 can be trapped through dissolution into the formation brine (solubility trapping), where it can react with the rock matrix and eventually precipitate into stable carbonate minerals (mineral trapping).

In terms of time, physical trapping occurs immediately after CO_2 is injected whereas residual trapping occurs because of relative permeability hysteresis. This process occurs after the injection operations stop as the CO_2 moves from the original injection point and the displaced brine imbibes back into the pore space previously occupied by CO_2 and traps a portion of the retreating CO_2. Because residual trapping occurs primarily after injection operation ceases, it does not play an important role in estimating storage efficiency or the amount of CO_2 that can be stored, see Fig. 24.

Solubility trapping occurs when injected CO_2 mixes with the formation waters. The amount of CO_2 that dissolves into the formation water is a function of temperature, water salinity, and pH, according to the following general rule: the higher

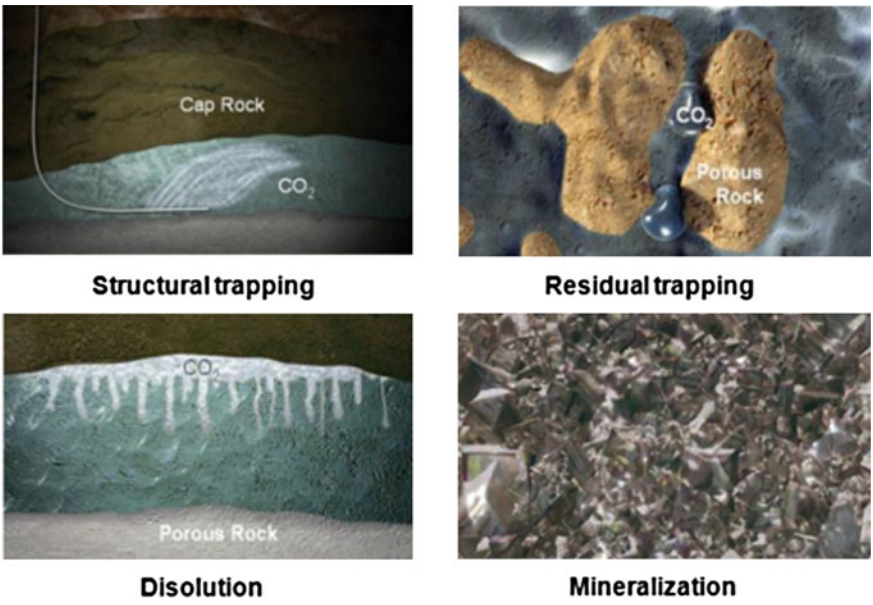


Fig. 24 Trapping mechanisms and their scales of time

the temperature or the salinity or the pH, the lower the CO_2 dissolution. Solubility trapping occurs immediately after injection begins and is dependent on the amount of mixing between the injected CO_2 and the formation waters.

The CO_2 dissolution in water will form H_2CO_3 that will be dissociated in bicarbonate (HCO^{3-}) and carbonate (CO_3^{2-}), decreasing the water pH in the reservoir. This pH modification will speed up the dissolution of specific minerals of the reservoir, producing the metallic cations liberation such as Mg, Ca, and Fe. These cations can react with bicarbonate anions to form carbonated minerals such as dolomite, calcite, and siderite; the last part of this process is known as mineral trapping.

4.3 Capacity Estimation for CO_2 Storage

Volumetric CO_2 storage resource is estimated using the pore volume of the storage target (a field, a portion or all of a saline formation, a geologic basin, etc.) and then multiplying the volume by an appropriate storage efficiency term (E) (see Eq. 1). The pore volume of an area is estimated by multiplying the porosity by the average high and total area. The efficiency term represents the fraction of the pore volume that CO_2 can occupy and is affected by boundary conditions, sweep efficiency, heterogeneity, etc. (Plataforma Tecnológica Española del CO_2 2014):

$$G_{CO_2} = A \cdot h_g \cdot \Phi_{tot} \cdot \ell \cdot E \tag{1}$$

where:

- G_{CO_2} = capacity of CO_2 in the reservoir (kg)
- A = normal are (m^2)
- h_g = averaged high the reservoir (m)
- Φ_{tot} = porosity (w/d)
- ℓ = CO_2 density in storage conditions (kg/m^3)
- E = efficiency factor

As the reader can see, volumetric estimations do not consider factors such as number of wells, injection flow rate variation with the time, and dependency between the injection pressures and the time, these factors being considered by the dynamic CO_2 storage resource.

4.4 Effects of Impurities

Considering that the capture process represents the most expensive element within the CCS chain, the purity has a huge importance on the actual costs of CCS. Consequently, the maximally allowable impurity level of the captured CO_2 during transport and storage has to be defined; nevertheless, at the time of writing various

knowledge gaps still exist with respect to the impact of CO_2 impurities on the CCS chain.

Physical effects (International Energy Agency 2019):

Physical effects are mainly determined by the presence and quantity of noncondensable impurities, such as Ar, O_2, N_2, and H_2 because they affect the fluid viscosity and density and, consequently, the injection process is altered (the injectivity is affected by a decreased density and a decreased mass flux over the same pressure drop).

On the other hand, the storage capacity is reduced because noncondensable impurities not only replace CO_2, but also reduce the density of the CO_2 stream and thus decrease the amount of CO_2 stored per unit volume as they are less compressible compared to pure CO_2.

In contrast to noncondensable impurities, easily condensable impurities, such as SO_2, increase both the storage capacity and injectivity through increasing the CO_2 stream density. It is noteworthy at this stage that, within a reservoir, decreased density also causes increased buoyancy of the CO_2 plume reducing the sweep efficiency of the injected CO_2 (decreasing any residual trapping) and decreasing the static storage capacity (the mass of CO_2 stored in the reservoir). And the larger the density difference between formation brine and injected CO_2, the larger the buoyancy. Accordingly, an amplified buoyancy would cause a larger lateral spreading of the CO_2 plume at the reservoir rock–cap rock interface due to the rising upward velocity next to reduced solubility as well as residual CO_2 trapping caused by reduced efficiency of CO_2 dissolution into formation brine.

Chemical effects:

Chemical effects are particularly important in the presence of water and oxygen and refer to changes in the geochemical behavior, i.e., reactivity and corrosion potential of the CO_2 stream toward reservoir and cap rock formations as well as technical installations (i.e., metallic parts and cements). Most significant species in that respect are SO_x, H_2S, and NO_x.

In case of coinjected SO_2, first the SO_2 will affect the partial pressure of CO_2 in supercritical state decreasing the solubility in the brine. Once dissolved into formation brine and oxidized it forms H_2SO_3 and/or H_2SO_4—sulfurous and/or sufuric acid, which are both stronger acids than carbonic acid (which forms from CO_2 dissolving into brine/water). Consequently, brine pH decreases resulting in higher mineral dissolution and enhancing mineral dissolution rates according to the following chemical reactions:

$$SO_2 + H_2O \rightleftarrows H_2SO_3$$

$$H_2SO_3 \rightleftarrows H^+ + HSO_3^- \rightleftarrows 2H^+ + SO_3^{-2}$$

$$CaCO_3(s) \rightleftarrows Ca^{+2} + CO_3^{-2}$$

$$Ca^{+2} + SO_3^{-2} \rightleftarrows CaSO_3(s)$$

In a SO_2–H_2S coinjection scenario, precipitation of elemental sulfur could be triggered and potentially block the pore space. These mineral precipitates may reduce the petrophysical rock properties such as porosity and permeability.

Just like SO_2, NO_2 is very soluble and coinjected NO_2 will promote dissolution of minerals by catalyzing the oxidation of SO_2 to sulfuric acid, on the one hand, but also by forming HNO_3–nitric acid with water molecules, on the other; two processes which again substantially lower the pH of the formation brine. In contrary to SOx, NOx will neither contribute to the formation of minerals and nor to a reduction of poroperm, respectively.

4.5 Monitoring

All areas of the CO_2 reservoir are kept under close survey at all times: the well, cap rock, and adjacent rock formations are monitored for changes in pressure and CO_2 concentration levels. This monitoring process takes place during all phases of a CO_2 reservoir's life, i.e., at the identification stage and the injection stage up to and after closure.

Monitoring systems include mainly thermal sensors to track temperature changes and seismic monitoring instruments such as tiltmeters to measure very small changes from the vertical level, either on the ground or in structures. Trackers and wired monitors are sent thousands of meters below the ground to check pressure and temperature changes near the reservoir (Zero Emission Platform 2014).

References

[Ademe 2014]. *Valorisation chimique du CO2*. ADEME. 2014. URL: http://www.ademe.fr/valorisation-chimique-co2-etat-lieux-quantification-benefices-energetiques-environnementaux-evaluation-economique-trois-voies-chimiques.

[Encyclopedia airliquide], AIR LIQUIDE, gas encyclopia, URL: http://encyclopedia.airliquide.com/encyclopedia.asp (2015).

[Hand 1982], D.W. Hand, Sci. Hortic. (Canterbury Engl.) **33** (1982) 14.

[Lorne 2015], D. Lorne URL: http://www.ifpenergiesnouvelles.fr/Espace-Decouverte/Tous-les-Zooms/Des-biocarburants-a-partir-de-microalgues (2015).

[Sun 2011], Sun & al, Comparative cost analysis of algal oil production for biofuels, Energy 36 (2011) p 5169–5179.

[Godec 2011], M.L. Godec, *Global Technology roadmap for CCS in Industry – Sectoral assessment – CO2 Enhanced Oil Recovery*, Advanced Resources International Inc. report for United Nations Industrial Development Organization, May 5, 2011.

[Visiongain 2013], *the Enhanced Oil recovery (EOR) Market 2013 – 2023 : Thermal, Gas and Chemical Production,* Visiongain report 2010.

[IEA 2008], *CO2 Capture and Storage, a key carbon abatement option*, p 81–85, 2008, IEA publications.

[TOTAL 2007] C&S conseil, *Project information dossier Lacq CO2 Capture and geological storage Pilot Project,* TOTAL S.A. Report, URL:http://www.total.com/sites/default/files/atoms/file/co2-lacq-total-project-information-dossier.

Carbon capture & storage association. (2015, 07). Retrieved from http://www.ccsassociation.org/what-is-ccs/capture/post-combustion-capture/.

Carbon capture & storage association. (2015, 07). Retrieved from http://www.ccsassociation.org/what-is-ccs/capture/pre-combustion-capture/.

Christian, B. (2015, 01). *EuroActiv*. Retrieved from Capture the carbon: http://www.euractiv.com/sections/energy/capture-carbon-311149.

Delgado, M. A., Diego, R., & Alvarez, I. (2014). CO2 balance in a compression and purification unit (CPU). *Energy Procedia*, 322–331.

Hamborg, E., Smith, V., & Cents, T. (2014). Results from MEA testing at the CO2 Technology Centre Mongstad. Part II: Verification of baseline results. *Energy Procedia*, 5994–6011.

International Energy Agency. (2019). *CO2 storage efficiency in deep saline formations: a comparison of volumetric and dynamic storage resource estimation methods.* IEA Environmental Projects.

Lockwood, F., Granados, L., & Leclerc, M. (2014). Oxy-combustion CPU - From pilots towards industrial-scale demonstration. *Energy Procedia*, 342–351.

Lockwood, T. (2014). *Developments in oxyfuel combustion of coal.* IEA Clean Coal Centre.

Marshall, C. (2015, 02). *E&E publishing*. Retrieved from DOE kills FutureGen project: http://www.eenews.net/stories/1060012843.

MIT. (2015, 07). *Carbon capture & sequestration technologies @ MIT*. Retrieved from https://sequestration.mit.edu/.

Plataforma Tecnológica Española del CO2. (2014). Almacenamiento de CO2: tecnologías, oportunidades y expectativas.

Rochelle, G. (2009). Amine Scrubbing for CO2 Capture. *Carbon Capture and Sequestration*, 1652–1654.

Stéphenne, K. (2014). Start up of world´s first commercial post-combustion coal fired CCS project: contribution of Shell Cansolv to SaskPower Boundary Dam ICCS Project. *Energy Procedia*, 6106–6110.

Zero Emission Platform. (2014, 06). *Zero Emission Platform*. Retrieved from http://www.zeroemissionsplatform.eu/ccs-technology/storage.html.

Optimal Design of Macroscopic Water and Energy Networks

Ramón González-Bravo, Fabricio Nápoles-Rivera and José María Ponce-Ortega

Abstract Water scarcity has led to an increase in the extraction of fresh water from aquifers, dams and lakes in certain regions where water availability is low. It has created serious problems in overexploitation of ground and surface water resources. This issue has been intensified due to population growth and increases in energy and water demands in the industry, agriculture, and households. In this chapter, a mathematical model for energy and water distribution networks in a macroscopic system is proposed. This model considers that the water and electricity demands can be satisfied by the existing power plants in the region and the installation of new power-desalination plants. Also, the model considers that the water demand can be satisfied by supplying water from dams, rivers, and aquifers. The model considers a macroscopic system that involves several cities in a water-stressed region. It accounts for variations in water demands throughout the year, for domestic, agricultural, and industrial users. The model considers both installation costs and operating costs of the new power-desalination plants, the installation of new storage tanks, pumping, and piping costs. The results show attractive solutions, where interesting economic profits can be obtained as well as the potential recharge of aquifers can be achieved.

List of Symbols

Indexes

g Location of agricultural users
i Existing aquifer
j Deep wells

R. González-Bravo · F. Nápoles-Rivera · J.M. Ponce-Ortega (✉)
Chemical Engineering Department, Universidad Michoacana de San Nicolás de Hidalgo, 58060 Morelia, Michoacán, Mexico
e-mail: jmponce@umich.mx

R. González-Bravo
e-mail: r.glezb@yahoo.com

F. Nápoles-Rivera
e-mail: fnapoles@umich.mx

M. Martín (ed.), *Alternative Energy Sources and Technologies*,
DOI 10.1007/978-3-319-28752-2_10

n Location of existing power-desalination plant (p-d plant)
o Location of industrial users
p Location of existing water storage tanks
q Possible location of new water storage tanks
r Location of domestic users
t Distribution time in months
u Possible location of new p-d plant
x Existing dam as natural resources

Positive variable

$a_{i,j,t}^{\mathrm{dw}}$	Water sent to the deep wells from aquifers, m^3/y
Annual Profit	Annual profit from water sales, $
$b_{n,t}^{E,\mathrm{rej}}$	Water sent to the sea as reject from existing p-d plants, m^3/y
$b_{u,t}^{N,\mathrm{rej}}$	Water sent to the sea as reject from new p-d plants, m^3/y
$B_{n,i,t}^{E,\mathrm{des}}$	Water received from existing p-d plants, m^3/y
$B_{u,i,t}^{N,\mathrm{des}}$	Water received from new p-d plants, m^3/y
$d_{r,j,t}^{\mathrm{dom}}$	Water distributed in central stations (c-s) from deep wells (domestic), m^3/y
$d_{g,j,t}^{\mathrm{agr}}$	Water distributed in c-s from deep wells (agricultural), m^3/y
$d_{o,j,t}^{\mathrm{ind}}$	Water distributed in c-s from deep wells (industrial), m^3/y
$D_{n,p,t}^{E,\mathrm{Esto}}$	Water received in storage tanks from existing p-d plant, m^3/y
$D_{u,p,t}^{N,\mathrm{Esto}}$	Water received in storage tanks from new p-d plants, m^3/y
$E_{r,t}^{\mathrm{dom}}$	Energy demand by domestic users, kW
$E_{g,t}^{\mathrm{agr}}$	Energy demand by agricultural users, kW
$E_{o,t}^{\mathrm{ind}}$	Energy demand by industrial users, kW
Energy Sales	Energy sales from p-d plants, $
$\mathrm{EProduction}_{n,t}^{E}$	Energy produced in existing p-d plants, kW
EPDopcost	Operating cost for existing p-d plants, $
$\mathrm{EProduction}_{u,t}^{N}$	Energy produced in new p-d plants, kW
$F_{i,t}^{\mathrm{agr}}$	Aquifers recharge from agricultural users, m^3/y
$G_{n,x,t}^{E,\mathrm{rel}}$	Water received in dams from existing p-d plants, m^3/y
$G_{u,x,t}^{N,\mathrm{rel}}$	Water received in dams from new p-d plants, m^3/y
$h_{r,t}^{\mathrm{dom}}$	Water in domestic c-s, m^3/y
$h_{g,t}^{\mathrm{agr}}$	Water in agricultural c-s, m^3/y
$h_{o,t}^{\mathrm{ind}}$	Water in industrial c-s, m^3/y
$\mathrm{InstCost}_{q}^{\mathrm{sto}}$	Storage installation cost, $
$\mathrm{InstCost}_{u}^{\mathrm{pdes}}$	Installation cost (new p-d plants), $
$M_{x,t}$	Water in dam, m^3/y
NPDinstcost	Installation cost for new p-d plants, $

NPDopcost	Operational cost for new p-d plants, $
$\mathrm{OpCost}_{u,t}^{\mathrm{pdes}}$	Operational cost (new p-d plants), $
$\mathrm{OpCost}_{n,t}^{E,\mathrm{des}}$	Operational cost (existing p-d plants), $
Piping Cost	Piping cost, $
Pumping Cost	Pumping cost, $
R	Natural recharge of water, m^3/y
$S_{p,t}^{E}$	Total water in existing storage tanks, m^3/y
$S_{q,t}^{N}$	Total water in new storage tanks, m^3/y
$s_{p,i,t}^{E}$	Water received from existing storage tanks, m^3/y
$s_{g,q,t}^{N}$	Water received from new storage tanks, m^3/y
$S_{q}^{\max}$	Maximum storage value, m^3
Storage Cost	Storage cost, $
$\mathrm{SW}_{n,t}^{\mathrm{in},E}$	The total water extracted from the sea for existing p-d plants, m^3/y
$\mathrm{SW}_{u,t}^{\mathrm{in},N}$	The total water extracted from the sea for new p-d plants, m^3/y
TAC	Total annual cost, $
$\mathrm{TEnergy}_t$	Total energy produced, kW
$v_{r,n,t}^{E}$	Water received from existing p-d plants, m^3/y
$v_{r,u,t}^{N}$	Water received from new p-d plants, m^3/y
Water Sales	Water sales from p-d plants, $
$W_{i,t}^{\mathrm{aq}}$	Water in aquifer, m^3/y
$w_{r,x,t}^{\mathrm{dom}}$	Water to c-s from dam (domestic), m^3/y
$w_{g,x,t}^{\mathrm{agr}}$	Water to c-s from dam (agricultural), m^3/y
$w_{o,x,t}^{\mathrm{ind}}$	Water to c-s from dam (industrial), m^3/y

Parameters

$a^{\mathrm{dw,max}}$	Maximum capacity of pipeline to deep wells from aquifers
AEC_t	Agricultural energy cost, $/kW
APC	Agricultural piping cost, $/m^3$
AQP	Deep well piping cost, $/m^3$
$\mathrm{agrdem}_{g,t}$	Water requirements of agricultural users, m^3/y
$B^{E,\mathrm{des,max}}$	Maximum capacity of pipeline to aquifers from existing p-d plants
$B^{N,\mathrm{des,max}}$	Maximum capacity of pipeline to aquifers from new p-d plants
BPC	Aquifer recharge piping cost, $/m^3$
$d^{\max}$	Maximum capacity of pipeline from deep well
$D^{E,\mathrm{sto,max}}$	Maximum capacity of pipeline from existing p-d plants (storage tanks)
$D^{N,\mathrm{sto,max}}$	Maximum capacity for pipeline from new p-d plants (storage tanks)
DEC_t	Unit domestic energy cost, $/kW
DPC	Unit domestic piping cost, $/m^3$
$\mathrm{domdem}_{r,t}$	Water requirements of domestic users, m^3/y
EPC	Storage recharge piping cost (existing p-d plant), $/m^3$

$G^{E,\mathrm{rel,max}}$	Maximum capacity for pipeline from existing p-d plants (dams)
$G^{N,\mathrm{rel,max}}$	Maximum capacity for pipeline from new p-d plants (dams)
GEP	Energy generation factor, kW/m^3
GPC	Unit dam recharge piping cost, \$/m^3
$h^{\max}$	Maximum capacity in distribution station
H_Y	Operating hours for the plant per year, h/y
IEC_t	Unit industrial energy cost, \$/kW
IPC	Unit industrial piping cost, \$/m^3
$\mathrm{inddem}_{o,t}$	Water requirements of industrial users, m^3/y
k_F	Factor used to annualize the inversion, y^{-1}
NPC	Unit storage recharge piping cost (new p-d plant), \$/m^3
pca	Percentage of water that seeps into the ground
PPA	Unit agricultural pumping cost, \$/m^3
PAQ	Unit deep well pumping cost, \$/m^3
PPC	Unit aquifer recharge pumping cost, \$/m^3
PPD	Unit domestic pumping cost, \$/m^3
PPE	Unit storage recharge piping cost (existing p-d plant), \$/m^3
PPG	Unit dam recharge pumping cost, \$/m^3
PPI	Unit industrial pumping cost, \$/m^3
PPN	Unit storage recharge piping cost (new p-d plants), \$/m^3
PPS	Unit storage pumping cost, \$/m^3
$s^{E,\max}$	Maximum capacity for pipeline from existing storage tanks
$s^{N,\max}$	Maximum capacity for pipeline from new storage tanks
SPC	Unit storage piping cost, \$/m^3
$\mathrm{SW}_u^{\max}$	Maximum value of seawater extracted from the sea, m^3/y
$v^{E,\max}$	Maximum capacity for pipeline to c-s from existing p-d
$v^{N,\max}$	Maximum capacity for pipeline to c-s from new p-d
$w^{\max}$	Maximum capacity for pipeline to c-s from dams
WAC_t	Unit cost of water for agricultural use, \$/m^3
WDC_t	Unit cost of water for domestic use, \$/m^3
WIC_t	Unit cost of water for industrial use, \$/m^3
Z	Unit fixed cost, \$
α	Economic scale factor
β	Reject factor of power-desalination plants
Θ_q^{sto}	Limit capacity of the tank, m^3
Θ_u^{pdes}	Limit capacity of new power-desalination plants, m^3/y

Binary Variables

y_q^{sto}	Existence of storage tanks
y_u^{pdes}	Existence of p-d plants
y	Existence of pipeline from c-s to domestic user

1 Introduction

Water and energy are closely linked. Energy production uses large amounts of water and providing freshwater requires energy for pumping, treating, and transporting water to the consumers (DOE 2014). In most of the power production plants, water is used as cooling utility (hydroelectric, nuclear, and refining); giving us an opportunity to couple desalination systems to recover waste heat to produce water and electricity simultaneously (UN-Water 2014). Coupling power production and water desalination is an option to provide water in regions with high water stress. However, water management and energy distribution represent the major challenges in power-desalination schemes (Averyt et al. 2011).

Finding optimal solutions to satisfy water demands has been studied by several authors. Nápoles-Rivera et al. (2013) proposed a mathematical programing approach for the sustainable water management (distribution and storage) in a macroscopic system, in which the authors propose that the water can be obtained by superficial water bodies, groundwater, rainwater, and treated wastewater. Later, they extended this approach to include seasonal increasing demands, population growth, and uncertainty in the rainfall patterns (Nápoles-Rivera et al. 2015). Zheng and Zecchin (2014) introduced a multi-objective optimization method for designing water distribution systems, in this work the minimization of the entire network cost is subjected to the selection of pipe diameters and the maximization of the minimum pressure across the network. Georgescu and Georgescu (2014) presented an optimization framework for scheduling pumping stations for water distribution networks in the city of Oradea in Romania, the results showed economic benefits due to the reduction of the energy consumption by the pumps. Other authors have focused on the optimization of power-desalination schemes, where energy, economic, and environmental effects were studied. Najafi (2014) proposed a dual-purpose plant for power generation and desalination using multistage flash desalination. Agashichev and El-Nashar (2005) proposed a triple hybrid model combining power production, multi-stage flash desalination and reverse osmosis, this work presents an economic and ecological analysis for various technological parameters and economic assumptions. Almansoori and Saif (2014) presented a model coupling reverse osmosis and power generation to find the optimal arrangement as well as the optimal operating conditions. Esfahani and Yoo (2014) proposed a link between power generation and multi-effect thermal vapor compression to minimize the total annual desalination cost. On the other side, several efforts have been made in the field of water management involving desalination and power production. Liu et al. (2011) presented an MILP optimization model integrating production, distribution, and storage of desalinated seawater, treated wastewater and reclaimed water, in the optimization task the authors subdivided the regions to propose an optimal water network. Atilhan et al. (2011) presented a systems integration approach for designing and planning decisions involving desalination technologies for water distribution networks in a macroscopic scale, this model determines the type of technology as well as the location and capacity of the desalination plant. Then, the

same authors extended this approach to incorporate fluctuations in water demands and the availability of energy resources for desalination issues (Atilhan et al. 2012).

In this chapter, the proposed model incorporates traditional (dams, wells, river, etc.) and alternative (desalinated seawater, harvested rainwater, and reclaimed water) water sources to satisfy the water demands. Also, the model incorporates the possibility of installing seawater desalination plants integrated with power plants. This way, the energy required in the desalination plant is provided by the heat excess from the power plant, then the associated cost is absorbed by the sale of the electricity produced in the power plant. Furthermore, to improve the sustainability of the system, the proposed optimization formulation considers the recharge of overexploited aquifers. In addition, the proposed model accounts for the variation of the water demands through the year. Important challenges such as the optimal location for installing new units, and the scheduling and operation of the system are considered to maximize the overall benefit. The design problem consists in finding the optimal water distribution network, which satisfies the electricity and fresh water demands in a macroscopic system, considering the seasonal water and electricity demands, pumping and piping costs, the optimal capacity and location of the new integrated desalination power plants, and the size and location of new storage tanks. To solve the above-mentioned problem, a superstructure and its corresponding optimization formulation is proposed, the objective function consists in maximizing the overall profit for the sales of electricity and water satisfying the corresponding demands. Next section describes the proposed optimization formulation.

2 Model Design

The optimization model is based on the superstructure shown in Fig. 1. This model accounts for the energy distribution and the water management in a macroscopic system. The proposed model seeks to satisfy water and energy demands by installing a new power-desalination plant. The indexes are defined as follows: i represents the existing aquifer, j defines the existing deep wells in each region, t is the time periods (i.e., months), p is the location of existing water storage tanks, q is the possible location of new water storage tanks, n is the location of existing power-desalination plant, u is the possible location of new power-desalination plants, r represents the region of domestic users, g is a region of agricultural users, o is the region of industrial users and x is the existing dam as natural resources.

2.1 Equations for Existing Power Plants and New Power-Desalination Plants

The power-desalination plants are modeled as black boxes, where the power production of an existing power plant and the power/desalination production factor of

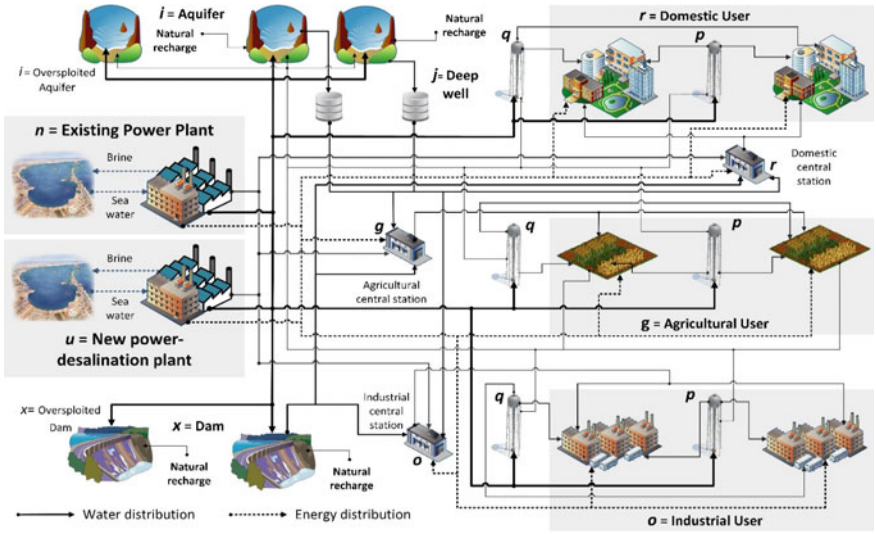


Fig. 1 Proposed superstructure for electricity and water distribution

new power-desalination plants are known. The energy produced in new power-desalination plants is a function of the total feed water.

$$\text{EProduction}_{u,t}^{N} = \text{SW}_{u,t}^{\text{in},N} \cdot \text{GEP}, \quad \forall u \in U, \ \forall t \in T \tag{1}$$

This function has a linear behavior that depends on the capacity of the new power-desalination plant ($\text{SW}_{u,t}^{\text{in},N}$) multiplying by a factor (GEP).

In case of new power-desalination plants, the total water rejected from the desalination technology can be calculated multiplying the total seawater times a factor (β), this factor indicates the conversion from seawater to fresh water, which is defined as the flux sent to the sea as reject. This is modeled as follows for new power-desalination plants:

$$b_{u,t}^{N,\text{rej}} = \beta * \text{SW}_{u,t}^{\text{in},N} \quad \forall u \in U, \ \forall t \in T \tag{2}$$

Then, the total energy produced (TEnergy_t) is the sum of the energy generated in each of the new ($\text{EProduction}_{u,t}^{N}$) and existing ($\text{EProduction}_{n,t}^{E}$) power-desalination plants; notice that this value depends on the time and this value varies in each time period due to the energy demand.

$$\text{TEnergy}_t = \sum_{u} \text{EProduction}_{u,t}^{N} + \sum_{n} \text{EProduction}_{n,t}^{E}, \quad \forall t \in T \tag{3}$$

The total produced energy (TEnergy_t) is a function of the energy demand in industrial zones ($E^{\text{ind}}_{o,t}$), domestic users ($E^{\text{dom}}_{r,t}$) and agricultural regions ($E^{\text{agr}}_{g,t}$).

$$\text{TEnergy}_t = \sum_r E^{\text{dom}}_{r,t} + \sum_g E^{\text{agr}}_{g,t} + \sum_o E^{\text{ind}}_{o,t} \tag{4}$$

On the other hand, the total water extracted from the sea for existing power-desalination plants ($\text{SW}^{\text{in},E}_{n,t}$) is equal to the water sent to the existing storage tanks ($D^{E,\text{Esto}}_{n,p,t}$), plus the water sent to new storage tanks ($D^{E,\text{Nsto}}_{n,q,t}$), the water sent to distribution stations for: domestic ($v^{E,\text{dom}}_{r,n,t}$), industrial ($v^{E,\text{ind}}_{o,n,t}$) and agricultural ($v^{E,\text{agr}}_{g,n,t}$) uses, also the water sent to recharge aquifers ($B^{E,\text{des}}_{n,i,t}$), the water sent to recharge dams ($G^{E,\text{rel}}_{n,x,t}$) and the water sent to the sea as reject ($b^{E,\text{rej}}_{n,t}$).

$$\begin{aligned}\text{SW}^{\text{in},E}_{n,t}(1-\beta) = {} & \sum_p D^{E,\text{Esto}}_{n,p,t} + \sum_q D^{E,\text{Nsto}}_{n,q,t} \quad \forall n \in N,\ \forall t \in T \\ & + \sum_r v^{E,\text{dom}}_{r,n,t} + \sum_g v^{E,\text{agr}}_{g,n,t} + \sum_o v^{E,\text{ind}}_{o,n,t} \\ & + \sum_i B^{E,\text{des}}_{n,i,t} + \sum_x G^{E,\text{rel}}_{n,x,t} + b^{E,\text{rej}}_{n,t}\end{aligned} \tag{5}$$

The total water extracted from the sea for new power-desalination plants ($\text{SW}^{\text{in},N}_{u,t}$) is equal to the water distributed to existing storage tanks ($D^{N,\text{Esto}}_{u,p,t}$), plus the water distributed to new storage tanks ($D^{N,\text{Nsto}}_{u,q,t}$), the water required to distribution stations for: domestic ($v^{N,\text{dom}}_{r,u,t}$), industrial ($v^{N,\text{ind}}_{o,u,t}$) and agricultural ($v^{N,\text{agr}}_{g,u,t}$) uses, also the water sent to recharge aquifers ($B^{N,\text{des}}_{u,i,t}$), the water sent to recharge dams ($G^{N,\text{rel}}_{u,x,t}$) and the water sent to the sea as reject ($b^{N,\text{rej}}_{u,t}$).

$$\begin{aligned}\text{SW}^{\text{in},N}_{u,t}(1-\beta) = {} & \sum_p D^{N,\text{Esto}}_{u,p,t} + \sum_q D^{N,\text{Nsto}}_{u,q,t} \quad \forall u \in U,\ \forall t \in T \\ & + \sum_r v^{N,\text{dom}}_{r,u,t} + \sum_g v^{N,\text{agr}}_{g,u,t} + \sum_o v^{N,\text{ind}}_{o,u,t} \\ & + \sum_i B^{N,\text{des}}_{u,i,t} + \sum_x G^{N,\text{rel}}_{u,x,t} + b^{N,\text{rej}}_{u,t}\end{aligned} \tag{6}$$

Binary variables (y^{pdes}_u) determine the existence and location of new power-desalination plants, if the binary variable is equal to one, then the power-desalination plant is needed, if the binary variable is equal to zero, the power-desalination plant is not needed, also the power-desalination plant has a maximum capacity ($\Theta^{\text{pdes,max}}_u$) and a minimum capacity ($\Theta^{\text{pdes,min}}_u$). Furthermore, the model considers the installation of the new power-desalination plant at the first period, taking into account the maximum capacity of all periods.

$$y_u^{\text{pdes}} \cdot \Theta_u^{\text{pdes,min}} \leq \text{SW}_u^{\max} \leq y_u^{\text{pdes}} \cdot \Theta_u^{\text{pdes,max}} \tag{7}$$

The installation cost ($\text{InstCost}_{u,t}^{\text{pdes}}$) is a function of the unit costs (Z_3 and Z_4), the maximum seawater capacity ($\text{SW}_u^{\max}$) and a factor used to annualize the inversion (k_F). The operating cost ($\text{OpCost}_{u,t}^{\text{pdes}}$) is a function of a unit operating cost (Z_5), the total water extracted from the sea ($\text{SW}_{u,t}^{\text{in},N}$), the recovery factor ($1-\beta$) and a factor used to account for the operational time per year (H_Y). The operating costs include maintenance, utility expenses, operating labor cost, product costs, raw material cost, fuel, and energy costs.

$$\text{InstCost}_u^{\text{pdes}} = k_F\left[Z_3 \cdot y_u^{\text{pdes}} + Z_4\left(\text{SW}_u^{\max}\right)^{\alpha}\right] \quad \forall u \in U \tag{8}$$

$$\text{OpCost}_{u,t}^{\text{pdes}} = H_Y\left[Z_5(1-\beta)\text{SW}_{u,t}^{\text{in},N}\right] \quad \forall u \in U,\ \forall t \in T \tag{9}$$

where $\text{SW}_u^{\max}$ is greater or equal to the water processed in the new power-desalination plant $\text{SW}_{u,t}^{\text{in},N}$ in any time period t.

$$\text{SW}_{u,t}^{\text{in},N} \leq \text{SW}_u^{\max} \quad q \in Q,\ t \in T \tag{10}$$

The operating cost is a function of the total water extracted from the sea ($\text{SW}_{n,t}^{\text{in},E}$) multiplying by the total recovery ($1-\beta$) and the operational unit cost (Z_6), H_Y is a factor used to account the operational time per year.

$$\text{OpCost}_{n,t}^{E,\text{des}} = H_Y\left[Z_6(1-\beta)\text{SW}_{n,t}^{\text{in},E}\right] \quad \forall n \in N,\ \forall t \in T \tag{11}$$

Then, the installation and operation cost for new and existing power-desalination plants can be calculated as follows:

$$\text{NPDinstcost} = \sum_u \text{InstCost}_u^{\text{pdes}} \tag{12}$$

$$\text{NPDopcost} = \sum_u \sum_t \text{OpCost}_{u,t}^{\text{pdes}} \tag{13}$$

$$\text{EPDopcost} = \sum_n \sum_t \text{OpCost}_{n,t}^{E,\text{des}} \tag{14}$$

2.2 Water Balances in Natural Resources

Water withdrawals and recharge are modeled using water balances, the mass balances involve the water level in a certain time and the mass balance in the previous time, in this case the natural resources involve the water contained in aquifers and dams.

The total water in a dam ($M_{x,t}$) minus the water in the same dam in a previous time period ($M_{x,t-1}$) is equal to the sum of the water received from existing power-desalination plants ($G^{E,\text{rel}}_{n,x,t}$) and the water received from new power-desalination plants ($G^{N,\text{rel}}_{u,x,t}$), plus the water from natural recharge ($R^{\text{dam}}_{x,t}$), minus the water sent to satisfy water needs for domestic users ($w^{\text{dom}}_{r,x,t}$), industrial users ($w^{\text{agr}}_{g,x,t}$), and agricultural users ($w^{\text{ind}}_{o,x,t}$).

$$M_{x,t} - M_{x,t-1} = \sum_{n} G^{E,\text{rel}}_{n,x,t} + \sum_{u} G^{N,\text{rel}}_{u,x,t} + R^{\text{dam}}_{x,t} \quad \forall x \in X,\ \forall t \in T, t \neq 1 \\ - \sum_{r} w^{\text{dom}}_{r,x,t} - \sum_{g} w^{\text{agr}}_{g,x,t} - \sum_{o} w^{\text{ind}}_{o,x,t} \tag{15}$$

The total water accumulated in an aquifer over a given time period ($W^{\text{aq}}_{i,t}$) minus the one in the previous time period ($W^{\text{aq}}_{i,t-1}$) is equal to the sum of the water received from the existing storage tanks ($s^{E,\text{aq}}_{p,i,t}$), plus the water received from new storage tanks ($s^{N,\text{aq}}_{q,i,t}$), plus the water received from existing power-desalination plants ($B^{E,\text{des}}_{n,i,t}$), plus the water received from new power-desalination plants ($B^{N,\text{des}}_{u,i,t}$), plus the water recharge from the agricultural users ($F^{\text{agr}}_{i,t}$), plus the water from natural recharge ($R^{\text{aq}}_{i,t}$), minus the water that is sent to the deep wells ($a^{\text{dw}}_{i,j,t}$).

$$W_{i,t} - W_{i,t-1} = \sum_{p} s^{E,\text{aq}}_{p,i,t} + \sum_{q} s^{N,\text{aq}}_{q,i,t} + \sum_{n} B^{E,\text{des}}_{n,i,t} \quad \forall i \in I,\ \forall t \in T, t \neq 1 \\ + \sum_{u} B^{N,\text{des}}_{u,i,t} + F^{\text{agr}}_{i,t} + R^{\text{aq}}_{i,t} - \sum_{j} a^{\text{dw}}_{i,j,t} \tag{16}$$

Part of the water in aquifers is extracted using deep wells, the water in deep wells ($a^{\text{dw}}_{i,j,t}$) is sent to the distribution stations including domestic station ($d^{\text{dom}}_{r,j,t}$), agricultural station ($d^{\text{agr}}_{g,j,t}$), and industrial station ($d^{\text{ind}}_{o,j,t}$).

$$\sum_{i} a^{\text{dw}}_{i,j,t} = \sum_{r} d^{\text{dom}}_{r,j,t} + \sum_{g} d^{\text{agr}}_{g,j,t} + \sum_{o} d^{\text{ind}}_{o,j,t} \quad \forall j \in J,\ \forall t \in T \tag{17}$$

2.3 Water Balance in Distribution Stations

Water distribution to users is through pumping stations and there is a distribution station for each user (r = domestic, g = agricultural and o = industrial). The total water in a domestic station over a given time period ($h^{\text{dom}}_{r,t}$) is equal to the water received from existent power-desalination plants ($v^{E,\text{dom}}_{r,n,t}$), plus the water received from new power-desalination plans ($v^{N,\text{dom}}_{r,u,t}$), deep wells ($d^{\text{dom}}_{r,j,t}$) and dams ($w^{\text{dom}}_{r,x,t}$).

$$h_{r,t}^{\mathrm{dom}} = \sum_{n} v_{r,n,t}^{E,\mathrm{dom}} + \sum_{u} v_{r,u,t}^{N,\mathrm{dom}} + \sum_{j} d_{r,j,t}^{\mathrm{dom}} + \sum_{x} w_{r,x,t}^{\mathrm{dom}} \quad \forall r \in R,\ \forall t \in T \tag{18}$$

The total water in an agricultural station ($h_{g,t}^{\mathrm{agr}}$) is equal to the water received from existent power-desalination plants ($v_{g,n,t}^{E,\mathrm{agr}}$), plus the water received from new power-desalination plans ($v_{g,u,t}^{N,\mathrm{agr}}$), deep wells ($d_{g,j,t}^{\mathrm{agr}}$) and dams ($w_{g,x,t}^{\mathrm{agr}}$).

$$h_{g,t}^{\mathrm{agr}} = \sum_{n} v_{g,n,t}^{E,\mathrm{agr}} + \sum_{u} v_{g,u,t}^{N,\mathrm{agr}} + \sum_{j} d_{g,j,t}^{\mathrm{agr}} + \sum_{x} w_{g,x,t}^{\mathrm{agr}} \quad \forall g \in G,\ \forall t \in T \tag{19}$$

The total water in an industrial station ($h_{o,t}^{\mathrm{ind}}$) is equal to the water received from existent power-desalination plants ($v_{o,n,t}^{E,\mathrm{ind}}$), plus the water received from new power-desalination plans ($v_{o,u,t}^{N,\mathrm{ind}}$), deep wells ($d_{o,j,t}^{\mathrm{ind}}$) and dams ($w_{o,x,t}^{\mathrm{ind}}$).

$$h_{o,t}^{\mathrm{ind}} = \sum_{n} v_{o,n,t}^{E,\mathrm{ind}} + \sum_{u} v_{o,u,t}^{N,\mathrm{ind}} + \sum_{j} d_{o,j,t}^{\mathrm{ind}} + \sum_{x} w_{o,x,t}^{\mathrm{ind}} \quad \forall o \in O,\ \forall t \in T \tag{20}$$

2.4 Pumping Cost

Pumping cost is one of the most important issues, the pumping cost can be calculated using the next expression:

$$\begin{aligned}
\text{Pumping cost} = H_Y \Bigg[& \sum_{r}\sum_{t} h_{r,t}^{\mathrm{dom}}\mathrm{PPD1}_{r,t} + \sum_{p}\sum_{r}\sum_{t} s_{p,r,t}^{E,\mathrm{dom}}\mathrm{PPD2}_{p,r,t} + \sum_{q}\sum_{r}\sum_{t} s_{q,r,t}^{N,\mathrm{dom}}\mathrm{PPD3}_{q,r,t} \\
& + \sum_{g}\sum_{t} h_{g,t}^{\mathrm{agr}}\mathrm{PPA1}_{g,t} + \sum_{g}\sum_{p}\sum_{t} s_{g,p,t}^{E,\mathrm{agr}}\mathrm{PPA2}_{g,p,t} + \sum_{q}\sum_{g}\sum_{t} s_{q,g,t}^{N,\mathrm{agr}}\mathrm{PPA3}_{q,g,t} \\
& + \sum_{o}\sum_{t} h_{o,t}^{\mathrm{ind}}\mathrm{PPI1}_{o,t} + \sum_{o}\sum_{p}\sum_{t} s_{o,p,t}^{E,\mathrm{ind}}\mathrm{PPI2}_{o,p,t} + \sum_{q}\sum_{o}\sum_{t} s_{q,o,t}^{N,\mathrm{ind}}\mathrm{PPI3}_{q,o,t} \\
& + \sum_{r}\sum_{j}\sum_{t} d_{r,j,t}^{\mathrm{dom}}\mathrm{PPD4}_{r,j,t} + \sum_{g}\sum_{j}\sum_{t} d_{g,j,t}^{\mathrm{agr}}\mathrm{PPA4}_{g,j,t} + \sum_{o}\sum_{j}\sum_{t} d_{o,j,t}^{\mathrm{ind}}\mathrm{PPI4}_{o,j,t} \\
& + \sum_{r}\sum_{x}\sum_{t} w_{r,x,t}^{\mathrm{dom}}\mathrm{PPD5}_{r,x,t} + \sum_{g}\sum_{x}\sum_{t} w_{g,x,t}^{\mathrm{agr}}\mathrm{PPA5}_{g,x,t} + \sum_{o}\sum_{x}\sum_{t} w_{o,x,t}^{\mathrm{ind}}\mathrm{PPI5}_{o,x,t} \\
& + \sum_{n}\sum_{i}\sum_{t} B_{n,i,t}^{E,\mathrm{des}}\mathrm{PPC1}_{n,i,t} + \sum_{u}\sum_{i}\sum_{t} B_{u,i,t}^{N,\mathrm{des}}\mathrm{PPC2}_{u,i,t} + \sum_{r}\sum_{n}\sum_{t} v_{r,n,t}^{E,\mathrm{dom}}\mathrm{PPD6}_{r,n,t} \\
& + \sum_{r}\sum_{u}\sum_{t} v_{r,u,t}^{N,\mathrm{dom}}\mathrm{PPD7}_{r,u,t} + \sum_{g}\sum_{n}\sum_{t} v_{g,n,t}^{E,\mathrm{agr}}\mathrm{PPA6}_{g,n,t} + \sum_{g}\sum_{u}\sum_{t} v_{g,u,t}^{N,\mathrm{agr}}\mathrm{PPA7}_{g,u,t} \\
& + \sum_{o}\sum_{n}\sum_{t} v_{o,n,t}^{E,\mathrm{ind}}\mathrm{PPI6}_{o,n,t} + \sum_{o}\sum_{u}\sum_{t} v_{o,u,t}^{N,\mathrm{ind}}\mathrm{PPI7}_{o,u,t} + \sum_{n}\sum_{q}\sum_{t} D_{n,q,t}^{E,\mathrm{Nsto}}\mathrm{PPE1}_{n,q,t} \\
& + \sum_{n}\sum_{p}\sum_{t} D_{n,p,t}^{E,\mathrm{Esto}}\mathrm{PPE2}_{n,p,t} + \sum_{u}\sum_{q}\sum_{t} D_{u,q,t}^{N,\mathrm{Nsto}}\mathrm{PPN1}_{u,q,t} + \sum_{u}\sum_{p}\sum_{t} D_{u,p,t}^{N,\mathrm{Esto}}\mathrm{PPN2}_{u,p,t} \\
& + \sum_{n}\sum_{x}\sum_{t} G_{n,x,t}^{E,\mathrm{rel}}\mathrm{PPG1}_{n,x,t} + \sum_{u}\sum_{x}\sum_{t} G_{u,x,t}^{N,\mathrm{rel}}\mathrm{PPG2}_{u,x,t} + \sum_{i}\sum_{j}\sum_{t} a_{i,j,t}^{\mathrm{dw}}\mathrm{PAQ1}_{i,j,t} \\
& + \sum_{p}\sum_{i}\sum_{t} s_{p,i}^{E,\mathrm{aq}}\mathrm{PPS1}_{p,i,t} + \sum_{q}\sum_{i}\sum_{t} s_{q,i,t}^{N,\mathrm{aq}}\mathrm{PPS2}_{q,i,t} \Bigg]
\end{aligned} \tag{21}$$

The pumping cost factors are determined through the Darcy-Weisbach equation:

$$\text{Pumping Cost Factor} = \frac{1}{0.0000576} f \frac{L}{D^5} \frac{(\#\text{ of hours})(\$/\text{kWh})}{\eta}, \tag{22}$$

where f is the friction factor, L is the pipe equivalent length, D is the pipe inner diameter and η is the combined pump and motor efficiency. The friction factor is based on the pipe roughness, pipe diameter, and the Reynolds number. The idea is to approximate pumping cost to the real cost knowing the flux and fixing the pipe type, pipe length, and pump efficiency.

2.5 Piping Costs

Water distribution is carried out through pipes, and they represent a key part of the model because the existence of pipes determines the optimal water network. The distribution can be calculated using the next relationship:

$$\begin{aligned}
\text{Piping cost} = k_F \Bigg[& \sum_r \text{DPC1}_r \cdot y_r^{h,\text{dom}} + \sum_p \sum_r \text{DPC2}_{p,r} \cdot y_{p,r}^{sE,\text{dom}} + \sum_q \sum_r \text{DPC3}_{q,r} \cdot y_{q,r}^{sN,\text{dom}} \\
& + \sum_g \text{APC1}_g \cdot y_g^{h,\text{agr}} + \sum_p \sum_g \text{APC2}_{p,g} \cdot y_{g,p}^{sE,\text{agr}} + \sum_q \sum_g \text{APC3}_{q,g} \cdot y_{g,q}^{sN,\text{agr}} \\
& + \sum_o \text{IPC1}_o \cdot y_o^{h,\text{ind}} + \sum_p \sum_o \text{IPC2}_{p,o} \cdot y_{o,p}^{sE,\text{ind}} + \sum_q \sum_o \text{IPC3}_{q,o} \cdot y_{q,o}^{sN,\text{ind}} \\
& + \sum_r \sum_j \text{DPC4}_{r,j} \cdot y_{r,j}^{d,\text{dom}} + \sum_g \sum_j \text{APC4}_{g,j} \cdot y_{g,j}^{d,\text{agr}} + \sum_o \sum_j \text{IPC4}_{o,j} \cdot y_{o,j}^{d,\text{ind}} \\
& + \sum_r \sum_x \text{DPC5}_{r,x} \cdot y_{r,x}^{w,\text{dom}} + \sum_g \sum_x \text{APC5}_{g,x} \cdot y_{g,x}^{w,\text{agr}} + \sum_o \sum_x \text{IPC5}_{o,x} \cdot y_{o,x}^{w,\text{ind}} \\
& + \sum_n \sum_i \text{BPC1}_{n,i} \cdot y_{n,i}^{BE,\text{des}} + \sum_u \sum_i \text{BPC2}_{u,i} \cdot y_{u,i}^{BN,\text{des}} + \sum_r \sum_n \text{DPC6}_{r,n} \cdot y_{r,n}^{vE,\text{dom}} \\
& + \sum_r \sum_u \text{DPC7}_{r,u} \cdot y_{r,u}^{vN,\text{dom}} + \sum_g \sum_n \text{APC6}_{g,n} \cdot y_{g,n}^{vE,\text{agr}} + \sum_g \sum_u \text{APC7}_{g,u} \cdot y_{g,u}^{vN,\text{agr}} \\
& + \sum_o \sum_n \text{IPC6}_{o,n} \cdot y_{o,n}^{vE,\text{ind}} + \sum_o \sum_u \text{IPC7}_{o,u} \cdot y_{o,u}^{vN,\text{ind}} + \sum_n \sum_q \text{EPC1}_{n,q} \cdot y_{n,q}^{DE,\text{Nsto}} \\
& + \sum_n \sum_p \text{EPC2}_{n,p} \cdot y_{n,p}^{DE,\text{Esto}} + \sum_u \sum_q \text{NPC1}_{u,q} \cdot y_{u,q}^{DN,\text{Nsto}} + \sum_u \sum_p \text{NPC2}_{u,p} \cdot y_{u,p}^{DN,\text{Esto}} \\
& + \sum_n \sum_x \text{GPC1}_{n,x} \cdot y_{n,x}^{GE,\text{rel}} + \sum_u \sum_x \text{GPC2}_{u,x} \cdot y_{u,x}^{GN,\text{rel}} + \sum_i \sum_j \text{AQP1}_{i,j} \cdot y_{i,j}^{a,\text{dw}} \\
& + \sum_p \sum_i \text{SPC1}_{p,i} \cdot y_{p,i}^{sE,\text{aq}} + \sum_q \sum_i \text{SPC2}_{q,i} \cdot y_{q,i}^{sN,\text{aq}}
\end{aligned} \Bigg] \tag{23}$$

The piping cost factors are determined through the next equation:

$$\text{Piping Cost Factor} = k_m L D^m \tag{24}$$

where L is the pipe length, D^m is the pipe diameter, k_m and m are pipe cost parameters that depend on the pipe material. If the diameter of the pipes, the distance from the storage points to final users and height are fixed, then the flowrate determines the piping cost. The existence of the pipe is determined by the binary variable, to activate the pipe cost, and each stream has its own binary variable to calculate the distribution cost. Maximum values to determine piping cost can be calculated as follows:

For the pipeline to the user from distribution station.

$$h_{r,t}^{\text{dom}} \leq h^{\text{dom,max}} \cdot y_r^{h,\text{dom}} \tag{25}$$

$$h_{g,t}^{\text{agr}} \leq h^{\text{agr,max}} \cdot y_g^{h,\text{agr}} \tag{26}$$

$$h_{o,t}^{\text{ind}} \leq h^{\text{ind,max}} \cdot y_o^{h,\text{ind}} \tag{27}$$

For the pipelines to the user from existing storage tanks.

$$s_{p,r,t}^{E,\text{dom}} \leq s^{E,\text{dom,max}} \cdot y_{p,r}^{sE,\text{dom}} \tag{28}$$

$$s_{g,p,t}^{E,\text{agr}} \leq s^{E,\text{agr,max}} \cdot y_{g,p}^{sE,\text{agr}} \tag{29}$$

$$s_{o,p,t}^{E,\text{ind}} \leq s^{E,\text{ind,max}} \cdot y_{o,p}^{sE,\text{ind}} \tag{30}$$

For the pipelines to the user from new storage tanks.

$$s_{q,r,t}^{N,\text{dom}} \leq s^{N,\text{dom,max}} \cdot y_{q,r}^{sN,\text{dom}} \tag{31}$$

$$s_{g,q,t}^{N,\text{agr}} \leq s^{N,\text{agr,max}} \cdot y_{g,q}^{sN,\text{agr}} \tag{32}$$

$$s_{q,o,t}^{N,\text{ind}} \leq s^{N,\text{ind,max}} \cdot y_{q,o}^{sN,\text{ind}} \tag{33}$$

For the pipelines to the distribution station from the deep wells.

$$d_{r,j,t}^{\text{dom}} \leq d^{\text{dom,max}} \cdot y_{r,j}^{d,\text{dom}} \tag{34}$$

$$d_{g,j,t}^{\text{agr}} \leq d^{\text{agr,max}} \cdot y_{g,j}^{d,\text{agr}} \tag{35}$$

$$d_{o,j,t}^{\text{ind}} \leq d^{\text{ind,max}} \cdot y_{o,j}^{d,\text{ind}} \tag{36}$$

For the pipelines to the distribution station from the dams.

$$w_{r,x,t}^{\text{dom}} \leq w^{\text{dom,max}} \cdot y_{r,x}^{w,\text{dom}} \tag{37}$$

$$w_{g,x,t}^{\text{agr}} \leq w^{\text{agr,max}} \cdot y_{g,x}^{w,\text{agr}} \tag{38}$$

$$w_{o,x,t}^{\text{ind}} \leq w^{\text{ind,max}} \cdot y_{o,x}^{w,\text{ind}} \tag{39}$$

For the pipelines to recharge aquifers from the power-desalination plant.

$$B_{n,i,t}^{E,\text{des}} \leq B^{E,\text{des,max}} \cdot y_{n,i}^{BE,\text{des}} \tag{40}$$

$$B_{u,i,t}^{N,\text{des}} \leq B^{N,\text{des,max}} \cdot y_{u,i}^{BN,\text{des}} \tag{41}$$

For the pipelines to distribution stations from existing power-desalination plants.

$$v_{r,n,t}^{E,\text{dom}} \leq v^{E,\text{dom,max}} \cdot y_{r,n}^{vE,\text{dom}} \tag{42}$$

$$v_{g,n,t}^{E,\text{agr}} \leq v^{E,\text{agr,max}} \cdot y_{g,n}^{vE,\text{agr}} \tag{43}$$

$$v_{o,n,t}^{E,\text{ind}} \leq v^{E,\text{ind,max}} \cdot y_{o,n}^{vE,\text{ind}} \tag{44}$$

For the pipelines to the distribution stations from new power-desalination plants.

$$v_{r,u,t}^{N,\text{dom}} \leq v^{N,\text{dom,max}} \cdot y_{r,u}^{vN,\text{dom}} \tag{45}$$

$$v_{g,u,t}^{N,\text{agr}} \leq v^{N,\text{agr,max}} \cdot y_{g,u}^{vN,\text{agr}} \tag{46}$$

$$v_{o,u,t}^{N,\text{ind}} \leq v^{N,\text{ind,max}} \cdot y_{o,u}^{vN,\text{ind}} \tag{47}$$

For the pipelines to water storage tanks from existing power-desalination plants.

$$D_{n,q,t}^{E,\text{Nsto}} \leq D^{E,\text{Nsto,max}} \cdot y_{n,q}^{DE,\text{Nsto}} \tag{48}$$

$$D_{n,p,t}^{E,\text{Esto}} \leq D^{E,\text{Esto,max}} \cdot y_{n,p}^{DE,\text{Esto}} \tag{49}$$

For the pipelines to the water storage tanks from new power-desalination plants.

$$D_{u,q,t}^{N,\text{Nsto}} \leq D^{N,\text{Nsto,max}} \cdot y_{u,q}^{DN,\text{Nsto}} \tag{50}$$

$$D_{u,p,t}^{N,\text{Esto}} \leq D^{N,\text{Esto,max}} \cdot y_{u,p}^{DN,\text{Esto}} \tag{51}$$

For the pipelines to recharge dams from power-desalination plants.

$$G_{n,x,t}^{E,\mathrm{rel}} \leq G^{E,\mathrm{rel,max}} \cdot y_{n,x}^{GE,\mathrm{rel}} \tag{52}$$

$$G_{u,x,t}^{N,\mathrm{rel}} \leq G^{N,\mathrm{rel,max}} \cdot y_{u,x}^{GN,\mathrm{rel}} \tag{53}$$

Pipeline to deep wells from aquifers.

$$a_{i,j,t}^{\mathrm{dw}} \leq a^{\mathrm{dw,max}} \cdot y_{i,j}^{a,\mathrm{dw}} \tag{54}$$

Pipeline to recharge aquifers from storage tanks.

$$s_{p,i,t}^{E,\mathrm{aq}} \leq s^{E,\mathrm{aq,max}} \cdot y_{p,i}^{sE,\mathrm{aq}} \tag{55}$$

$$s_{q,i,t}^{N,\mathrm{aq}} \leq s^{N,\mathrm{aq,max}} \cdot y_{q,i}^{sN,\mathrm{aq}} \tag{56}$$

2.6 Water Demands and Energy Demands

The water requirements of domestic users ($\mathrm{domdem}_{r,t}$) are satisfied with the water from the domestic station ($h_{r,t}^{\mathrm{dom}}$), plus the water from existing water storage tanks ($s_{p,r,t}^{E,\mathrm{dom}}$) and new water storage tanks ($s_{q,r,t}^{N,\mathrm{dom}}$).

$$\mathrm{domdem}_{r,t} = h_{r,t}^{\mathrm{dom}} + \sum_{p} s_{p,r,t}^{E,\mathrm{dom}} + \sum_{q} s_{q,r,t}^{N,\mathrm{dom}}, \quad \forall r \in R,\ \forall t \in T \tag{57}$$

The water requirements of industrial users ($\mathrm{inddem}_{o,t}$) are satisfied with water from the corresponding agricultural station ($h_{o,t}^{\mathrm{ind}}$), plus water from existing water storage tanks ($s_{o,p,t}^{E,\mathrm{ind}}$) and new water storage tanks ($s_{o,q,t}^{N,\mathrm{ind}}$).

$$\mathrm{inddem}_{o,t} = h_{o,t}^{\mathrm{ind}} + \sum_{p} s_{o,p,t}^{E,\mathrm{ind}} + \sum_{q} s_{o,q,t}^{N,\mathrm{ind}}, \quad \forall o \in O,\ \forall t \in T \tag{58}$$

The water requirements of agricultural users ($\mathrm{agrdem}_{g,t}$) are satisfied with the water from the corresponding agricultural station ($h_{g,t}^{\mathrm{agr}}$), plus the water from existing water storage tanks ($s_{g,p,t}^{E,\mathrm{agr}}$) and new water storage tanks ($s_{g,q,t}^{N,\mathrm{agr}}$).

$$\mathrm{agrdem}_{g,t} = h_{g,t}^{\mathrm{agr}} + \sum_{p} s_{g,p,t}^{E,\mathrm{agr}} + \sum_{q} s_{g,q,t}^{N,\mathrm{agr}}, \quad \forall g \in G,\ \forall t \in T \tag{59}$$

Part of the used water for agriculture (*pca*) seeps into the ground, thus may recharge the aquifer ($F_{i,t}^{\text{agr}}$) as follows.

$$F_{i,t}^{\text{agr}} = \sum_{g(i)} \text{pca} * \text{agrdem}_{g,i,t} \quad \forall i \in I,\ \forall t \in T \tag{60}$$

where the water that seeps into the ground in a given agricultural user (g) associated to certain aquifer (i).

2.7 Equations for Storage Tanks

It is important to consider the existing water storage and the possible installation of new storage tanks if is necessary. New and existing storage tanks are modeled using accumulation balances. The total water balance in existing storage tanks ($S_{p,t}^{E}$) minus the water in a previous time period ($S_{p,t-1}^{E}$) is equal to the water received from existing power-desalination plants ($D_{n,p,t}^{E,\text{Esto}}$) plus the water received from new power-desalination plants ($D_{u,p,t}^{N,\text{Esto}}$), minus the water sent to domestic users ($s_{p,r,t}^{E,\text{dom}}$), agricultural users ($s_{g,p,t}^{E,\text{agr}}$), industrial users ($s_{o,p,t}^{E,\text{ind}}$) and the one sent to aquifers ($s_{p,i,t}^{E,\text{aq}}$).

$$\begin{aligned} S_{p,t}^{E} - S_{p,t-1}^{E} = &\sum_{n} D_{n,p,t}^{E,\text{Esto}} + \sum_{u} D_{u,p,t}^{N,\text{Esto}} - \sum_{r} s_{p,r,t}^{E,\text{dom}} \quad \forall p \in P,\ \forall t \in T, t \neq 1 \\ &- \sum_{g} s_{g,p,t}^{E,\text{agr}} - \sum_{o} s_{o,p,t}^{E,\text{ind}} - \sum_{i} s_{p,i,t}^{E,\text{aq}} \end{aligned} \tag{61}$$

The total water balance in new storage tanks ($S_{q,t}^{N}$) minus the water in a previous time period ($S_{q,t-1}^{N}$), is equal to the water received from existing power-desalination plants ($D_{n,q,t}^{E,\text{Nsto}}$) plus water received from new power-desalination plants ($D_{u,q,t}^{N,\text{Nsto}}$), minus water sent to domestic users ($s_{q,r,t}^{N,\text{dom}}$), agricultural users ($s_{g,q,t}^{N,\text{agr}}$), industrial users ($s_{o,q,t}^{N,\text{ind}}$) and the one sent to the aquifers ($s_{q,i,t}^{N,\text{aq}}$).

$$\begin{aligned} S_{q,t}^{N} - S_{q,t-1}^{N} = &\sum_{n} D_{n,q,t}^{E,\text{Nsto}} + \sum_{u} D_{u,q,t}^{N,\text{Nsto}} - \sum_{r} s_{q,r,t}^{N,\text{dom}} \quad \forall q \in Q,\ \forall t \in T, t \neq 1 \\ &- \sum_{g} s_{g,q,t}^{N,\text{agr}} - \sum_{o} s_{o,q,t}^{N,\text{ind}} - \sum_{i} s_{q,i,t}^{N,\text{aq}} \end{aligned} \tag{62}$$

The existence and location of new storage tanks in a certain region depend on the water demand of the users, the location, and existence is modeled through binary variables ($y_{q,t}^{\text{sto}}$), if the binary variable is equal to one, then the tank is needed,

if the binary variable is equal to zero, the tank is not required, also the existence depends on the maximum capacity of the tank ($\Theta_q^{\text{sto,max}}$) and the minimum capacity of the tank ($\Theta_q^{\text{sto,min}}$).

$$y_q^{\text{sto}} \cdot \Theta_q^{\text{sto,min}} \leq S_q^{\max} \leq y_q^{\text{sto}} \cdot \Theta_q^{\text{sto,max}}, \quad \forall q \in Q \tag{63}$$

The associated installation cost ($\text{InstCost}_{q,t}^{\text{sto}}$) is a function of a fixed cost of the tank (Z_1), the unit variable cost (Z_2), and a factor used to annualize the cost (k_F). In addition, $S_q^{\max}$ represents the maximum capacity for the storage tank.

$$\text{InstCost}_q^{\text{sto}} = k_F \left[Z_1 \cdot y_q^{\text{sto}} + Z_2 \left(S_q^{\max} \right)^{\alpha} \right], \quad \forall q \in Q \tag{64}$$

where $S_q^{\max}$ is greater than the existing water in storage tank $S_{q,t}^N$ in any time period t.

$$S_{q,t}^N \leq S_q^{\max} \quad q \in Q, t \in T \tag{65}$$

where, the overall storage cost can be calculated by the next equation:

$$\text{Storage Cost} = \sum_q \text{InstCost}_q^{\text{sto}} \tag{66}$$

2.8 Objective Function

The objective function consists in maximizing the gross annual profit, which considers the energy and water sales for domestic, industrial, and agricultural users, minus the total annual cost.

$$\text{Annual Profit} = \text{Water Sales} + \text{Energy Sales} - \text{TAC} \tag{67}$$

where the annual water sales can be calculated through the next equation:

$$\text{Water Sales} = H_Y \sum_t \left[\begin{array}{l} \left(\sum_r h_{r,t}^{\text{dom}} + \sum_p \sum_r s_{p,r,t}^{E,\text{dom}} + \sum_q \sum_r s_{q,r,t}^{N,\text{dom}} \right) \text{WDC}_t \\ + \left(\sum_g h_{g,t}^{\text{agr}} + \sum_g \sum_p s_{g,p,t}^{E,\text{agr}} + \sum_q \sum_g s_{q,g,t}^{N,\text{agr}} \right) \text{WAC}_t \\ + \left(\sum_o h_{o,t}^{\text{ind}} + \sum_o \sum_p s_{o,p,t}^{E,\text{ind}} + \sum_q \sum_o s_{q,o,t}^{N,\text{ind}} \right) \text{WIC}_t \end{array} \right] \tag{68}$$

And the annual energy sales can be calculated by the next relationship:

$$\text{Energy Sales} = H_Y \sum_t \left[\sum_r E_{r,t}^{\text{dom}} \cdot \text{DEC}_t + \sum_g E_{g,t}^{\text{agr}} \cdot \text{AEC}_t + \sum_o E_{o,t}^{\text{ind}} \cdot \text{IEC}_t \right] \tag{69}$$

Also, the total annual cost (TAC) accounts for the sum of storage cost (Storage Cost), piping cost (Piping Cost), pumping cost (Pumping Cost), installation cost for new power-desalination plants (NPDinstcost), operating costs for new power-desalination plants (NPDopcost) and operating costs for existing power-desalination plants (EPDopcost) as follows:

$$\begin{aligned} \text{TAC} = {} & \text{Storage Cost} + \text{Piping Cost} + \text{Pumping Cost} \\ & + \text{NPDinstcost} + \text{NPDopcost} + \text{EPDopcost} \end{aligned} \tag{70}$$

3 Case Study

As a case study, the problem of water scarcity in the city of Hermosillo in Mexico was selected to demonstrate the applicability of the proposed optimization approach. The urbanized area of Hermosillo is located in the lower part of the Sonora River basin, near the "El Molinito" and "Abelardo L. Rodriguez" dams. These dams provided water to the city, but these sources have been exhausted. The main reason for this problem has been the lack of rain, prolonged drought periods, and rapid population growth in the region (SEMARNAT 2012). For this case, the industrial, agricultural, and domestic demands are considered. Also, it is considered the population involved in the region, the dams as natural resources, the existing aquifers in the region and the existing power plants. Figure 2 shows the main water bodies and streams of the state of Sonora.

3.1 Domestic Users

In the domestic sector, the water demands for the cities of Hermosillo, Obregon, and Guaymas are considered. Figure 3 shows the electricity demands in each city. Notice that the electricity demands increase in the hottest months because of the use of cooling systems. Figure 4 shows the variation of water demands for domestic users in each month, in which the water demands decrease in the hottest months because of the scarcity of this resource.

Fig. 2 Major river basins on the state of Sonora (CONAGUA 2013)

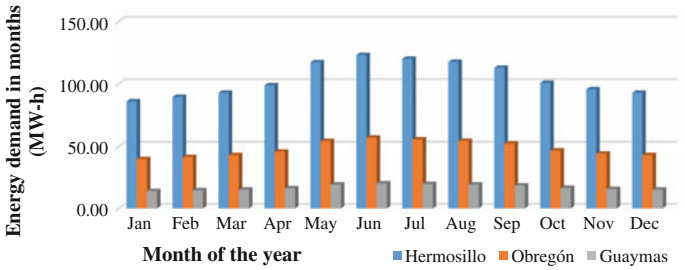


Fig. 3 Electricity demands for domestic users (SENER 2014)

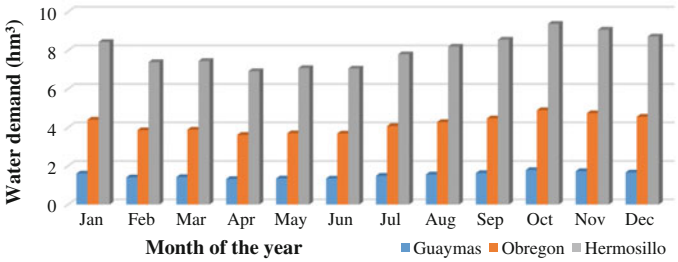


Fig. 4 Variation of water demands for domestic users (CEA 2008)

3.2 *Agricultural Users*

Four irrigation districts are considered in the case study, 084 Guaymas, 051 Costa de Hermosillo, 041 Yaqui River, and 018 Colonias Yaquis. Figure 5 shows the variation of water demands in each region through the year, this variation is based on the water requirements and the irrigation stages of each crop considering two growing seasons, March–August and October–January. The electricity demands for agricultural users are shown in Fig. 6.

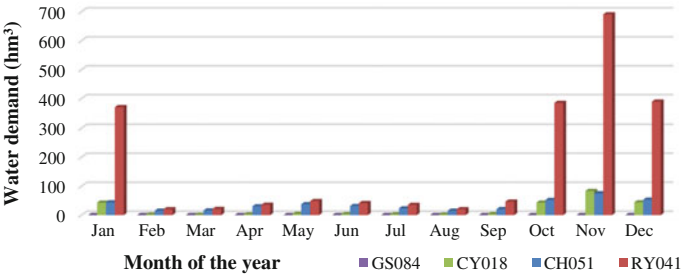


Fig. 5 Variation of water demands for agricultural users (CONAGUA 2014)

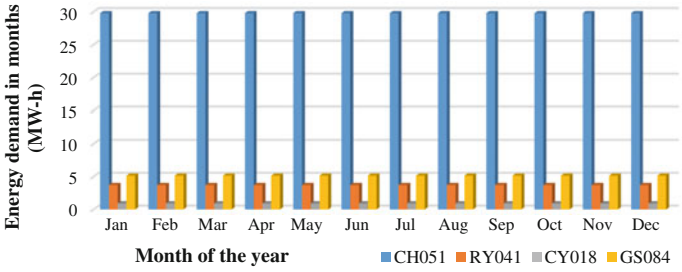


Fig. 6 Electricity demand for agricultural users (SENER 2014)

3.3 Industrial Users

The water demand in the industrial sector is estimated as 69 hm^3, of which 40 million m^3 correspond to groundwater and 29 hm^3 to surface water. Figure 7 shows the water demands for industrial users, where these demands are relatively constant throughout the year. Figure 8 shows the energy demands for industrial users.

3.4 Surface Water

Five hydrologic regions are considered in the state of Sonora, from which three are involved in the considered case study (Sonora River region, Mátape River region,

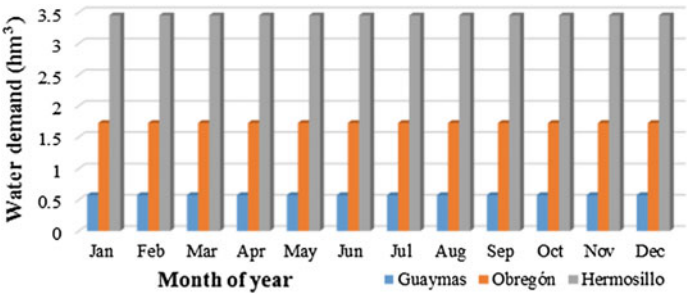


Fig. 7 Variation of water demands for industrial users (CEA 2008)

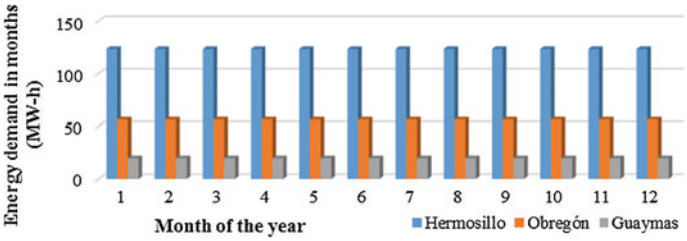


Fig. 8 Electricity demand in industrial users (SENER 2014)

Table 1 Characteristics of the main hydrological flows

Characteristic	Sonora river	Yaqui river	Matape river
Natural runoff (Million hm^3/year)	439	3,163	90
Area (km^2)	27,740	72,540	6,606
Length (km)	421	410	205

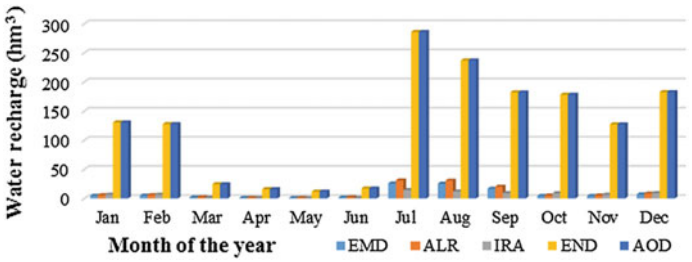


Fig. 9 Natural recharge for dams (CONAGUA 2013)

Table 2 Main characteristics for the dams involved in the case study

Characteristic	Abelardo L. R.	El Novillo	El Molinito	Álvaro Obregón	Ignacio R. A.
Capacity (hm^3)	2110.5	2833.10	130.04	29810.20	703.38
Net volume (hm^3)	0.01	1851.83	7.16	2134.96	517.87

and Yaqui River region). These regions match with the watersheds in the State. Table 1 shows the main data for these streams (CONAGUA 2013).

Along the Sonora River are located the Abelardo L. R. and El Molinito dams. In the Matape River is located the Ignacio R. Alatorre dam. While in the trajectory of the Yaqui River are located the Alvaro Obregon and El Novillo dams. Figure 9 shows the profile of the natural recharge for the dams involved in the case study. Where, most of the natural recharge is increased in the months of July to February, mainly in the region of the Yaqui River. The main data for these dams are shown in Table 2 (CONAGUA 2013).

3.5 Aquifers

The most important aquifers of the State are on the Pacific Coastal Plain (Costa de Hermosillo, Guaymas, and Valle del Yaqui). The aquifer Costa de Hermosillo is located in the irrigation district Costa de Hermosillo 05. It has a total recharge of 250 hm^3/year. It has 527 hm^3/year of extractions by pumping, giving a deficit of 277 hm^3/year. The condition is that the aquifer is overexploited. The aquifer of Guaymas is located in the irrigation district Guaymas 084. It has an annual recharge of 72.66 hm^3/year. It has 910.24 hm^3/year of extractions by pumping, giving a deficit of 26.58 hm^3/year. The aquifer of Valle de Yaqui is located in the irrigation districts of Yaqui River 041 and Colonias Yaquis 018. It has an annual recharge of 415.50 hm^3. It has 354.00 hm^3/year of extractions by pumping, giving a total storage of 61.50 hm^3/year. Figure 10 shows the natural recharge of the aquifers through the year.

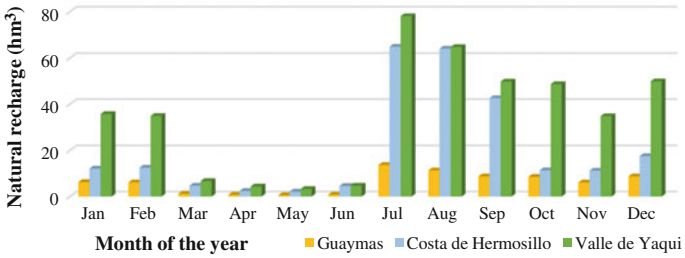


Fig. 10 Natural recharge for aquifers (CONAGUA 2013)

3.6 Existing Power Plants and New Power-Desalination Plants

In the State of Sonora, there are two thermal power plants that are the major sources of electricity in the Pacific north region of Mexico. The first one is the power plant "Puerto Libertad," located at 194 km from the city of Hermosillo. The power plant currently has a total installed capacity of 632,000 kW. It is able to provide a daily capacity of 15,168 kWh (SENER 2014). The second power plant is the "Guaymas II," which is located at 130 km at south of Hermosillo, and 108 km at north-west of the city of Obregon. The plant has a total installed capacity of 484,000 kW (SENER 2014). In the case study, there are considered two possible locations to install a new power-desalination plant, the Costa de Hermosillo beach and the Valle de Yaqui Beach. The configuration consists of a gas turbine, a heat recovery steam generator; an auxiliary boiler and a seawater reverse osmosis desalination plant. Table 3 shows the main data for the optimization analysis.

Table 3 Parameters for optimization

Concept	Value
Maximum capacity of storage tanks (m^3) (Nápoles-Rivera et al. 2013)	50,000
Fixed cost of storage tanks ($) (Nápoles-Rivera et al. 2013)	20,000
Variable cost of storage tanks ($/$m^3$) (Nápoles-Rivera et al. 2013)	17.5
Unitary price of electricity (Domestic users) ($/kWh)	0.16
Unitary price of electricity (Industrial users) ($/kWh)	0.32
Unitary price of electricity (Agricultural users) ($/kWh)	0.23
Unitary price of water (Domestic users) ($/m3)	1.07
Unitary price of water (Industrial users) ($/m3)	1.31
Unitary price of water (Agricultural users) ($/m3)	0.63
Specific capital cost of power generation ($/kW) (Agashichev and El-Nashar 2005)	500

(continued)

Table 3 (continued)

Concept	Value
Specific O&M cost of power generation ($/kW) (Agashichev and El-Nashar 2005)	0.1
Maximum capacity of power generation (MW)	1500
Specific capital cost of reverse osmosis ($/m^3/day) (Agashichev and El-Nashar 2005)	1000
Specific O&M cost of reverse osmosis ($/m^3/day) (Agashichev and El-Nashar 2005)	120
Power/Desalinated water supply ratio (MW/m^3/day) (Atilhan et al. 2012)	4.4×10^{-3}
Rejected water from desalination, β (Ladewig and Asquith 2013)	0.4
Overall efficiency for pump and motor efficiency, η (Swamee and Sharma 2008)	0.8
Hours of operation per year, H_Y (h/year)	8760
Factor used to annualize the inversion, k_F (year^{-1})	0.239

4 Optimization Results

The proposed optimization formulation was coded in the GAMS software (Brooke et al. 2015). The model consists of 9,523 equations, 9,903 continuous variables and 719 binary variables. The associated mixed-integer nonlinear programming (MINLP) problem was solved using the solver DICOPT in conjunction with the solvers CONOPT and CPLEX, with an average CPU time of 15,350 s in a computer with an i7 Intel core processor at 3.20 GHz and 32 GB of RAM. In addition, the model was tested using global solvers (BARON and LINDO GLOBAL) and the same results were obtained. Notice that the nonlinearities are found mainly in the installation cost equations.

Figure 11 shows the optimal solution where the electricity demands are fully satisfied. A new power-desalination plant is installed in the Hermosillo beach at the first time period with a total capacity of 1219 MW. In the city of "Hermosillo," the domestic water demand is satisfied with water from the "El Molinito" and "Abelardo L. Rodríguez" dams. The industrial water demand is satisfied with water from the "Abelardo L. Rodríguez" dam and by extracted water from deep wells in the area. The water in the irrigation district 051 for agricultural users is obtained by direct extraction from the "Costa de Hermosillo" aquifer. The level of water in the aquifer "Costa de Hermosillo" is benefited by recharging from the new power-desalination plant using injection wells. An increase in the water reserves of aquifers is shown, even for those that currently are overexploited, this is because of the installation of a new power-desalination plant. In the city of "Obregon," the industrial and domestic demands are satisfied by the "Álvaro Obregón" dam. In this case, "El Novillo" and "Álvaro Obregón" dams contain enough water to satisfy both demands. As mentioned earlier, both dams belong to the same water concession, this means that the flow from the "El Novillo" dam can recharge the "Álvaro Obregón" dam. Also, the volume of water contained in both dams is able

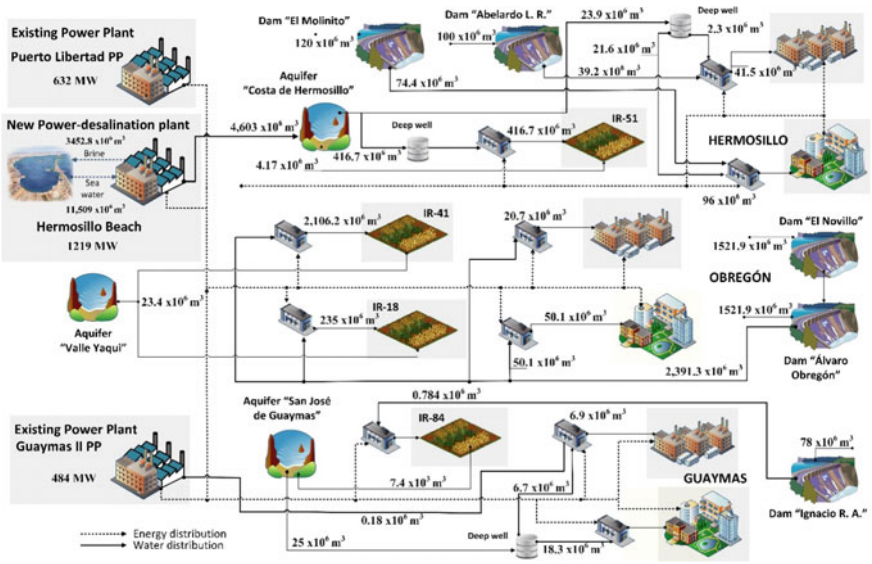


Fig. 11 Optimal solution for the case study

of supplying the water demand in the irrigation district 018 "Colonias Yaquis" and the irrigation district 041 "Valle Yaqui." In the city of "Guaymas," the domestic demands are satisfied with water from deep wells of the region. The industrial water demands are supplied with water from deep wells and water generated in the power-desalination plant of "Guaymas II." Agricultural water demands are satisfied with water from the "Ignacio R. Alatorre" dam. The "Guaymas" aquifer and "Ignacio L. Alatorre" slightly increase their volume due to the natural recharge.

The total annual cost in this case is 2,956 $MM/year, with pumping cost of 197 $MM/year and piping cost of 2.6 $MM/year. The installation cost of the new power-desalination plant is 106 $MM/year. The operation cost for the existing power plant is 0.25 $MM/year. The operation cost of the new power-desalination plant is the main contribution to the total annual cost with 2,651 $MM/year. The total water sales generate 2,004 $MM/year, which represents 63 % of total revenues. The total energy sales generate 1,199 $MM/year, which represents 37 % of total revenues, giving a total annual profit of 246.15 $MM/year.

5 Concluding Remarks

This chapter has presented a mathematical programing model for the optimal design of water distribution networks in macroscopic systems involving power and desalination processes. The proposed model seeks to satisfy the electricity and water demands and simultaneously considers the artificial recharge for overexploited

resources. Also, the water demand can be satisfied by extracting water from natural resources (e.g., dams, deep wells, storage tanks, rivers, and aquifers). The results have shown that the designed networks yield attractive solutions from the economic and sustainability points of view. Additionally, the results have shown that the total annual cost is mainly influenced by the operating cost of the new power-desalination plants and the pumping costs. Furthermore, attractive solutions have been identified, where interesting economic profits can be obtained and the recharge of the aquifers can be achieved. Finally, the proposed optimization formulation is general and it can be applied to other case studies.

References

Agashichev S. P., El-Nashar A. M., Systemic approach for techno-economic evaluation of triple hybrid (RO, MSF and power-generation) scheme including accounting of CO_2 emission. Energy. 2005; 30: 1283–1303.

Almansoori A., Saif Y., Structural optimization of osmosis processes for water and power production in desalination applications. Desalination. 2014; 344: 12–27.

Atilhan S., Linke P., Abdel-Wahab A., A systems integration approach to the design of regional water desalination and supply networks. International Journal of Process Systems Engineering. 2011; 1(2):125–135.

Atilhan S., Mahfouz A., Batchelor B., Linke P., Abdel-Wahab A., Nápoles-Rivera F., Jiménez-Gutiérrez A., El-Halwagi M. M., A systems-integration approach to the optimization of macroscopic water desalination and distribution networks: a general framework applied to Qatar's water resources. Clean Technologies and Environmental Policy. 2012; 14 (2): 161–171.

Averyt, K., Fisher J., Huber-Lee A., Lewis A., Macknick J., Madden N., Rogers J., and Tellinghuisen S., (2011) Freshwater use by U.S. power plants: Electricity's thirst for a precious resource. A report of the Energy and Water in a Warming World initiative. Cambridge, MA: UCS. 2011.

Brooke A., Kendrick D., Meeruas A., Raman R., GAMS-language guide. Washington DC: GAMS Development Corporation, 2015.

CEA, 2008. Sonora State Water Commission. Water Statistics in the State of Sonora Edition 2008.

CONAGUA, 2013. National Water Commission. Statistics on water in Mexico, 2013.

CONAGUA, 2014. National Water Commission. Agricultural Statistics of Irrigation Districts, 2014.

DOE, 2014. Department of Energy United States of America. The Water-Energy Nexus, 2014.

Esfahani I. J., Yoo C.-K., Feasibility study and performance assessment for the integration of a steam-injected gas turbine and thermal desalination system. Desalination. 2014; 332: 18–32.

Georgescu S.-C., Georgescu A.-M., Application of HBMOA to pumping stations scheduling for a water distribution network with multiple tanks. Procedia Engineering. 2014; 70: 715–723.

Ladewig B., Asquith B., Desalination Concentrate Management: Characteristics of Membrane Concentrate. Springer. 2013; 2: 5–15.

Liu S., Konstantopoulou F., Gikas P., Papageorgiou L. G., A mixed integer optimization approach for integrated water resources management. Computers and Chemical Engineering.2011; 35 (5): 858–875.

Najafi B., Shirazi A., Aminyavari M., Rinaldi F., Taylor R. A., Exergetic, economic and environmental analyses and multi-objective optimization of an SOFC-gas turbine hybrid cycle coupled with an MSF desalination system. Desalination. 2014; 334: 46–510.

Nápoles-Rivera F., Serna-González M., El-Halwagi M. M., Ponce-Ortega J. M., Sustainable water management for macroscopic systems. Journal of Cleaner Production. 2013; 47: 102–117.

Nápoles-Rivera F., Rojas-Torres M. G., Ponce-Ortega J. M., Serna-González M., El-Halwagi M. M., Optimal design of macroscopic water networks under parametric uncertainty. Journal of Cleaner Production, 2015; 88: 172–184.

SEMARNAT, 2012. National Council of Environment and Natural Resources. Report of the Environment in Mexico Edition 2012.

SENER, 2014. Secretary of Energy. Statistics and Indicators of the Electricity Sector.

Swamee P. K., Sharma, A. K., Design of Water Supply Pipe Networks: Cost considerations. Wiley. 2008; 4: 79–95.

UN-Water, (2014) WWAP (United Nations World Water Assessment Programme). 2014. The United Nations World Water Development Report 2014: Water and Energy. ISBN 978-92-3-104259-1. Paris, UNESCO.

Zheng F., Zecchin A., An efficient decomposition and dual-stage multi-objective optimization method for water distribution systems with multiple supply sources. Environmental Modelling and Software. 2014; 55: 143–155.

Part IV
Operations

Retrofit of Total Site Heat Exchanger Networks by Mathematical Programming Approach

Lidija Čuček and Zdravko Kravanja

We can create a more sustainable, cleaner and safer world by making wiser energy choices.

Robert Alan Silverstein

Abstract Heat Integration and Total Site Integration can provide considerable energy savings and emission reduction within industrial plants and entire Total Sites. Several approaches have been developed, such as an approach based on physical insights—Pinch Analysis, an approach based on Mathematical Programming, and recently combined or hybrid approaches based on both Mathematical Programming and Pinch Analysis. There are two designs, grassroots design for new process plants and Total Sites and retrofit design for existing plants and Total Sites. Tools for grassroots designs usually differ from those for retrofits, and retrofit analysis is considerably more difficult as the existing layout should be taken into consideration. This chapter provides an overview of the recent developments in Total Site Integration focusing mainly on the retrofits of existing Total Sites by Mathematical Programming approach and hybrid approaches. First, the Total Site Integration, approaches for Total Site Integration and software tools for Total Site Integration are introduced, followed by the retrofitting of existing heat exchangers within Total Sites. Furthermore, two illustrative examples are shown. Finally, the concluding remarks and sources of further information are provided.
Suggested citation Čuček L., Kravanja Z., 2016, Retrofit of Total Site Heat Exchanger Networks by Mathematical Programming Approach, In: M. Martín (Editor), Alternative Energy Sources and Technologies: Chapter 11, Springer International Publishing Switzerland, doi: 10.1007/978-3-319-28752-2_11

L. Čuček · Z. Kravanja (✉)
Faculty of Chemistry and Chemical Engineering, University of Maribor,
Smetanova Ulica 17, SI-2000 Maribor, Slovenia
e-mail: zdravko.kravanja@um.si

L. Čuček
e-mail: lidija.cucek@um.si

M. Martín (ed.), *Alternative Energy Sources and Technologies*,
DOI 10.1007/978-3-319-28752-2_11

Keywords Energy efficiency · Energy savings · Fixed and flexible designs · Heat exchanger network · Heat integration · Hybrid approach · Mathematical programming · Retrofit · Total site · Total site integration

Abbreviations

CC	Composite Curve
CHP	Combined Heat and Power
GCC	Grand Composite Curve
HE	Heat Exchange
HEN	Heat Exchanger Network
LIES	Locally Integrated Energy Sectors
LP	Linear Programming
LPS	Low Pressure Steam
MER	Minimum Energy Requirement or Maximum Energy Recovery
MILP	Mixed Integer Linear Programming
MINLP	Mixed Integer Nonlinear Programming
MP	Mathematical Programming
MP/PA	Mathematical Programming/Pinch Analysis
MPS	Medium Pressure Steam
NLP	Nonlinear Programming
NPV	Net Present Value
PA	Pinch Analysis
RAM	Reliability, Availability and Maintenance
TAC	Total Annual Cost
TS	Total Site
TSP	Total Site Profile

1 Introduction

Heat and Total Site Integration, and waste heat utilisation within industrial processes and Total Sites are efficient ways of conserving energy and achieving emission reductions. Heat Integration across plants is an extension of conventional Heat Integration within a single plant for further improving energy efficiency (Wang et al. 2015). Currently, energy consumption by both the industrial and residential sectors worldwide is very high (Boldyryev et al. 2013) and is still growing. The industrial sector has the largest share of total energy consumption amongst all the sectors (US EIA 2014). The industrial sector has high potential for reducing energy consumption. It has been estimated that most industrial plants throughout the world are namely still using up to 50 % more energy than necessary (Alfa Laval 2011). Typical savings of fuel within various industrial sectors can be shown in, e.g. CETC-Varennes (2003). On average, Total Site Integration shows potential for 20–25 % of additional energy savings (Hackl et al. 2011).

Energy is predominantly supplied by fossil fuels, whilst renewables provided about 10 % of the world's energy needs in 2014 (BP 2015).[1] Global energy consumption has increased by 2.1 % over the last 10 years on average (BP 2015). Environmental footprints have also been rising considerably (Čuček et al. 2015a). Energy saving, energy efficiency, renewable energy and environmental protection have become top priorities in several parts of the world. Significant efforts are being spent on increasing the share of renewables and on increasing energy efficiency.

Several approaches can be used for identifying the energy efficiency options and thus conserving energy and achieving emission reduction within process plants and Total Sites. In general they are divided into an approach based on heuristics ('rule of thumbs'), an approach based on thermodynamics and physical insights—Pinch Analysis (PA), and an approach based on numerical optimisation—Mathematical Programming (MP) (Klemeš and Kravanja 2013). Over recent years combined PA/MP and MP/PA approaches have also been developed for overcoming the drawbacks of both approaches.

Two situations exist regarding the analyses of heat exchanger networks (HENs) within processes and Total Sites: (i) grassroots or minimum energy requirement/maximum energy recuperation (MER) designs for new plants and Total Sites, and (ii) retrofit (also reconstruction, revamp or redesign) for existing plants and Total Sites. It should be noted that new grassroots plants and Total Sites should (and could) be designed in such a way to consider sustainability. MER designs have to be refined just by removing small heat exchangers with small heat transfer due to reduced cost and network complexity (Gundersen 2013a). Existing plants and Total Sites on the other hand should be retrofitted for improved sustainability (Pereira et al. 2012). Considerably more engineering time in the process industries is spent on improving existing production facilities than on designing new plants (Gundersen 2013a). Also, the tools for grassroots design cannot be readily used for the retrofit case (Gundersen 2013a).

There is a vast amount of literature available describing tools and methods for designing new plants—grassroots design, and much less sources are available relating to the improvements of existing plants and even less on retrofitting existing Total Sites. Performing retrofit analysis within existing process plants and Total Sites is namely considerably more difficult than for grassroots networks (Yee and Grossmann 1991) due to several reasons: (i) the existing equipment and layout must be considered, (ii) plant downtime is required which can be critical and (iii) existing heat exchange (HE) units should be more accurately modelled (Klemeš et al. 2013a). There are also many ways of improving the existing designs, such as changes in the uses of utilities, topological modifications, installing of additional areas, repiping of streams, reassignments of matches and heat transfer enhancement (Wang et al. 2012).

[1]Primary energy needs comprise commercially traded fuels. Renewables include also hydroelectric power generation.

This chapter focuses on retrofits of existing plants and Total Sites and mostly on recent developments. Only the MP approach to Total Site Integration is presented in this chapter. For the identification of targets in the case of retrofits it is also combined with the PA approach due to the transparency and physical insights of the obtained solution.

It should be noted that Heat Integration and Total Site Integration using the PA approach and the synthesis of new HENs at process and Total Site levels are beyond the scope of this chapter. Most of the sources available have focused on the synthesis of new HENs and Heat and Total Site Integration using PA as the main methodology. The reader is referred to the sources of further information provided in Sect. 8.

2 Total Site Integration

Total Site Heat Integration is a methodology for the integration of heat recovery among multiple processes and/or plants interconnected by common utilities on a site (Liew et al. 2012). Industrial processes rarely operate in isolation and are usually organised in larger areas termed Total Sites. The Total Site, also sometimes termed 'sidewide integration' is a set of processes linked together through a central utility. A site utility system consumes fuel centrally and supplies the necessary steam to the process consumers. It supplies the site processes with hot and cold utilities and power. Processes represent different sectors, such as industrial, residential, service, business and agricultural sectors (see Fig. 1).

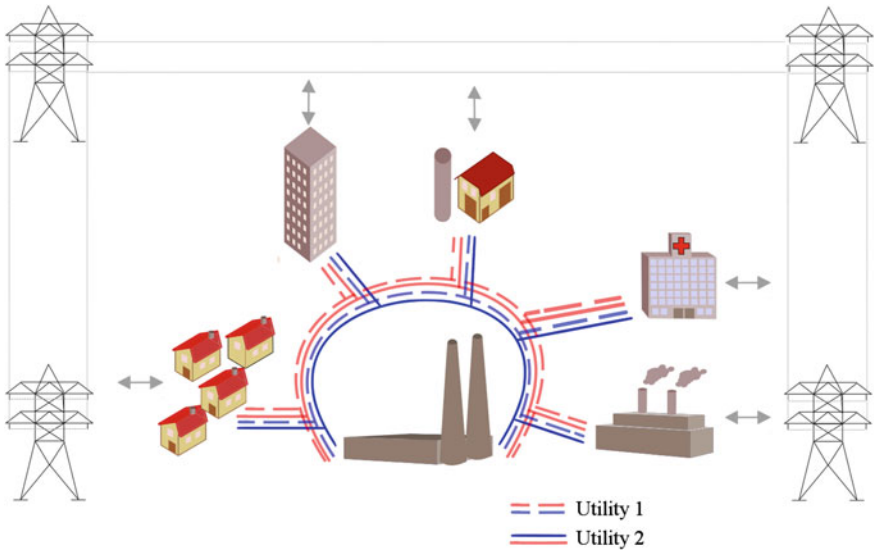


Fig. 1 Scheme of a Total Site

Total Site Integration is composed of two major components, heat recovery and power cogeneration using the utility system. This chapter however only describes Total Site Integration for increased heat recovery, for power cogeneration potential, see, e.g. Bandyopadhyay et al. (2010).

The Total Site concept describing heat integration across processes was first introduced by Dhole and Linnhoff (1993) and extended by Klemeš et al. (1997). Total Site has been defined as a collection of industrial processes only, and connected by a common utility system supplied by fossil fuels (Varbanov 2013). A broader Total Site concept integrates other sectors besides industrial ones, such as residential, service, business and agricultural sectors, termed as Locally Integrated Energy Sectors—LIES (Perry et al. 2008). Furthermore, Total Site methodology has been extended to incorporate renewable energy sources accounting for variability by supply and demand sides (Varbanov and Klemeš 2011) by introducing Time Slices into the Total Site and a heat storage system for accommodating the variations. Figure 2 shows an extended Total Site incorporating various non-renewable and renewable energy sources. Total Site Integration deals even with the surroundings of the Total Site in the cases of district heating and cooling (Gundersen 2013a). The scope of Total Site Integration has also been extended recently to the regional level (Čuček et al. 2013).

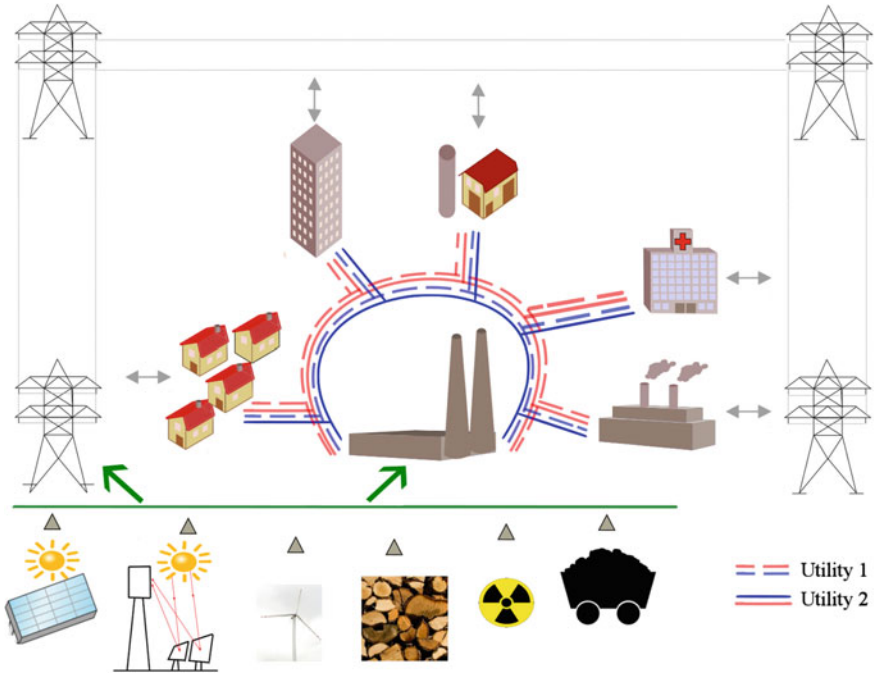


Fig. 2 Locally Integrated Energy Sectors incorporating non-renewable and renewable energy sources

There are two different strategies for Total Site synthesis: (i) sequential, when Heat Integration is performed first within a process and after between the processes (at the Total Site level) and (ii) simultaneous, where the Heat Integration is performed simultaneously within and between the processes. The simultaneous strategy can reveal additional opportunities for heat recovery that might not be obtained when performing sequential strategy (Nemet 2015).

Direct and indirect heat transfer inside and between processes could occur. Direct integration is performed using process streams, and indirect using intermediate fluid circuits such as steam, hot oil, etc. Direct integration can be performed by transporting either hot or cold process stream of one process to a heat exchanger placed in the other process, undertake heat exchange and afterwards return them to the original process (Nemet et al. 2015c). Direct integration offers maximum energy benefits in the cases of short distances between plants (Bandyopadhyay et al. 2010) and requires less heat transfer area in those cases where the number of loops is small. Rodera and Bagajewicz (2001) e.g. found that direct Total Site Integration can achieve more energy savings in specific cases than indirect Total Site Integration. Direct heat transfer requires more heat exchange loops indicating a higher capital cost compared with indirect integration (Wang et al. 2015). The indirect heat transfer on the other hand requires a more complex network but due to safety and operational concerns it might be preferred. Also, indirect integration is preferred when the distances between plants are long and there are large number of heat exchange loops (Wang et al. 2013).

There have been several successful Total Site Integration studies and/or implementations involving large petrochemical sites, such as BP UK, BP US and MOL Hungary (Chew et al. 2013), Kashima industrial Site, one of the biggest heavy chemical complexes in Japan (Matsuda et al. 2009), Swedish chemical cluster involving 360 process streams in Total Site Analysis (Hackl et al. 2011), large-scale steel plant with 8 Mt/y of crude steel (Matsuda et al. 2012), etc. Besides retrofit analyses of large-scale sites, analyses have also been performed for medium-sized and small-sized Total Sites. Such examples are medium-size refineries using PA (Nemet et al. 2015a) and the hybrid approach combining MP and PA under flexible design with considering more than 100 HE units (Čuček et al. 2015b), within a small community considering industry, service sector and residential area (Kostevšek et al. 2015), hospital complex including the general hospital, a regional laundry centre, a sports centre, and other buildings (Herrera et al. 2003), food producing company (Vujanović et al. 2015), etc.

However, despite many successful industrial Total Site Integration studies, there still remained several issues that remain unresolved and are of vital importance for the industries, such as considering design (plant layout, pressure drop), operation (start-up, shutdown), reliability, availability and maintenance (RAM), and regulation and policy (Chew et al. 2013).

3 Approaches for Total Site Integration

This section briefly presents systematic approaches for Total Site Integration: Pinch Technology or PA is based on information about the thermodynamic states of process streams, MP (deterministic and stochastic) which is usually an equation-oriented approach, and hybrid approaches combining both insights from the PA and MP. Besides those approaches, there are other lesser known tools which could be used for Heat and Total Site integration, such as P-graph (Friedler et al. 1992), which is capable of addressing nominal and flexible HEN designs (Friedler et al. 2009), however it lacks the demonstrative examples.

3.1 Pinch Technology

Pinch Technology or PA approach has been the more widely accepted concept both in academia and in industry (Klemeš et al. 2013a). It has the most important feature of obtaining performance targets ahead of design. The most typically used approach for Total Site heat recovery using PA is the graphical approach, however recently a numerical procedure has also been developed known as the Total Site Problem Table Algorithm (Liew et al. 2012). Total Site targeting methodology using PA includes data extraction methods, construction of Total Site Profiles (TSPs), Total Site Composite Curves (Total Site CCs) and the Site Utility Grand Composite Curves (Site Utility GCCs) (Klemeš et al. 2014).

TSPs are constructed for analysing the heat recovery potential through the central utility system. The heat can be recovered by first transporting heat to an intermediate utility and then by second heat transfer to different processes' heat demands. The Sink Site Profile and Source Site Profiles are constructed from the segments of the processes' individual GCCs (see, e.g. Klemeš et al. 2014). The segments with heat demand are presented in the Sink Site Profile, and the Source Site Profile is an indicator for streams with heat surpluses. When constructing TSPs the segments' temperatures should be shifted in order to obtain feasible heat transfer. Also the source segments should be rotated by constructing the Source Profile on the left-hand side of the TSP.

In the case of indirect heat transfer, the heat recovery can be maximised based on the constructed TSPs by selecting a proper temperature level of the intermediate utility, at which the maximal amount of available heat from the Source Site Profile is recovered to the Sink Site Profile.

Indirect Total Site Integration could be performed using sequential strategy, where the Heat Integration is first performed within each process and the remaining heating and cooling demands are integrated on a Site level (pockets are removed before Site Integration). Another option is possible using simultaneous strategy where the integration is performed simultaneously within and between the processes (in that case the pockets are not removed).

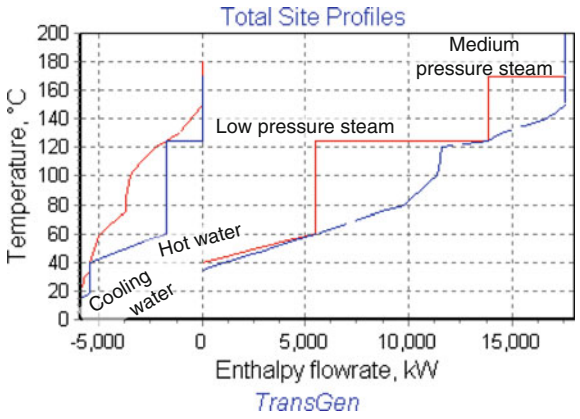


Fig. 3 Total Site Profiles with levels of intermediate utilities (modified from Čuček and Kravanja 2014)

Figure 3 illustrates TSPs using an example of LIES (modified from Perry et al. 2008). This LIES consists of two industrial plants, a hospital complex and residential dwellings (Perry et al. 2008). The Site Source and Site Sink Profiles are presented in red and blue dashed lines, and the Source and Sink CCs in blue and red solid lines. Without Total Site Integration, around 5.9 MW would need to be supplied by cold utilities and around 17.6 MW by hot utilities. When performing indirect Total Site Integrations there are significant opportunities for energy savings, as can be seen from Fig. 3. Around 5.4 MW of intermediate utilities (3.7 MW of hot water and 1.7 MW of low pressure steam—LPS) can be produced at the Source Side and transferred to the Sink Side, thereby reducing the consumptions of hot and cold utilities by the same amount. Possible savings accounting for Total Site Integration are significant in several cases.

The presented approach can be applied for producing targets for the heating and cooling demands (MER designs). In order to obtain MER design the following rules should be followed: (i) Heat should not be transferred across the pinch, (ii) Cold utilities should not be used above the pinch and (iii) Hot utilities should not be used below the pinch.

The presented approach is well suited for designing new plants—for grassroots design. It should be noted that usually design evolution, also termed energy relaxation, might be required by removing small heat exchangers (Gundersen 2013a). The approaches for retrofit analysis are quite different from those for grassroots analysis as the economics, safety and other constraints (distance, start-up etc.) should be satisfied and the existing equipment should be utilised as much as possible. Optimal HEN after retrofit is likely to be quite different from the optimal grassroots design (Gundersen 2000).

HEN retrofitting analysis using PA can be generally performed using several approaches (Kemp 2007): (i) Developing a MER design as for a new plant but favouring the matches which already exist in the current network, (ii) Start with the existing network and work towards an MER design, and (iii) Start with the existing

network and identify the most critical changes that give the most substantial energy reduction. A key insight of this approach is the Network Pinch (Asante and Zhu 1996).

The more commonly used approach for retrofits is the Network Pinch approach (Asante and Zhu 1996) and its extensions (e.g. Smith et al. 2010a). Network Pinch presents the bottleneck matches in existing HENs that should be overcome for increased heat recovery (Nie and Zhu 1999). This approach however combines physical insights and MP techniques to retrofit problems (Smith et al. 2010a).

3.2 Mathematical Programming

MP involves the formulation of a mathematical model and its solution using optimisation methods (Sreepathi and Rangaiah 2014). MP problems are more generally formulated as Mixed Integer Nonlinear Programming (MINLP) problems. In order to avoid being trapped within local solutions, simplifications are made in many cases to convert MINLP into Linear Programming problems (LP), Mixed Integer Linear Programming (MILP) or Nonlinear Programming (NLP) problems (Liu et al. 2014). The most successful version has been proposed in the past in the form of a MILP-NLP approach (Tantimuratha et al. 2000). In the continuation in Sect. 5 the MILP-MINLP approach is presented in more detail regarding the retrofitting of existing HENs within Total Sites.

Optimisation methods can be divided into deterministic and stochastic (probabilistic) methods (Klemeš et al. 2010). Deterministic methods refer to optimisation models where all the parameters have known constants. If the parameters are specified as uncertain quantities the values of which are characterised by probability distributions, the optimisation model is stochastic (Bradley et al. 1977).

Deterministic methods check a large number of possible optimum solutions by enumerating the combinations of design variables and finally find the best one. They could be based on insights from PA; and also represent sequential or simultaneous approaches to the design problems with considerably better capabilities than PA to handle the complex trade-offs, and they can easily handle issues such as forbidden matches (Gundersen 2013a). The major limitations of deterministic methods are numerical problems related to non-convexity (local optima) and computational complexity (combinatorial explosion).

Stochastic optimisation methods overcome the numerical problems and have the advantage of having more chances to obtain globally optimal solutions for NLP and MINLP problems due to the random natures of these kinds of optimisation methods (Liu et al. 2014). The commonly used algorithms in the synthesis of HEN are Genetic Algorithms and Simulated Annealing (Smith et al. 2010b). However, the drawbacks of stochastic methods are that they do not utilise insights from PA and do not guarantee an optimum solution (Gundersen 2013a). Stochastic methods are also non-rigorous and the quality of the solution depends on the time spent on the search (Anantharaman 2011).

Mathematical modelling methods have the advantage of solving large-scale problems (Smith et al. 2010a). Also they can properly and easily handle those situations where there are constrained matches, such as forbidden, required or restricted matches due to product purity and possible contaminations of streams, long distances, operability issues such as control, safety and start-up, and restricted matches. In addition, by retrofitting most of the existing HE matches remain unchanged. Furthermore, they can handle the complex trade-offs between operating and investment costs under fixed or steady-state operating conditions and under flexible conditions (Čuček et al. 2015b).

There are several types of MP formulations that could address Heat and Total Site Integration. The better-known and applied formulations are: the transportation problem (Cerda et al. 1983), transhipment problem (Papoulias and Grossmann 1983), method based on a pinch point location (Duran and Grossmann 1986) and a stagewise problem (Yee and Grossmann 1990).

The transportation problem was the first MP formulation that allowed the considering of multiple utilities and constrained (forbidden, required or restricted) matches (Grossmann 1992). The transportation problem enables the obtaining of minimum utility usage during HEN synthesis (Cerda et al. 1983). It considers all feasible links for heat exchange between each pair of hot and cold streams over their corresponding temperature intervals (Cerda and Westerberg 1983).

The transhipment and expanded transhipment models—the transhipment models considering constrained matches (Papoulias and Grossmann 1983)—are alternative models to the transportation problem. Heat is shipped from hot streams (sources) to cold streams (destinations) through intermediate temperature intervals (warehouses). Heat flows from hot streams to a temperature interval and then to cold streams within the same interval or as a residual heat to the next lower interval. The transhipment model is an equivalent to the calculation of the Problem Table and is an alternative representation of the Heat Cascade (Gundersen 2013a), which also forms the basis of the GCC and TSPs. The transhipment model is based on the assumption that the flow rates and temperatures have fixed values. Figure 4 illustrates an example of heat flow representation in a transhipment and an expanded transhipment model for two hot and two cold streams (modified from Biegler et al. 1997). The data for the example are taken from Smith (2005) from Fig. 16.2.

The transhipment model has been extended to a multiperiod model in order to account for different operating modes (Floudas and Grossmann 1986), for optimisation of steam levels of the Total Site (Shang and Kokossis 2004), for incorporating rules for obtaining optimistic and pessimistic scenarios by exploring the Plus–Minus principles (Čuček and Kravanja 2014), physical insights for existing and target designs for plants and Total Sites under varying operational conditions (Čuček et al. 2014), trade-offs between investment and operating costs and thus possibilities for proposing limited numbers of retrofitting solutions or all profitable solutions at process and Total Site levels (Čuček and Kravanja 2015b), etc.

Transportation and transshipment models have the advantages of being linear and thus the obtained solutions are globally optimal. They enable targeting of the

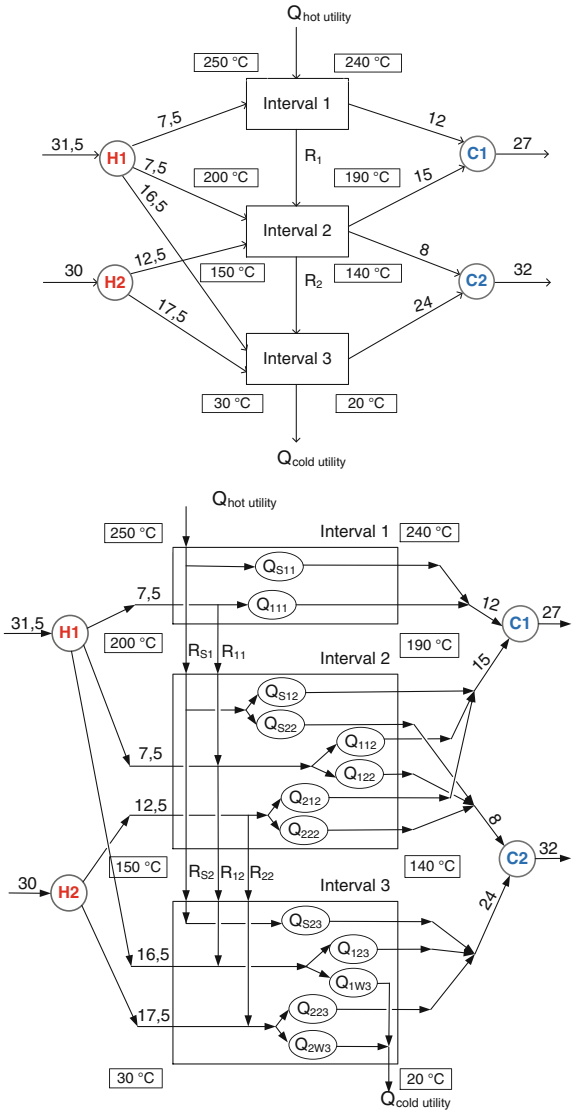


Fig. 4 Representation of heat flows in transhipment (*above*) and expanded transshipment model (*below*) (modified from Biegler et al. 1997)

utility costs and could be extended for predicting the fewest number of heat exchange units. The transhipment model leads to problems of smaller size compared to transportation problems (Grossmann 1992).

The disadvantage of the transportation and transhipment models are that they require the fixed values of the temperatures and flow rates. Thus, they are less applicable for the grassroots. Heat or Total Site Integration should be considered simultaneously with the optimisation in order to account for interactions between the process or Total Site and HEN, and thus for appropriate trade-offs between

investment, energy requirement and raw material consumption. However, transportation and transhipment models could well be applicable for the retrofits as the process or Total Site's conditions are known.

A model based on pinch location method (Duran and Grossmann 1986) has the advantage of performing simultaneous heat integration with process flowsheet optimisation. The minimum energy targets are obtained from the optimised process flowsheet. Heat integration can be performed under variable flow rates and temperatures, and multiple utilities could be considered. The main concept of the method is that it assumes each stream (stream's inlet temperature) as a potential pinch candidate. This method has the advantage of being easy to embed within any MP synthesis model for performing simultaneous flowsheet synthesis and heat or Total Site integration (Grossmann et al. 2000). It can also be successfully used in process simulators, e.g. FLOWTRAN (Lang et al. 1988) for simultaneous optimisation and heat integration.

The stagewise model (Yee and Grossmann 1990) is based on the MP for the synthesis of HENs. The model is based on a superstructure comprising all structural alternatives and involving combinatorial decisions, see Fig. 5 for a superstructure of two hot and two cold streams.

During each stage potential matches between each hot and cold stream could occur, and the temperature driving forces are optimisation variables. It can handle stream splitting/no splitting, constrained matches and multiple utilities. It can also consider variable inlet and outlet temperatures, and can be used for simultaneous process/Total Site optimisation and its HEN synthesis (Yee et al. 1990). However, the drawback is that this model is nonlinear programming hard and has complex combinatorics and non-convexities involved. It could be used only for the smaller problems. The lowest total annual cost is usually the main objective (Ahmetović and Kravanja 2013). There are several extensions of the model by accounting for pressure drops (Frausto-Hernández et al. 2003), pumping cost (Mizutani et al.

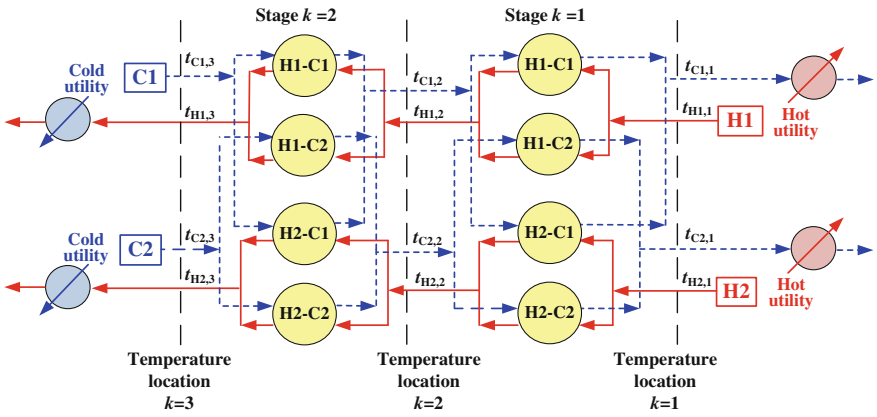


Fig. 5 Stagewise superstructure for two hot and two cold streams (modified from Yee and Grossmann 1990 and Klemeš and Kravanja 2013)

2003), different types of heat exchangers (Soršak and Kravanja 2002), retrofit (Soršak and Kravanja 2004), flexible HENs (Aaltola 2002), global optimisation (Bogataj and Kravanja 2012), multiperiod synthesis involving multiple utilities (Isafiade et al. 2015), multi-objective synthesis (López-Maldonado et al. 2011), multi-objective design of HENs with dimensionality reduction (Vaskan et al. 2012), future energy prices and entire lifespans (Nemet et al. 2013), Total Site synthesis (Nemet et al. 2015b), etc.

MP strategies could be classified as either sequential or simultaneous synthesis methods (Furman and Sahinidis 2002). Sequential synthesis methods use the strategy of dividing the HEN synthesis problem into a series of subproblems that are much easier to solve. Generally, this is achieved using the following three rules: (i) the minimum utility usage/cost, (ii) the minimum number of exchanger units and (iii) the minimum area/capital cost of the network. Rule 1 has precedence over Rule 2, and Rule 2 has precedence over Rule 3 (Biegler et al. 1997). PA is the better-known example of the sequential approach (Peng and Cui 2015). Examples of the sequential-based MP approach are the transportation and transhipment problems. The sequential approach has the disadvantage of inappropriate trade-offs between energy, number of units and area requirements. One of the major developments for automated sequential HEN synthesis was done by Floudas et al. (1986). HEN configuration features minimum investment costs subject to minimum utility costs and the fewest number of units. The sequential approach is still considered the more practical way to solve industrial-scale HEN problems (Chen et al. 2015).

The simultaneous synthesis approach optimises utility costs, selection of matches, and area costs simultaneously. Examples of the simultaneous mathematical approach include the stagewise superstructure (Yee and Grossmann 1990), simultaneous synthesis method using simulated annealing (Dolan et al. 1989), the interval-based MINLP superstructure model (Isafiade and Fraser 2008), etc. The disadvantage of the simultaneous approach is that it cannot be solved for the larger scale Total Sites even with the rapid advancement in computing power and optimisation technology. Namely, the size of the problem namely still does not meet the industrial needs (Anantharaman 2011).

3.3 Hybrid Approaches Combining Pinch Analysis and Mathematical Programming

Hybrid methods are those methods where two or all of the approaches (heuristics, PA and MP) are combined and taken advantage of the strengths of the individual techniques (Fraser 2013). Grossmann and Daichendt (1996) pointed out as one of the major challenges to be solved that different approaches (heuristic search, optimisation and targeting approaches) should be integrated.

Table 1 Main strengths of pinch analysis and mathematical programming

Pinch analysis	Mathematical programming
A systematic procedure	Optimality, feasibility and integrality of solutions (Kravanja 2010)
Based on physical insights	
Targeting before the design	Handle large number of optimisation variables
It guarantees an optimum solution	LPs and MILPs are robust
Consideration of physical laws	Systematic and potentially automatic even for large industrial problems under flexible designs (Čuček et al. 2015b)
Graphical and algorithmic	
Conceptually simple and easily understandable	
	Enable simultaneous process/Total Site optimisation with its Heat/Total Site Integration
High industrial acceptance	Easy consideration of multiple criteria
Proven energy savings using Pinch Technology	Rigorous optimisation of the structure, sizes of heat exchangers and utility usage
Could be applied to continuous and batch processes, for grassroots and retrofits	Easily handle constrained matches
Widely used in academia and in industry	Could automatically include the rules for obtaining optimistic and pessimistic scenarios by exploring the "Plus–Minus Principle" (Čuček and Kravanja 2014)
Widespread inclusion in chemical engineering education curricula (Klemeš et al. 2013b)	
A routine tool for advanced design and optimisation in various industries (Klemeš 2013a)	Synthesis of retrofitted HENs including positions and temperatures of heat exchanger units within networks under fixed and flexible designs
	Synthesis and retrofit can be performed in 'one single run'
Flexibility in integrated plants could be considered (Liew et al. 2012)	Certain number of most optimal retrofitting modifications could be proposed (Čuček et al. 2015b)
A mature technology (Kemp 2007)	HEN design could potentially be automated (Gundersen 2013b)

The use of graphical techniques has proved to be a very powerful tool for visualising the results of the optimisation and for proposing further energy efficiency improvements (Maréchal and Kalitventzeff 1998). MP by itself can consume a lot of computational time, and does not provide any insight into the problem. However, MP methods allow the HEN design procedure to be automated. Tables 1 and 2 present the main strengths and limitations of PA and MP approaches.

Klemeš and Kravanja (2013) presented possible opportunities for employing the hybrid approach combining both MP and PA. Those main opportunities are:

- When both approaches are employed, the searching space of MP is narrowed. Physical laws and optimality, feasibility and integrality of solutions are considered (Kravanja 2010);
- Combination of MP and PA can be very beneficial for data collection and verification, as the physical insights make the checking easier;
- PA is beneficial in the first step followed by MP for problem formulation, and can guide MP close to global optima.

There are several examples of successful targeting and retrofits analyses of existing Total Sites which are mainly based on hybrid approaches, at least to a small extent. Such examples are Network Pinch (Asante and Zhu 1997), Hypertargets

(Briones and Kokossis 1999), HEN synthesis using MINLP stagewise model with PA and relaxation (Angsutorn et al. 2014), heat integration based on two steps: conceptual using PA and an optimisation step using MINLP (Krajnc et al. 2006), synthesis of a utility system within Total Site by applying a combination of targeting and MP techniques (Varbanov et al. 2005), transhipment-based software tool TransGen including insights from the PA (Čuček and Kravanja 2014) and an improved version (Čuček and Kravanja 2015b), etc.

The methods which represented breakthrough in the Heat and Total Site Integration based on both PA and/or MP were developed in the 1980s and early 1990s, while after that no major breakthroughs have been made (Gundersen 2013b). Yet, there have been numerous publications on HENs both for grassroots and retrofit design during the last 20 plus years (Gundersen 2013a), and several of them are presented in this chapter as well. Recent investigations regarding the HEN design have focused mainly on the developments of procedures and methods for solving HENs with greater numbers of process streams, global optimisation of HEN, more detailed HEN designs, the synthesis of flexible HENs, multi-objective synthesis of HENs (Klemeš et al. 2013a), and adding industrial realism to the problem definition (Gundersen 2013b).

4 Software Tools for Total Site Integration

Total Site Integration problems are complex in terms of scale and relationship (Klemeš et al. 2013a). Therefore, software tools have been developed in order to solve complex problems of Process and Total Site Integration and other engineering fields (Bulatov 2013). The software tools developed are mostly based on a hybrid approach combining PA and MP. A software tools overview including process integration and retrofit analysis tools is provided in Lam et al. (2011). A comprehensive list of software tools for Process Integration, modelling and optimisation is provided in Klemeš et al. (2010b). An overview of software tools specifically suited for Heat Integration is presented by Bulatov (2013) in Handbook of Process Integration (PI), Minimisation of energy and water use, waste and emissions (Klemeš 2013b).

The main features of recently developed software tools applicable at Total Site level are described. It should be noted that this is probably not a comprehensive list of the software tools that exist worldwide. Only those software tools are described for which their developers have explicitly mentioned that software tools could be used for Total Site Integration.

i-Steam™ is the latest software tool from Process Integration Limited (PIL 2012). It has several features such as: (i) targeting and graphical analysis for defining the best energy performance for a site under a range of different operating scenarios. It includes targets for the full range of available utilities; (ii) Modelling and simulation of utility systems by including a comprehensive suite of unit operational models such as steam and gas turbines; (iii) Comprehensive modelling

Table 2 Main limitations of pinch analysis and mathematical programming

Pinch analysis	Mathematical programming
Design process is time-consuming, especially for large-scale problems and presents multiple alternatives (Bagajewicz et al. 2013)	Could be very complicated
	Difficulties in problem formulation
	It can miss some options
Sequential approach: (i) minimum utility targeting; (ii) network design for minimum number of units; (iii) use loops and paths to evolve (Bagajewicz et al. 2013)	NLPs and MINLPs are locally optimal and computationally extensive, stochastic optimisation may partially overcome those limitations
	Solutions are usually hard to interpret
	Requires advanced knowledge of Total Site Integration and MP
Capital investment is not straightforward in case of retrofit designs	For large-scale problems models are either rigorous and not solvable, or deficient and solvable (Björk and Nordman 2005)
Larger scale retrofit problems offer combinatorial challenge	
	Lack of easy-to-use software (Shenoy 1995)
It requires expertise to perform retrofit design	Possibility that solutions are wrong due to some mistakes in the model
Difficult to perform retrofit under varying operational conditions, especially at Total Site level	
More difficult to consider multiple criteria	

and simulation: Combined Heat and Power (CHP) + Steam pipeline + Processes which enables, e.g. detailed modelling of steam pipeline system; (iv) Optimisation of the steam system, (v) Driver selection: selection between motor and turbine for the power demands based on the steam system, (vi) Performance regression tool producing full and part-load performance models for boilers and steam turbines (PIL 2012). It is the successor of SITE-int (PIL 2012) and STAR (2015).

EFENIS-Site developed by the Centre for Process Integration at the School of Chemical Engineering and Analytical Science (CEAS) of The University of Manchester (CPI Manchester 2015) is a software package for the design of site utility and cogeneration systems. The interactions between the processes on the site and the steam system, steam turbines, gas turbines (with auxiliary firing options), boiler house, local fired heaters and cooling systems are all analysed using EFENIS-Site. Issues addressed by EFENIS-Site include: (i) Understanding site utility infrastructures; (ii) Optimising existing utility system configurations; (iii) Targeting cogeneration potential; (iv) Choosing the more appropriate cogeneration system; (v) Optimising site steam pressures and loads; (vi) Design and operation of steam turbine networks; (vii) Minimising energy costs for the site and (viii) Reducing flue gas emissions from the site (EFENIS, 2015a).

SuperTarget™ is a PA-based software tool from KBC Advanced Technologies (KBC 2014). It is used to improve Heat and Total Site Integration in new design and retrofit projects by reducing operating costs and optimally targeting capital investment. It enables the determining of energy targets, cogeneration potential and

optimal utility levels, and usage for sitewide heat and power analysis. It is a modular product with options for Process (energy optimisation within a single-process unit), Site (heat and power targeting across the entire manufacturing site) and Column (heat distribution in distillation columns) analysis (KBC 2014). It is also a tool for day-to-day application and makes PA a routine part of process design (Klemeš et al. 2010). It is still the preferred tool of energy Pinch experts worldwide (Bulatov 2013).

EnIgMa is an energy integration manager tool with generic interface developed at The Hamburg University of Technology (TUHH 2015). Currently, it is developed as a prototype. It directly links to the Data Interface which is linked to the Process Data, a Flowsheet Simulator, a HEN Optimisation Framework and a Project Data Bank. The prototype covers, amongst other features, the data extraction and data storage via interfaces, e.g. Aspen Plus and Excel, and implementation of an optimisation framework via an interface to e.g. SyntHex (XRG Simulation GmbH 2014).

TransGen is a software tool for the energy targeting and retrofitting of existing industrial processes and Total Sites operating under steady-state and dynamic conditions developed at the University of Maribor (Čuček and Kravanja 2015c). It is based on and significantly extended from an expanded transhipment model (Papoulias and Grossmann 1983) and combines the advantages of both MP and PA. It has several more significant features (Čuček and Kravanja 2015b) such as: (i) Enabling the performing of existing, target and modified (proposed) HEN designs; (ii) Insights from Pinch Technology; (iii) Data independence; (iv) Globally optimal solutions (MILP form); (v) Possibility for solving Total Site problems of any scale; (vi) Internal and direct/indirect Total Site Integration (Čuček and Kravanja 2014), (vii) Considering both investment and operating costs; (viii) Solutions can be obtained in 'real' time (Čuček and Kravanja 2015b).

HENSYN is a code for the detailed synthesis of the retrofitted HENs under steady-state and dynamic operating conditions developed at the University of Maribor (Čuček and Kravanja 2015b). It is based on and significantly extended from a stagewise model (Yee and Grossmann 1990). The more significant features are: (i) Obtaining the detailed syntheses regarding the trade-offs between investment and operating costs; (ii) Synthesis of modified HENs including positions and temperatures of heat exchanger units within networks under fixed and flexible designs and (iii) Accounting for different heat exchanger types (Soršak and Kravanja 2004).

5 Retrofitting of Existing Heat Exchanger Networks Within Total Site

In this section the methodology and procedure for the retrofitting of existing HENs within Total Sites is presented. By applying this approach presented in the following it is possible to retrofit existing HENs under steady-state or fixed operating conditions and under flexible dynamic operating conditions.

The presented approach is based on the sequential synthesis where in the first step targets and directly the potential for Total Site integration are obtained, in the second step the modifications from existing HENs are identified, and in the final step modified HEN is synthesised. The procedure is mostly based on the MP approach. The formulation for the first and second steps is based on MILP, on an extended transhipment model, and the formulation for the third step is based on MINLP, on an extended stagewise model. The first step also includes physical insights from the PA. The solutions obtained during the second and third steps consider the trade-offs between energy and capital costs.

A software tool TransGen was developed for this purpose. TransGen is used for first targeting and second identification steps. Further, code HENSYN is used for the third synthesis step. In order for the procedure to be automated, the data regarding the final streams that are included in HEN are automatically generated in TransGen and loaded into HENSYN.

The obtained solutions could be obtained 'by two runs' in reasonable time from less than 1 s for the problem with a few hot and cold streams to about 1 day for the large-scale problem with more than 100 hot and cold streams under few different operating conditions. The presented approach for performing the most optimal retrofitting modifications was proven to be promising for the industrial Total Site EFENIS (2015b). This approach could be used for the retrofit of Total Site of any scale but is especially advantageous for the retrofitting of large-scale industrial Total Sites under varying conditions.

It should be noted that it is always likely that even continuously operating chemical plants operate under somewhat varying operating conditions. This is due to changes in the feed flow rates, usage of different feedstocks, operating at various capacity levels, varying ambient temperatures, etc. Due to those variations, the data required for Total Site Integration (such as temperatures, flow rates, etc.) should be obtained for those varying operating conditions. Thus flexible HEN should be synthesised which should be feasible for all operating conditions, and ensuring safe production. In terms of modelling it is formulated as a multiperiod optimisation problem.

As it was introduced above, the sequential three-step procedure based on hybrid MP/PA approach and applied software tool TransGen and code HENSYN for the retrofitting of existing HENs within Total Sites consist of the following steps:

- Targeting step using the software tool TransGen based on MP/PA and identification of the potential for Heat and Total Site Integration and waste heat utilisation by comparing with the existing energy consumption;
- Identifications of alternatives for modifications based on MP using TransGen and selection of modifications by forbidding the unfeasible matches. This step enables the obtaining of the most optimal retrofitting modifications regarding energy consumption reduction and intermediate utilities production in regard to trade-offs between operating and investment costs. All the proposed modifications should be verified and the procedure is repeated as long as all the proposed heat exchange matches are acceptable. Several loops could be required to obtain verified and feasible results;
- Synthesis of the retrofitted HEN design from those modifications identified during the second step using code HENSYN. The synthesis step enables obtaining of the structure of the retrofitted part of HEN and the basic parameters for the heat exchange units involved.

Figure 6 illustrates the applied procedure.

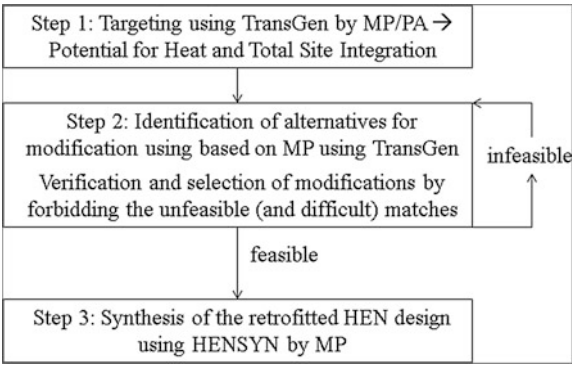


Fig. 6 Procedure for retrofitting of Total Sites using TransGen and HENSYN (modified from Čuček et al. 2015b)

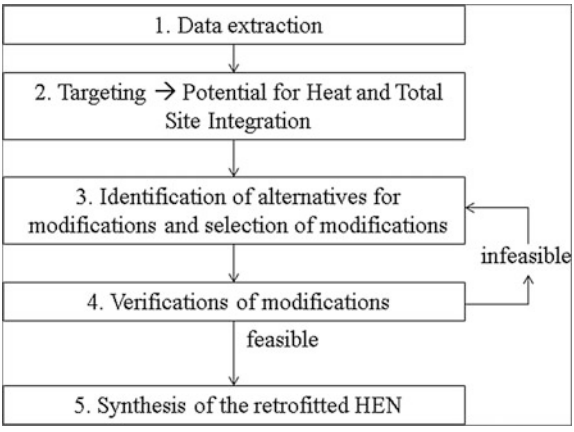


Fig. 7 Methodology for retrofitting industrial Total Sites (modified from Čuček and Kravanja 2015c)

The following methodology is used for Heat Integration and Total Site Optimisation:

1. Data extraction and incorporation of data into software tool TransGen;
2. Targeting using software tool TransGen and the comparison with the existing design and energy consumption;
3. Identification of alternatives for modifications and selection of modifications by forbidding the unfeasible matches. Identification and selection of modifications is performed using software tool TransGen;
4. Verifications of the obtained solutions from Step 3 and returning to Step 3 as long as all the proposed heat exchange matches are acceptable for the leaders and engineers of the plant and/or Total Site. Several loops could be required to obtain (near) optimal, verified and feasible results;
5. Synthesis of the retrofitted HEN using code HENSYN using the data for selected modifications.

The developed methodology is presented in Fig. 7 and demonstrated by illustrative and industrial examples.

5.1 Extraction of Data

Correct data extraction is the crucial issue in performing Total Site Integration (Klemeš 2013a). It is a very time-consuming and critical activity, as the quality and realism of the obtained design heavily depend on the used data (Gundersen 2013a). The data for retrofits are usually obtained from process simulations and directly from measurements, however in this case data reconciliation (Shenoy, 1995) is also required to adjust measured values and estimate non-measured streams (Vocciante et al. 2014).

There are several options (approaches) on how to extract the data according to the level of detail of the analysis and the availability of process data (Hackl et al. 2010). Those approaches are (Hackl et al. 2010):

- Black-box approach which is represented by utility demand only. Steam consumers excluded in the steam data could be handled by the black-box approach. The temperatures at which energy is required are unknown, and the energy requirements are represented by the temperature levels of the utility streams (Nguyen et al. 2014);
- Grey-box approach considers the process–utility interface and ignores process-to-process heat exchange. In this case the current level of Integration within single units is not changed and self-sufficient pockets of the GCCs are excluded (Nguyen et al. 2014). Total Site studies are usually based upon the grey-box approach (Nguyen et al. 2014);
- White-box approach, according to which a detailed Total Site analysis is carried out. The data include process-to-utility and process-to-process heat exchange, and thus all the heat requirements of the process streams are considered with

their appropriate temperature levels (Nguyen et al. 2014). The white-box approach has been selected for extracting the data in order to perform the detailed retrofit analysis using the methodology presented in Fig. 7.

Several data are required for the detailed target and retrofit analysis of the Total Site and should be obtained from the existing Total Site (from process simulations and/or from direct measurements). Those data are:

- For each hot stream (stream that is being cooled and/or condensed) and cold stream (stream that is being heated and/or evaporated): stream contents, state of aggregation and phase changes, supply and target temperatures, specific heat capacities, enthalpies of vaporisation or condensation, mass flow rates and estimated film heat transfer coefficients. In the case of flexible operation, which is modelled as a multiperiod problem, those data are required for each representative case;
- For hot and cold utilities: the pressures, temperatures, prices and heat transfer coefficients;
- For recovered intermediate utilities (such as hot water, steam, etc.) in the case where there is any possible and economic use of them: supply and return temperatures and the required quantities, and percentage of condensate or hot water return;
- For detailed analysis of the trade-offs between energy and capital cost: amortisation period for retrofit, discount rate, and if appropriate, the available/estimated money for retrofit investment or desirable payback period;
- For economic analysis related to modifications between the plants: the distances between plants within Total Sites;
- For the calculation of pipe cost: at least some reasonable estimates of densities and pressures of all hot and cold streams.

There are two possibilities regarding the extraction of data for detailed retrofitting analysis within existing Total Sites: (i) data based on composite streams and (ii) data based on segmented streams (data related to each existing HE unit). The obtained results when applying segmented data are more precise than in those cases when applying composite data (Čuček et al. 2014), and it is also considerably easier to perform data extraction with less (preferably without) pitfalls. However, the drawback of such data extraction is that larger amount of data are required.

Data extraction for composite streams should follow the basic principles of data extraction presented elsewhere, e.g. in Klemeš (2013a). However, it should be noted that it may not be a straightforward task. Typical pitfalls and suggestions on how to avoid them are described in Klemeš et al. (2010b), and for major challenges in data extraction see, e.g. Gundersen (2013a).

Data extraction for segmented streams considers the data relating to each HE unit (each hot and cold stream including utilities). It should be noted that for hot and cold utilities the net flow rates are unknown (El-Halwagi 2012), and are in terms of the modelling defined as variables (they should be found by energy targeting and retrofit modifications). The main advantage of having the extracted data for each

segmented stream is the possibility of performing the identification and selection of the more optimal retrofitting solutions by unfixing a certain number of existing HE matches. Also, there are some limitations within Total Sites that certain units should not be integrated because of economic and safety reasons, start-up, and due to system dynamics such as non-simultaneous operation (Kravanja and Čuček 2013). Such connections are defined as forbidden. Furthermore, all unfeasible and difficult modifications could be excluded within modified design just by forbidding those specific matches.

When extracting data for each segment of process streams it is possible to obtain existing, target and modified designs. Existing designs are obtained when all the matches are constrained, target designs are obtained when all the existing matches are unfixed or unconstrained, and modified designs with certain numbers of modifications are obtained by unfixing that number of existing matches.

5.2 *Targeting and Identification of Potential for Heat and Total Site Integration*

When having the required data collected and inserted into the TransGen software tool, the first step is to perform targeting. As was written above, target designs are obtained when there are no restrictions in HE matches (all the existing matches are unfixed or unconstrained). Targeting stands for the estimation of the key performance indicators for a process or Total Site (Varbanov 2013). It refers to the best performance ahead of design based only on information available in the stream data, utility data and economic data (Gundersen 2013a). Targeting gives minimum consumption of external hot and cold utilities and maximum heat recovery by the specific minimum allowed temperature difference for heat transfer—$\Delta T_{\min}$. Maximum Energy Recovery or Minimum Energy Requirement (MER) design is obtained. By comparing the target design with the existing one it is possible to identify the potential for improvement in terms of energy consumption reduction.

Targeting is performed using the hybrid MP/PA approach in order to obtain physical insights that enable easier representations and verifications of the solution. Physical insights could be in the form of CCs, GCCs, Balanced CCs (modified CCs in which utility streams are included) for each plant, and TSPs for the entire Site. See the book sources provided in Sect. 8 for details regarding physical insights.

Besides physical insights the numerical solution is also obtained which provides requirements for each utility, utility consumption reduction and all the hot stream-to-cold stream, hot stream-to-cold utility and hot utility-to-cold stream matches. It should be noted that target design is usually very different from the existing design, especially if there are many HE units. As the existing equipment should be reutilised as much as possible and due to limitations on investment capital for retrofits, only minor changes and just a few modifications with considerable energy savings could be performed. Thus, an MER design usually has no value in the retrofit, not even as a starting point (Gundersen 2013b). However, the advantage

of the targeting step is obtaining the potential savings in terms of energy consumption. If the potential energy savings are significant, it is worth investigating the retrofit options.

5.3 Identification and Selection of Modifications

In those cases when from the targeting step significant potential for energy consumption reduction is obtained, the next step is identifications and selections of the modifications. The analysis regarding the more optimal modifications considering the trade-offs between energy and investment cost should be performed to obtain the correct setting (Smith 2005). All the modifications are compared to the existing design, and thus the results directly show improvements in terms of energy consumption, energy cost, emission savings, profit, etc.

The main objective during the identification step could be maximisation of the incremental profit or minimisation of, e.g. Total Annual Cost (TAC) by considering the trade-offs between investment and operating cost. Investment cost includes cost for additional HE areas and pipes at the Total Site level, and operating cost (savings) includes energy savings due to reduced energy consumption and intermediate utility production in cases where there is demand for it, e.g. production of hot water for district heating.

Analysis could be performed for specific units and for entire Total Sites by prohibiting of matches identified as being unsuitable for integration due to safety and other operational constraints. There could be several possibilities of obtained results regarding the selected modifications using TransGen:

- Optimum design with all the profitable solutions;
- Certain number of more optimal new/modified HE matches;
- Optimum modifications that fall within the limit of the specified payback period;
- Optimum modifications that could be performed with the available money for retrofit investment.

The incremental profit does not usually improve much when performing larger numbers of modifications. Usually significant profit could be obtained with just a few modifications, those modifications which bring the most significant savings (the ‘big apples’, Gundersen 2013b). Retrofits are, in practice, usually economical only for a small number of network modifications (Smith 2005) which can also be seen in the following on illustrative and industrial examples. As it was written above, a certain number of modifications are obtained by unfixing certain number of existing HE matches.

There are several options regarding retrofitting solutions which could be analysed using TransGen and are summarised in Table 3.

The options differ in terms of (i) reuse of the heat exchanger area where the entire load is released, (ii) minimum heat transfer for all the operating conditions (one operating condition when steady-state operation, and few or several when

Table 3 Options that could be analysed (modified from Čuček et al. 2015b)

Option	Reuse of existing area with entire load released	Minimum heat transfer for all the operating conditions	Excluded modifications
1	No	Specified	Unfeasible
2	No	–	Unfeasible
3	Yes	Specified	Unfeasible
4	Yes	–	Unfeasible
5	No	Specified	Unfeasible and difficult
6	No	–	Unfeasible and difficult
7	Yes	Specified	Unfeasible and difficult
8	Yes	–	Unfeasible and difficult

flexible operation) could be either 0 (not specified) or could be specified and should be at least, e.g. 0.5 MW, and (iii) identified modifications as being either unfeasible or unfeasible and difficult are excluded.

5.4 Verification of Modifications

The next step is verification of modifications proposed from the identification step. Some HE matches between specific process streams may be identified as unfeasible or difficult to be implemented. The reasons might be: difficulties to re-arrange the existing heat exchangers, operational difficulties, difficulties due to fouling, dynamics of operation such as non-continuous operation, safety reasons (e.g. contamination), space limitation, etc. Many identified options may be found to be unfeasible (Van Reisen 2008).

Total Site analysis is also significantly more difficult if there is any batch process in a Total Site as the production takes place only certain time periods and also there might be fluctuations in flow rates and consequently in required/excess heat. There is similar difficulty if different plants in Total Site operate at different times. In those cases there might be options for energy savings at Total Site level but only if there is appropriate heat storage available. In this case the heat losses should not be significant and should not be overlooked.

If any new HE match is identified as unfeasible or difficult (in the case that difficult modifications are also excluded), the steps 'Identification and Selection of Modifications' presented in Sect. 5.3 and 'Verification of Modifications' should be repeated. Those steps are repeated until all the proposed HE matches are acceptable for the leaders and engineers of the plant and/or Total Site. Several loops could be required to obtain (near) optimal, verified and feasible results. In the case there would be unfeasible modifications proposed, in terms of modelling those unfeasible matches should be defined as forbidden.

5.5 Synthesis of the Final Heat Exchanger Network

After all the suggested modifications are feasible to be implemented at the Total Site level, the final modifications can be subjected to a detailed capital-energy trade-offs requiring (MI)NLP optimisation (Smith 2005). This could be performed using HENSYN, and the final HEN could be synthesised from the results provided which are in a form that makes it possible to draw HENs.

It should be noted that the synthesis could be performed either by fixing the matches as identified during the second step or any matches could be obtained from identified retrofitted process streams during the second step. Furthermore there are two possibilities regarding retrofitting either reusing the heat exchanger area where the entire load is released or such heat exchanger area is not reused.

HENSYN currently enables the synthesis of the final HEN when the HE matches are such as identified during the second step, and heat exchanger area suggested not to be in use is not reused elsewhere. When performing the synthesis of the final HEN considering reuse of the heat exchanger area at some other location and forming any HE match, e.g. the model by Soršak and Kravanja (2004) could be used.

The HEN is also called a grid diagram, and provides a convenient and efficient representation of HENs. Only heat transfer operations are shown. Temperature increases from left to right in the grid. Hot streams are at the top and cold streams at the bottom. A HE match is represented by a vertical line joining two circles on the two streams being matched. An example of final HENs can be seen in Fig. 10 for illustrative example and in Fig. 13 for an industrial example.

6 Illustrative Examples

Two illustrative examples are shown in terms of the retrofit of an existing Total Site. The first illustrative example represents ‘existing’ simple Total Site consisting of two industrial plants which operate under dynamic conditions. Such an example is selected to be easily verified. It was also assumed that a certain amount of hot water should be produced. The HEN of a Total Site is proposed with five of the more optimal modifications. The second illustrative example shows the results obtained from the retrofit analysis of an existing refinery Total Site. All the results presented in the final HEN are verified to be feasible. In both retrofit studies significant potential is obtained for energy savings, reduced cost and reduced emissions.

6.1 Retrofit of a Small-Scale Total Site

A simple illustrative example is performed of a Total Site consisting of two industrial plants A and B (Perry et al. 2008) with a small number of streams.

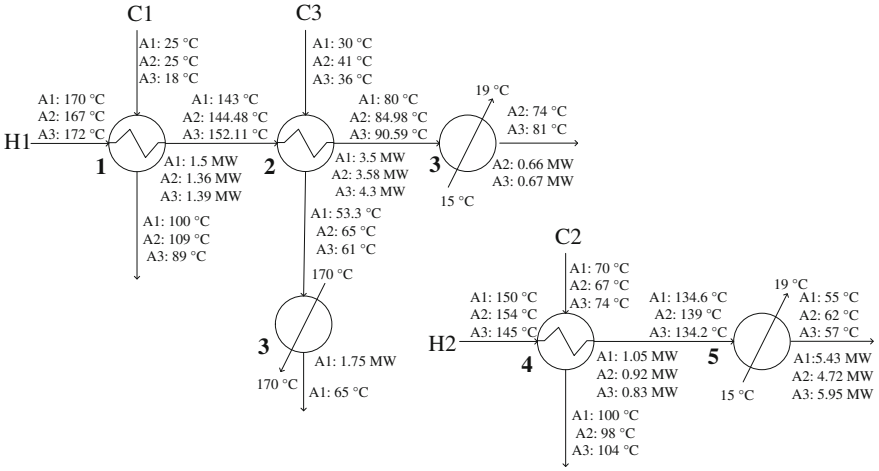


Fig. 8 Heat exchanger network design for plant A (modified from Čuček and Kravanja 2015a)

The HEN for plant A under varying operating conditions (three representative cases) is shown in Fig. 8, and plant B is shown in Fig. 9. Case 1 is in Figures represented as "A1", case 2 as "A2" and case 3 as "A3".

Industrial plants operate under flexible varying conditions. It is assumed that they are located 1.5 km apart. All process streams are assumed to be liquids and for simplicity it is also assumed that the pressures of the process streams are at 1 bar. Also, it is assumed that 2 MW of hot water should be produced.

The first step (targeting step) is to perform targeting. By comparing the target and existing designs, it is possible to identify the potentials for improvements within process plants and Total Sites in terms of energy consumption reduction. The main results obtained for the target and existing designs are shown in Table 4 from which it can be seen that there is significant potential for Total Site Integration.

The second step (identification step) is the step where modifications are identified. The analysis is performed regarding the more optimal modifications (new HE matches) by considering the trade-offs between energy and investment costs. All the modifications could be compared to the existing design, and thus the results could directly show possible improvements in terms of energy consumption and cost. Analysis is performed for a certain number of the more optimal new HE matches, e.g. for five new HE matches. The main results regarding modified designs with five modifications (new HE matches) are shown in Table 5.

It can be seen that the results regarding energy consumption are obtained within ranges. Most of the new HE matches are selected at Total Site level, whilst 2 MW of hot water was produced from hot stream 5 at plant A. Further, verification of the modifications is performed and the second step is repeated until all the proposed HE matches are feasible and accepted. The identification step could be repeated several

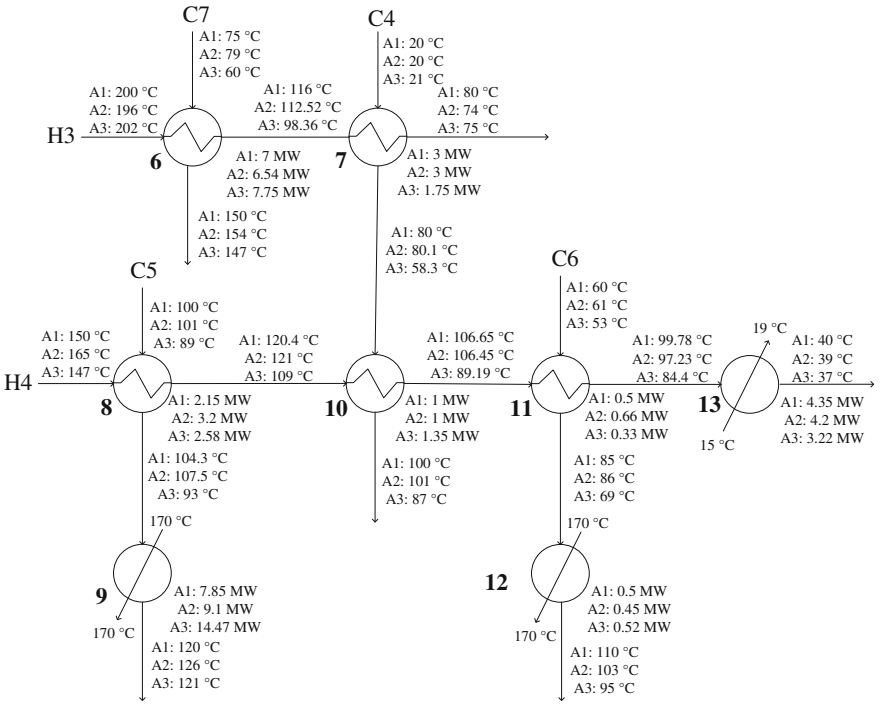


Fig. 9 Heat exchanger network design for plant B (modified from Čuček and Kravanja 2015a)

Table 4 Main results regarding existing and target designs

	Existing design	Target design
Average energy cost	3.4 M€/y	0.9 M€/y
Average energy consumption	21.4 MW	4.7 MW
Hot utility consumption		
Case 1|Case 2|Case 3	10.1 MW|9.6 MW|15.4 MW	1.7 MW|3.7 MW|6.7 MW
Cold utility consumption		
Case 1|Case 2|Case 3	9.8 MW|9.6 MW|9.8 MW	0.1 MW|2 MW|0.2 MW

times by excluding either unfeasible or unfeasible and difficult modifications, before the final solution could be obtained. In this illustrative example it is assumed that all the modifications are acceptable and thus this step is not repeated. Results from the first and second steps were obtained using software tool TransGen.

The final third step (synthesis step) is synthesis of the retrofitted HEN. The final step is performed using code HENSYN which considers more appropriate trade-offs between investment and operating costs. The synthesis is performed only for those selected retrofitting modifications during the second identification step. Figure 10

Table 5 Main results regarding modified design with five new heat exchange matches (Čuček and Kravanja 2015a)

	Modified design with five new HE matches
Average profit	1.25 M€/y
Average energy savings	7.96 MW
Retrofit area investment	0.72 M€
Pipe investment	0.83 M€
Payback time	0.94 y
Modified HE units (fraction of rearranged HE unit's energy)	2 (29–40 %), 5 (87–96 %), 6 (30–43 %), 9 (32–43 %), 13 (27–60 %)
New HE matches at unit level	H6–C9 (1.9–3.3 MW), H5–hot water (2 MW)
New HE matches at Total Site level	H2–C9 (1–1.7 MW), H5–C6 (2.2–3.7 MW), H13–C2 (1.1–1.9 MW)
Hot utility consumption (absolute and relative reduction)	
Case 1|Case 2|Case 3	6.7 MW (3.4 MW–33 %)|6.7 MW (2.9 MW–30 %)|10.6 MW (4.9 MW– 31.6 %)
Cold utility consumption (absolute and relative reduction)	
Case 1|Case 2|Case 3	3.9 MW (5.9 MW–60 %)|4.3 MW (5.3 MW–55 %)|2.2 MW (7.6 MW–77 %)

shows the retrofitted HEN design when performing five of the more profitable modifications. New heat exchangers in HEN are shown in yellow, and the existing ones the areas of which should be reduced are shown in grey. Heaters are highlighted with red and coolers with blue colours. The HEN includes also all the temperatures separately for each scenario, exchanged heat, and areas and types of selected heat exchangers.

Two heat exchanger types were selected, fixed plate shell and tube, and shell and tube with U tubes. It can be seen that there are five new HE units. From those five, three are exchanging heat between the plants A and B, one within plant B, and one is used for hot water production (previously heat exchanger 5). Hot water production is completely satisfied.

6.2 Retrofit of an Existing Refinery Total Site

This section presents the retrofit analysis of a large-scale industrial Total Site under fixed and varying conditions. The refinery operates under varying operating conditions due to ambient temperature fluctuations and crude oil feedstocks. The most significant differences are between the summer and winter months, and due to two different types of crude oil feedstocks, sweet and sour crudes.

Four of the more significantly different operating conditions are selected in order to cover the extreme variations for the purpose of retrofitting. Four units U1–U4 within the refinery are analysed under four operating conditions (cases) C1–C4.

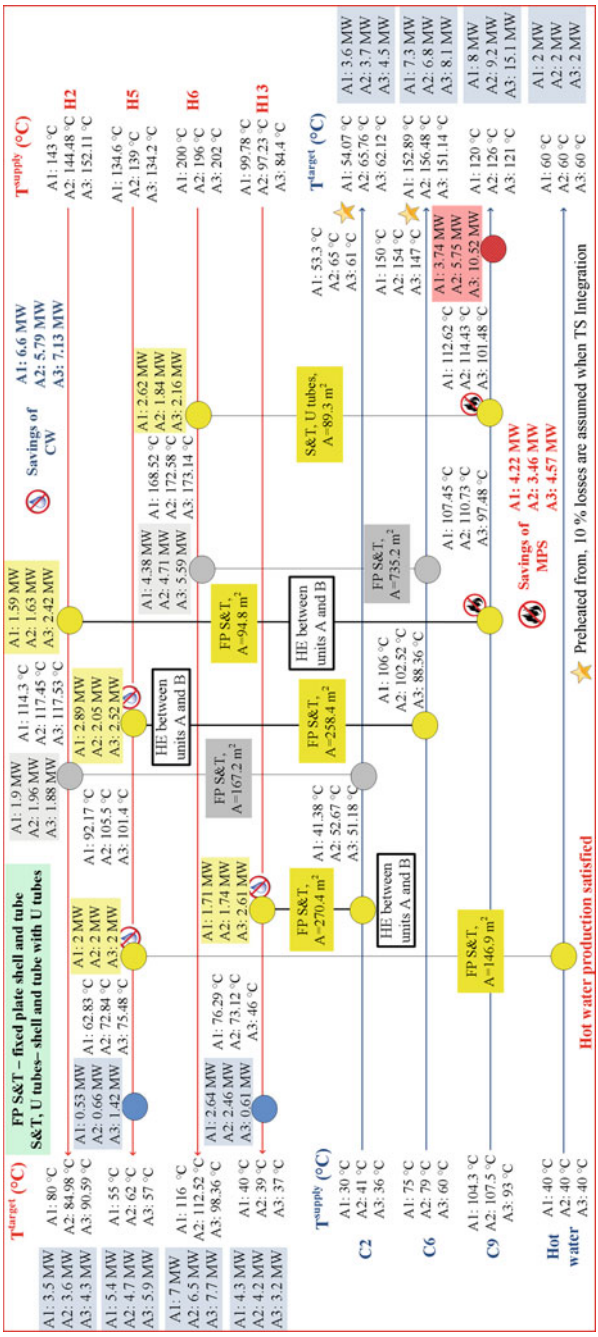


Fig. 10 HEN of a small-scale Total Site with five new heat exchange matches (modified from Čuček and Kravanja 2015a)

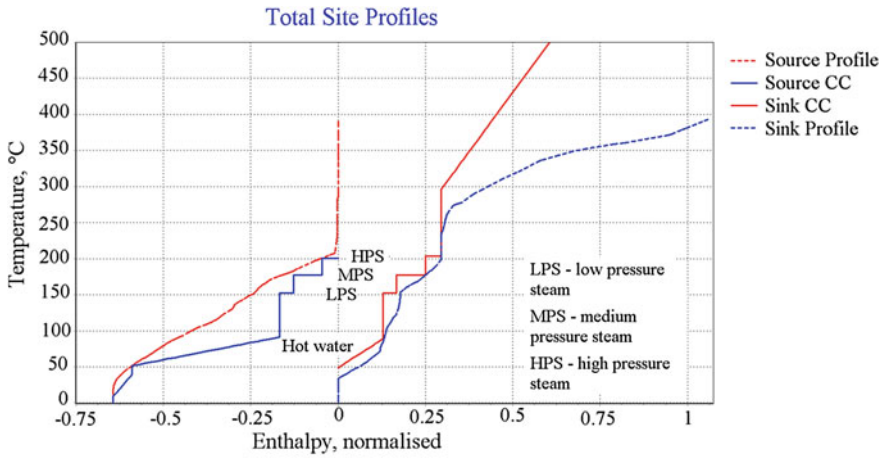


Fig. 11 Total Site Profiles with hot water production for operating condition C4 (modified from Čuček et al. 2015b)

The motivation for retrofitting is twofold, the main one being to find sources for producing a certain amount of hot water for district heating. The second task is to perform Heat and Total Site Integration to find out if there are still remaining potentials for reducing external utilities' consumptions. It should be noted that the data are confidential and cannot be presented. The obtained results are regarded also as confidential and are presented in normalised or 'hidden' form. However, the shapes of the curves and all the ratios between results are preserved (Čuček et al. 2015b).

The first step when having the required data collected is the targeting step to obtain the potential for improvement in terms of utility consumption reduction and intermediate utility production. Special emphasis is on estimating the hot water production potential (Čuček et al. 2015b). Figure 11 shows an example of TSPs for case C4, and also shows the amount of hot water which could be produced in the targeted design. Significant potentials for hot water production and for energy consumption reduction are shown from the targeting step.

The next step is the identification and selection of retrofitting modifications. The analysis is performed regarding the more optimal modifications with the main objective of maximising the incremental profit by considering the trade-offs between energy and investment costs.

Prohibition of matches between some units is considered for those identified as being non-integrated. Production and consumption of steam at different pressure levels is also considered as being fixed at current production and consumption levels. Integration between units is based on the distance between units multiplied by 3 (multiplied by 6 for both directions), and only direct integration is considered by transporting cold stream. Heat loss of 10 % is considered when the integration is performed between units. The average heat exchange duties are considered for all hot and cold streams if there are variations between the duties at the cold and hot sides.

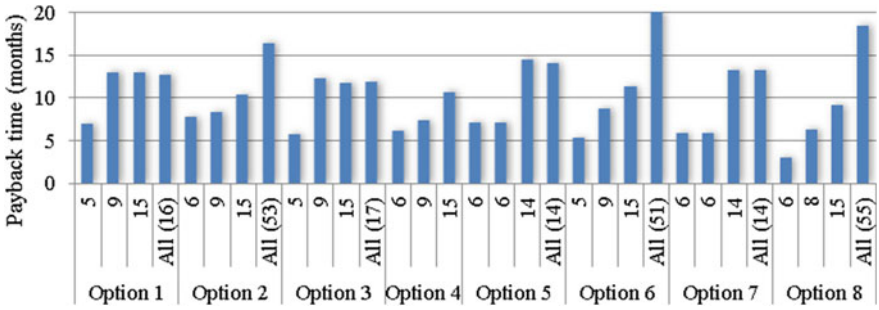


Fig. 12 Payback period for proposed modifications (maximum 6, 9, 15 and all profitable) (Čuček et al. 2015b)

Analysis was performed using the software tool TransGen for a certain number of the most optimal new HE matches (e.g. 6, 9 and 15) by relaxing a certain number of the restrictions in terms of existing HE units. According to feedback regarding the feasibilities of the proposed modifications, the identification step was repeated several times by excluding unfeasible and difficult modifications until all the proposed HE matches were acceptable, and the final solution had been obtained. Several options were analysed for retrofitting, see Table 3.

Figure 12 illustrates an example of the obtained main results regarding those options in terms of the payback period. It shows the results obtained from the second set of proposed modifications by forming maximum 6, 9, and 15 and all new profitable HE matches.

It can be seen from Fig. 12 that the payback periods vary from a few months up to around 1.5 years when performing all profitable modifications. However, it should be noted that payback periods do not represent the final solution (due to confidentiality reasons), and some of the modifications identified are unfeasible or difficult and are thus excluded from the next iterations in terms of feasible retrofit modifications.

When all the retrofitting modifications have been identified as feasible, the synthesis of retrofitted HEN is obtained during the third step, based on the information regarding those feasible modifications. Figure 13 shows an example of the obtained HEN with a maximum of nine new HE matches. Only HE units that involve stream segments where existing heat exchangers should be modified are shown, either being removed or with reduced area.

New heat exchangers in HEN are again shown in yellow, existing ones which should be modified are shown in grey, heaters in red and coolers in blue. The HEN also includes all the temperatures (in °C) separately for each scenario and exchanged duties (A…AR). The proposed types of heat exchangers identified are also shown in Fig. 13, and the calculated reduction in the existing heat exchanger area. The enthalpy values are denoted as A, B…AR as they are confidential. The first and second steps were again performed using the software tool TransGen, and the final step using code HENSYN.

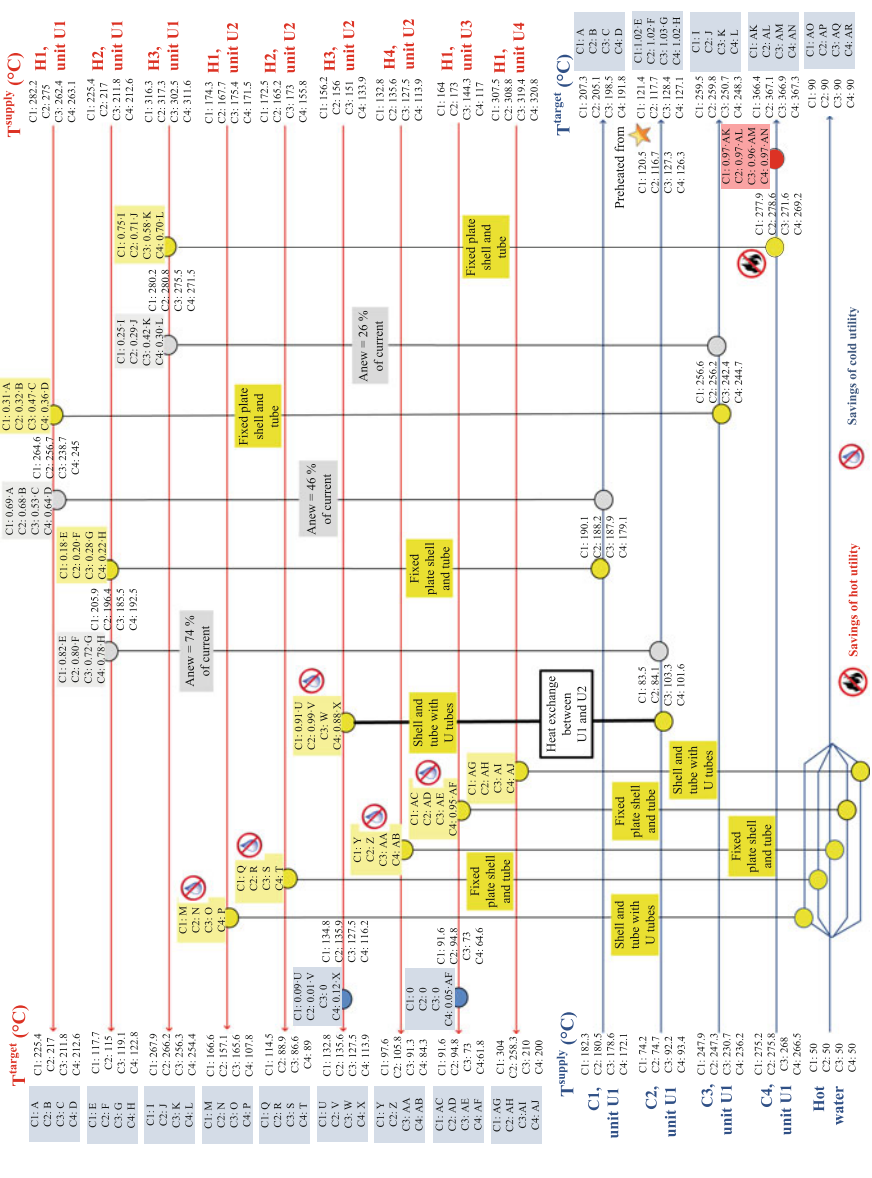


Fig. 13 Heat exchanger network with a maximum of nine new heat exchange matches (modified from Čuček et al. 2015b)

7 Concluding Remarks

The primary objective of this chapter was to provide an introduction and main approaches to Total Site Integration and present the recently developed methodologies for the retrofitting of the existing Total Sites. Those methodologies have also been demonstrated on illustrative and industrial examples.

The methodologies and procedures for retrofitting existing Total Sites has been presented which enable retrofitting of large-scale industrial plants and Total Sites under fixed or steady-state operating conditions and under uncertain dynamic operating conditions with any number of existing HE units. The industrial example of an existing refinery Total Site analysed under four operating conditions included, e.g. more than 100 heat exchange units. It was shown that significant energy savings are possible over a short payback period. The developed automated procedure for the purpose of retrofitting large-scale industrial Total Sites consists of the following steps:

1. Targeting and identification of potential for Heat and Total Site Integration;
2. Identification and selection of retrofitting modifications;
3. Final synthesis of retrofitted HEN.

It should be noted that the developed and demonstrated methodology and approach could be used for any process plant and/or Total Site including domestic and service sectors also. There is a large potential for energy savings worldwide within the industrial, domestic and service sectors.

Recently, Anantharaman (2011) stated: 'Though there has been significant developments in HENs using MP methods, synthesis of large scale HENs problems without simplifications and heuristics have been lacking. This is an area that requires more research before MP-based approaches can be used in the industry'. Presenting the methodology and approach for overcoming the lack of synthesis and retrofitting large-scale HENs problems at plant and Total Site levels using MP was one of the main intentions of this chapter. It should also be noted that the presented methodology and approach for retrofitting large-scale HENs within flexible Total Sites has been proven to be an interesting approach that may contribute to significant energy and cost savings and pollution reduction EFENIS (2015b).

8 Sources of Further Information

Heat Integration and Total Site Analysis have been used extensively since its development, and thus there are several journal and conference papers, and books published. Furthermore there are numerous useful sources available on the web such as, e.g. by CETC-Varennes (2003) and Gundersen (2000).

There are several excellent books which cover the topic of Process Integration mainly focusing on the PA as the main methodology. Those books provide

invaluable source of information for learning Total Site Integration by PA approach and the basic concepts of MP, and Process Integration by MP approach. However, it should be noted that the list of books is not comprehensive but it provides several of the better-known and respected sources.

The more well-known book on Process Integration is the 'Red Book' entitled User Guide to Process Integration for the Efficient Use of Energy, authored by the team of founders of Process Integration technology (Linnhoff et al. 1982). Its revised edition with a review of developments since 1982 was published in 1994 (Linnhoff et al. 1994). Those books remain an excellent source for teaching and learning about Heat Integration and several other Process Integration technology applications (Bulatov, 2013).

Basic concepts in process synthesis including heat and power integration, optimisation approaches to process synthesis and design including synthesis of HENs and simultaneous optimisation and heat integration are provided in the book entitled Systematic methods of chemical process design by Biegler et al. (1997). It also provides more details regarding the several types of MP formulations that could address Total Site Integration such as the transhipment problem, stagewise problem, etc.

One of the sources which describes in details the principles of Process Integration both for synthesis and retrofits by PA in several chapters is the 'Blue book' (Smith 1995) and updated entitled Chemical Process Design and Integration (Smith 2005). It deals with the designs and integrations of chemical processes with the emphasis on conceptual issues. It provides a practical guide to chemical process design and integration for undergraduate and postgraduate students of chemical engineering, practicing process designers and chemical engineers and applied chemists working in process development. It deals with all aspects of chemical engineering, such as economics, optimisation, thermodynamics, reactors, separators, continuous and batch processes, HENs, Heat and Total Site Integration, environmental emissions, water systems, and inherent safety.

An updated and extended second edition of the 'Red Book' entitled Pinch Analysis and Process Integration, A User Guide on Process Integration for the Efficient Use of Energy was published in 2007 (Kemp 2007), and is also called 'Black Book'. It deals with key concepts of PA, with special focus on data extraction, energy targeting and HEN design and heat and power systems including the analysis of Total Sites, etc. It also demonstrates how to apply the technology in practice including stream data extraction. Several case studies are provided. This book also provides a spreadsheet for PA.

A more recent book entitled Sustainability in the Process Industry: Integration and Optimization (Klemeš et al. 2010) provides the basics regarding process integration and optimisation and provides recent developments in those topics until 2010. It provides support for graduate and postgraduate students worldwide and for practitioners from various fields of the processing industry. It is supported by several case studies of varying complexities including eight industry-based case studies.

Of outstanding interest might be a recently published Handbook of Process Integration (PI): Minimisation of energy and water use, waste and emissions (Klemeš 2013b). It provides in several sections the recent advances in Total Site Integration and its extensions, e.g. to address varying energy supply and demand. It also presents the application of Pinch Technology to Industrial Total Sites and software tools for Heat Integration amongst others. The leading scientists and researchers in the field of Process Integration contributed to the book.

The recent book Process Integration and Intensification: Saving Energy, Water and Resources by Klemeš et al. (2014) aims at a detailed description of selected Process Integration and Intensification principles and thus enabling the reader to solve both illustrative and industrial problems. It is supported by several working sessions to allow the reader to practise and to avoid potential pitfalls.

Besides those high-quality books on Process Integration there are several review papers which may be interesting for the reader. The comprehensive overview of the syntheses of HENs was presented by Gundersen and Naess (1988). A critical review with an annotated bibliography for the literature on HENs up to the year 2000 was published by Furman and Sahinidis (2002). A review of the recent developments in retrofit methodologies for HENs is provided by Smith et al. (2010). Morar and Agachi (2010) reviewed the important contributions in heat integration techniques during 1975–2008. Anantharaman (2011) presented a HENs bibliography up to the year 2010. Klemeš and Kravanja (2013) provided a short overview of historical developments, achievements, and future challenges of PA and MP for Heat Integration. Furthermore, Klemeš et al. (2013b) presented an overview of the recent achievements in Process Integration and future challenges. The main issues that can influence the practical implementation of Total Site Heat Integration in industry are provided in Chew et al. (2013). Ahmetović et al. (2015) reviewed energy integration studies besides water integration studies published over the last two decades.

There is a dedicated conference for Process Integration—the Conference on Process Integration, Modelling and Optimisation for Energy Savings and Pollution Reduction—PRES (Conference PRES 2015). Process Integration also plays an important role at some other conferences such as at the European Symposium on Computer Aided Process Engineering—ESCAPE (Conference ESCAPE 2015), Conference on Sustainable Development of Energy, Water and Environment Systems—SDEWES (Conference SDEWES 2015) and many others. The reader may refer to the conference proceedings and special issues dedicated to those conferences for further information regarding Total Site Integration.

Acknowledgments The authors acknowledge the financial support from EC FP7 project ENER/FP7/296003/EFENIS ‘Efficient Energy Integrated Solutions for Manufacturing Industries—EFENIS’, from SCOPES joint research project CAPE-EWWR ‘Computer Aided Process Engineering applied to energy, water, and waste reduction during process design and operation’, and from the Slovenian Research Agency (Program Numbers P2-0032 and P2-0377). The authors are also grateful to EFENIS partners, especially to Prof. Jiří Jaromír Klemeš, Dr. Valter Mantelli, Prof. Petar Sabev Varbanov and DDr. Andreja Nemet.

Glossary

Grassroots design	It represents the design of a new plant or Total Site. It is also called minimum energy requirement design and has the widest freedom of choice.
Heat Integration	It means integrating different processes to achieve energy savings (Klemeš et al. 2010). It examines the potential for improving and optimising the heat exchange in order to reduce the amount of external utilities used and thus to reduce costs and emissions.
Mathematical Programming	It is also called 'mathematical optimisation', or 'optimisation'. It stands for the use of specific methods for determining the best solutions to a problem, subject to given constraints. It is the act of obtaining the best result under given circumstances (Rao 2009). It is the key methodology used for sustainable process design and synthesis (Klemeš et al. 2010). The general form of the mathematical programming is the mixed integer nonlinear program (MINLP). The optimality, feasibility and integrality of the obtained solutions are the main capabilities of Mathematical Programming (Kravanja 2010).
Pinch Analysis or Pinch Technology	It is a systematic methodology for analysing and optimising energy savings in processes, plants and Total Sites. It sets the targets by determining the thermodynamic maximal possible rate of heat recovery (Nemet et al. 2015b).
Process Integration	It consists of a family of methodologies for reducing the consumption of resources and emissions within processes, plants and Total Sites (Klemeš et al. 2010). One of the definitions of Process Integration is that it represents (Gundersen 2000): *'Systematic and General Methods for Designing Integrated Production Systems ranging from Individual Processes to Total Sites, with special emphasis on the Efficient Use of Energy and reducing Environmental Effects'*. Methodologies of Process Integration are: Heat Integration, Total Site Integration, Mass Integration, Water Integration, Hydrogen Pinch, etc.

Retrofit design It stands for the modification of the existing plant(s) or Total Site. It is also called revamp, reconstruction or redesign. The motivation to retrofit could be to increase capacity, allow for different feed or product specifications, reduce operating cost, improve safety, increase capacity, etc. (Smith 2005). The fewer there are modifications the better. There are many ways of improving the existing designs, such as changes in the usages of utilities, topological modifications, installing of additional areas, repiping of streams, reassignments of matches and heat transfer enhancement (Wang et al. 2012).

Code HENSYN It is used for the final synthesis of the retrofitted heat exchanger network (HEN) for the identified more promising retrofitting modifications obtained from the software tool TransGen. It consideres in detail the trade-offs between investment and operating cost. It can be applied for the retrofitting of existing large-scale industrial plants and Total Sites under nominal and uncertain conditions (Čuček and Kravanja 2015a).

Software tool TransGen Software tool for (i) targeting and comparing the target and existing designs for obtaining the potential for Heat and Total Site Integration and intermediate utility production, and (ii) identifications of alternatives for retrofitting modifications, and selections of modifications by forbidding unfeasible matches. It is an extension of an expanded transhipment model (Papoulias and Grossmann 1983) and is based on a combined Mathematical Programming/Pinch Analysis (MP/PA) approach. It is suitable for analysing energy targets and existing HEN designs, and especially for proposing optimal modifications for the retrofitting of each plant and Total Sites under nominal and uncertain conditions. Retrofitting of HENs could be performed for problems of any size (Čuček and Kravanja 2015b).

Total Site It is a set of factories incorporating several processes, serviced by and linked through a central utility system (Dhole and Linnhoff 1993). The concept of Total Site has been extended by including processes from residential, business, service and agricultural sectors (Perry et al. 2008), renewable energy sources and storage systems for accommodating the variations (Varbanov and Klemeš 2011), surrounding of the Total Site by district heating and cooling (Gundersen 2013a) and to a regional level (Čuček et al. 2013).

Total Site Integration It offers energy conservation opportunities across different individual processes and for designing and optimising the central utility system (Bandyopadhyay et al. 2010). There are two possibilities for Total Site Integration: indirect via intermediate utility and direct integration between process streams where either hot or cold process stream is transported between the processes.

References

Aaltola J (2002) Simultaneous synthesis of flexible heat exchanger network. Applied Thermal Engineering 22 (8):907–918.

Ahmetović E, Ibrić N, Kravanja Z, Grossmann IE (2015) Water and energy integration: A comprehensive literature review of non-isothermal water network synthesis. Computers & Chemical Engineering 82:144–171.

Ahmetović E, Kravanja Z (2013) Simultaneous synthesis of process water and heat exchanger networks. Energy 57:236–250.

Alfa Laval (2011) Waste heat recovery – Optimizing your energy system. local.alfalaval.com/en-gb/about-us/news/Documents/PPI00443EN%20Optimising%20your%20energy%20system.pdf. Accessed 12.8.2015.

Anantharaman R (2011) Energy Efficiency in Process Plants with emphasis on Heat Exchanger Networks: Optimization, Thermodynamics and Insight. Department of Energy and Process Engineering, Norwegian University of Science and Technology, Trondheim, Norway PhD thesis www.ivt.ntnu.no/ept/fag/tep4215/innhold/Rahul%20-%20thesis%20-%20final%20version%20-%2030%20May%202011.pdf Accessed 17.8.2015.

Angsutorn N, Siemanond K, Chuvaree R (2014) Heat Exchanger Network Synthesis using MINLP Stage-wise Model with Pinch Analysis and Relaxation. In: Jiří Jaromír Klemeš PSV, Peng Yen L (eds) Computer Aided Chemical Engineering, vol Volume 33. Elsevier, pp 139–144.

Asante NDK, Zhu XX (1996) An automated approach for heat exchanger network retrofit featuring minimal topology modifications. Computers & Chemical Engineering 20, Supplement 1:S7-S12.

Asante NDK, Zhu XX (1997) An Automated and Interactive Approach for Heat Exchanger Network Retrofit. Chemical Engineering Research and Design 75 (3):349–360.

Bagajewicz M, Valtinson G, Nguyen Thanh D (2013) Retrofit of Crude Units Preheating Trains: Mathematical Programming versus Pinch Technology. Industrial & Engineering Chemistry Research 52 (42):14913–14926.
Bandyopadhyay S, Varghese J, Bansal V (2010) Targeting for cogeneration potential through total site integration. Applied Thermal Engineering 30 (1):6–14.
Biegler LT, Grossmann IE, Westerberg AW (1997) Systematic methods of chemical process design. Prentice Hall PTR, New Jersey, US.
Björk K-M, Nordman R (2005) Solving large-scale retrofit heat exchanger network synthesis problems with mathematical optimization methods. Chemical Engineering and Processing: Process Intensification 44 (8):869–876.
Bogataj M, Kravanja Z (2012) An alternative strategy for global optimization of heat exchanger networks. Applied Thermal Engineering 43:75–90.
Boldyryev S, Varbanov PS, Nemet A, Klemeš J, Kapustenko P (2013) Capital Cost Assessment for Total Site Power Cogeneration. Computer Aided Chemical Engineering 32:361–366.
BP (2015) BP Statistical Review of World Energy, June 2015. www.bp.com/content/dam/bp/pdf/Energy-economics/statistical-review-2015/bp-statistical-review-of-world-energy-2015-full-report.pdf. Accessed 29.7.2015.
Bradley SP, Hax AC, Magnanti TL (1977) Applied Mathematical Programming. Addison-Wesley, Reading, MA, US.
Briones V, Kokossis AC (1999) Hypertargets: a Conceptual Programming approach for the optimisation of industrial heat exchanger networks — II. Retrofit design. Chemical Engineering Science 54 (4):541–561.
Bulatov I (2013) Software Tools for Heat Integration. In: Klemeš JJ (ed) Handbook of Process Integration (PI): Minimisation of Energy and Water Use, Waste and Emissions. Woodhead Publishing Limited, Cambridge, UK.
Cerda J, Westerberg AW, Mason D, Linnhoff B (1983) Minimum utility usage in heat exchanger network synthesis A transportation problem. Chemical Engineering Science 38 (3):373–387.
Cerda J, Westerberg AW (1983) Synthesizing heat exchanger networks having restricted stream/stream matches using transportation problem formulations. Chemical Engineering Science 38 (10):1723–1740.
CETC-Varennes (2003) Pinch Analysis: For the Efficient Use of Energy, Water and Hydrogen. CANMET Energy Technology Centre-Varennes. www.nrcan.gc.ca/sites/www.nrcan.gc.ca/files/canmetenergy/pdf/fichier.php/codectec/En/2009-052/2009-052_PM-FAC_404-DEPLOI_e.pdf. Accessed 10.8.2015.
Chen Y, Grossmann IE, Miller DC (2015) Computational strategies for large-scale MILP transshipment models for heat exchanger network synthesis. Computers & Chemical Engineering 82:68–83.
Chew KH, Klemeš JJ, Wan Alwi SR, Abdul Manan Z (2013) Industrial implementation issues of Total Site Heat Integration. Applied Thermal Engineering 61 (1):17–25.
Conference ESCAPE (2015) European Symposium on Computer-Aided Process Engineering. escape26.um.si/technical-themes/. Accessed 14.8.2015.
Conference PRES (2015) Conference Process Integration, Modelling and Optimisation for Energy Saving and Pollution Reduction. www.conferencepres.com/. Accessed 15.8.2015.
Conference SDEWES (2015) Conference on Sustainable Development of Energy, Water and Environment Systems. www.dubrovnik2015.sdewes.org/. Accessed 14.8.2015.
CPI Manchester (2015) Efenis-Site. CPI, CEAS, The University of Manchester, UK. demo.efenis.alexandra.dk/public-portal/admin.hta. Accessed 15.8.2015.
Čuček L, Klemeš JJ, Varbanov PS, Kravanja Z (2015a) Significance of environmental footprints for evaluating sustainability and security of development. Clean Technologies and Environmental Policy 17(8): 2125–2141.
Čuček L, Kravanja Z (2014) Efficient Transshipment-Based Framework for Energy Targeting and Retrofitting Industrial Total Sites. Chemical Engineering Transactions 39:1813–1818.
Čuček L, Kravanja Z (2015a) A Procedure for the Retrofitting of Large-Scale Heat Exchanger Networks for Fixed and Flexible Designs. Chemical Engineering Transactions 45:31–36.

Čuček L, Kravanja Z (2015b) Retrofitting of large-scale heat exchanger networks within total sites under uncertainty by considering trade-offs between investment and operating cost. Chemical Engineering Transactions 45:1723–1728.

Čuček L, Kravanja Z (2015c) Software Tools for Mixed Pinch and Mathematical Programming Analysis, Deliverable 1.10 for EFENIS EC FP7 project ENER/FP7/296003/EFENIS 'Efficient Energy Integrated Solutions for Manufacturing Industries'. Faculty of Chemistry and Chemical Engineering, University of Maribor.

Čuček L, Mantelli V, Yong JY, Varbanov PS, Klemeš JJ, Kravanja Z (2015b) A Procedure for the Retrofitting of Large-Scale Heat Exchanger Networks for Fixed and Flexible Designs applied to Existing Refinery Total Site. Chemical Engineering Transactions 45:109–114.

Čuček L, Varbanov PS, Klemeš JJ, Kravanja Z (2013) Multi-Objective Regional Total Site Integration. Chemical Engineering Transactions 35:97–102.

Čuček L, Yong JY, Mantelli V, Vocciante M, Varbanov PS, Klemeš JJ, Karlopoulos E, Kravanja Z (2014) Data Acquisition and Analysis of Total Sites under Varying Operational Conditions. Chemical Engineering Transactions 39:1819–1824.

Dhole VR, Linnhoff B (1993) Total site targets for fuel, co-generation, emissions, and cooling. Computers and Chemical Engineering 17, Supplement 1:S101 - S109.

Dolan WB, Cummings PT, LeVan MD (1989) Process optimization via simulated annealing: Application to network design. AIChE Journal 35 (5):725–736.

Duran MA, Grossmann IE (1986) Simultaneous optimization and heat integration of chemical processes. AIChE Journal 32 (1):123–138.

EFENIS (2015a) Efficient Energy Integrated Solutions for Manufacturing Industries, EC Project "EFENIS" Grant Agreement 296003, efenis.uni-pannon.hu/project-overview/#demonstration. Accessed 25.1.2016

EFENIS (2015b) Efficient Energy Integrated Solutions for Manufacturing Industries, EC Project "EFENIS" Grant Agreement 296003, Fundamentally improved total site analysis for demo site 4 and documentation of achieved benefits of improved methods – public summary, efenis.uni-pannon.hu/wp content/uploads/2013/02/Deliverable-8.3_public-summary_rev.pdf Accessed 25.1.2016

El-Halwagi MM (2012) Sustainable Design Through Process Integration: Fundamentals and Applications to Industrial Pollution Prevention, Resource Conservation, and Profitability Enhancement. Butterworth-Heinemann, Waltham, MA, USA.

Floudas CA, Ciric AR, Grossmann IE (1986) Automatic synthesis of optimum heat exchanger network configurations. AIChE Journal 32 (2):276–290.

Floudas CA, Grossmann IE (1986) Synthesis of flexible heat exchanger networks for multiperiod operation. Computers & Chemical Engineering 10 (2):153–168.

Fraser DM (2013) Retrofit Mass Integration of Acid Gas Removal Systems in Petrochemical Plants. In: Klemeš JJ (ed) Handbook of Process Integration (PI): Minimisation of Energy and Water Use, Waste and Emissions. Woodhead Publishing Limited, Cambridge, UK.

Frausto-Hernández S, Rico-Ramírez V, Jiménez-Gutiérrez A, Hernández-Castro S (2003) MINLP synthesis of heat exchanger networks considering pressure drop effects. Computers & Chemical Engineering 27 (8–9):1143–1152.

Friedler F, Tarján K, Huang YW, Fan LT (1992) Graph-theoretic approach to process synthesis: axioms and theorems. Chemical Engineering Science 47 (8):1973–1988.

Friedler F, Varbanov P, Klemeš J (2009) Advanced HENs design for multi-period operation using P-graph. Chemical Engineering Transactions 18 (1):457–462.

Furman KC, Sahinidis NV (2002) A Critical Review and Annotated Bibliography for Heat Exchanger Network Synthesis in the 20th Century. Industrial & Engineering Chemistry Research 41 (10):2335–2370.

Grossmann IE (1992) Mathematical methods for heat exchanger network synthesis. repository.cmu.edu/cgi/viewcontent.cgi?article = 1186&context = cheme. Accessed 10.08.2015.

Grossmann IE, Caballero JA, Yeomans H (2000) Advances in mathematical programming for the synthesis of process systems. Latin American Applied Research 30 (4):263–284.

Grossmann IE, Daichendt MM (1996) New trends in optimization-based approaches to process synthesis. Computers & Chemical Engineering 20 (6–7):665–683.
Gundersen T (2000) A process integration primer—Implementing agreement on process integration. Trondheim, Norway: International Energy Agency, SINTEF Energy Research.
Gundersen T (2013a) Heat Integration: Targets and Heat Exchanger Network Design. In: Klemeš JJ (ed) Handbook of Process Integration (PI): Minimisation of Energy and Water Use, Waste and Emissions. Woodhead Publishing Limited, Cambridge, UK.
Gundersen T (2013b) Analysis and Design of Heat Recovery Systems for Grassroots and Retrofit Situations. In: Klemeš JJ (ed) Handbook of Process Integration (PI): Minimisation of Energy and Water Use, Waste and Emissions. Woodhead Publishing Limited, Cambridge, UK.
Gundersen T, Naess L (1988) The synthesis of cost optimal heat exchanger networks: An industrial review of the state of the art. Computers & Chemical Engineering 12 (6):503–530.
Hackl R, Andersson E, Harvey S (2011) Targeting for energy efficiency and improved energy collaboration between different companies using total site analysis (TSA). Energy 36 (8):4609–4615.
Hackl R, Harvey S, Andersson E (2010) Total Site Analysis (TSA) Stenungsund. Department of Energy and Environment, Division of Heat and Power Technology, Chalmers University of Technology. publications.lib.chalmers.se/records/fulltext/local_131484.pdf. Accessed 16.8. 2015.
Herrera A, Islas J, Arriola A (2003) Pinch technology application in a hospital. Applied Thermal Engineering 23 (2):127–139.
Isafiade A, Bogataj M, Fraser D, Kravanja Z (2015) Optimal synthesis of heat exchanger networks for multi-period operations involving single and multiple utilities. Chemical Engineering Science 127:175–188.
Isafiade AJ, Fraser DM (2008) Interval-based MINLP superstructure synthesis of heat exchange networks. Chemical Engineering Research and Design 86 (3):245–257.
KBC (2014) SuperTargetTM. www.kbcat.com/energy-utilities-software/supertarget. Accessed 15. 8.2015.
Kemp IC (2007) Pinch Analysis and Process Integration: A User Guide on Process Integration for the Efficient Use of Energy, Second Edition. Elsevier Science, Oxford, UK.
Klemeš J, Dhole VR, Raissi K, Perry SJ, Puigjaner L (1997) Targeting and design methodology for reduction of fuel, power and CO2 on total sites. Applied Thermal Engineering 17:993–1003.
Klemeš J, Friedler F, Bulatov I, Varbanov P (2010) Sustainability in the Process Industry: Integration and Optimization: Integration and Optimization. McGraw-Hill Companies, New York, USA.
Klemeš JJ (2013a) Epilogue: The Importance of Problem Formulation and Data Extraction in Process Integration. In: Klemeš JJ (ed) Handbook of Process Integration (PI): Minimisation of Energy and Water Use, Waste and Emissions. Woodhead Publishing Limited, Cambridge, UK.
Klemeš JJ (2013b) Handbook of Process Integration (PI): Minimisation of energy and water use, waste and emissions. Woodhead Publishing Limited, Cambridge, UK.
Klemeš JJ, Kravanja Z (2013) Forty years of Heat Integration: Pinch Analysis (PA) and Mathematical Programming (MP). Current Opinion in Chemical Engineering 2 (4):461–474.
Klemeš JJ, Varbanov PS, Kravanja Z (2013a) Recent developments in Process Integration. Chemical Engineering Research and Design 91 (10):2037–2053.
Klemeš JJ, Kravanja Z, Varbanov PS, Lam HL (2013b) Advanced multimedia engineering education in energy, process integration and optimisation. Applied Energy 101:33–40.
Klemeš JJ, Varbanov PS, Alwi SRW, Manan ZA (2014) Process Integration and Intensification: Saving Energy, Water and Resources. Walter de Gruyter GmbH, Berlin, Germany.
Kostevšek A, Petek J, Čuček L, Klemeš JJ, Varbanov PS (2015) Locally Integrated Energy Sectors supported by renewable network management within municipalities. Applied Thermal Engineering 89: 1014–1022.
Krajnc M, Kovač-Kralj A, Glavič P (2006) Heat integration in a speciality product process. Applied Thermal Engineering 26 (8–9):881–891.

Kravanja Z (2010) Challenges in sustainable integrated process synthesis and the capabilities of an MINLP process synthesizer MipSyn. Computers & Chemical Engineering 34 (11):1831–1848.
Kravanja Z, Čuček L (2013) Process Synthesis: Minimum Utility Cost with Constrained Matches under Uncertainty, part of Deliverable 9.4 for EFENIS EC FP7 project ENER/FP7/296003/EFENIS 'Efficient Energy Integrated Solutions for Manufacturing Industries'. Faculty of Chemistry and Chemical Engineering, University of Maribor. efenis.uni-pannon.hu/partners-section/deliverables/. Accessed 20.1.2015.
Lam HL, Klemeš JJ, Kravanja Z, Varbanov PS (2011) Software tools overview: process integration, modelling and optimisation for energy saving and pollution reduction. Asia-Pacific Journal of Chemical Engineering 6 (5):696–712.
Lang YD, Biegler LT, Grossmann IE (1988) Simultaneous optimization and heat integration with process simulators. Computers & Chemical Engineering 12 (4):311–327.
Liew PY, Wan Alwi SR, Varbanov PS, Manan ZA, Klemeš JJ (2012) A numerical technique for Total Site sensitivity analysis. Applied Thermal Engineering 40:397–408.
Linnhoff B, Townsend DW, Boland P, Hewitt GF, Thomas BEA, Guy AR, Marsland RH (1982) A User Guide on Process Integration for the Efficient Use of Energy. IChemE, Rugby, UK.
Linnhoff B, Townsend DW, Boland P, Hewitt GF, Thomas BEA, Guy AR, Marsland RH (1994) A User Guide on Process Integration for the Efficient Use of Energy, Revised Edition. IChemE, Rugby, UK.
Liu X-w, Luo X, Ma H-g (2014) Studies on the retrofit of heat exchanger network based on the hybrid genetic algorithm. Applied Thermal Engineering 62 (2):785–790.
López-Maldonado LA, Ponce-Ortega JM, Segovia-Hernández JG (2011) Multiobjective synthesis of heat exchanger networks minimizing the total annual cost and the environmental impact. Applied Thermal Engineering 31 (6–7):1099–1113.
Maréchal F, Kalitventzeff B (1998) Energy integration of industrial sites: tools, methodology and application. Applied Thermal Engineering 18 (11):921–933.
Matsuda K, Hirochi Y, Tatsumi H, Shire T (2009) Applying heat integration total site based pinch technology to a large industrial area in Japan to further improve performance of highly efficient process plants. Energy 34 (10):1687–1692.
Matsuda K, Tanaka S, Endou M, Iiyoshi T (2012) Energy saving study on a large steel plant by total site based pinch technology. Applied Thermal Engineering 43:14–19.
Mizutani FT, Pessoa FLP, Queiroz EM, Hauan S, Grossmann IE (2003) Mathematical Programming Model for Heat-Exchanger Network Synthesis Including Detailed Heat-Exchanger Designs. 2. Network Synthesis. Industrial & Engineering Chemistry Research 42 (17):4019–4027.
Morar M, Agachi PS (2010) Review: Important contributions in development and improvement of the heat integration techniques. Computers & Chemical Engineering 34 (8):1171–1179.
Nemet A (2015) Synthesis of Processes and Process Subsystems for Entire Lifetime. Doctoral dissertation, Faculty of Chemistry and Chemical Engineering, University of Maribor, Maribor, Slovenia dk.um.si/IzpisGradiva.php?id = 47864&lang = eng Accessed 27.7.2015.
Nemet A, Klemeš J, Varbanov P, Mantelli V (2015a) Heat Integration retrofit analysis—an oil refinery case study by Retrofit Tracing Grid Diagram. Front Chem Sci Eng 9 (2):163–182.
Nemet A, Klemeš JJ, Kravanja Z (2013) Optimising entire lifetime economy of heat exchanger networks. Energy 57:222–235.
Nemet A, Klemeš JJ, Kravanja Z (2015b) Designing a Total Site for an entire lifetime under fluctuating utility prices. Computers & Chemical Engineering 72:159–182.
Nemet A, Lee K, Kravanja Z, Klemeš JJ, Varbanov PS, Moon I (2015c) Safety Analysis Embedded in Total Site Synthesis. Chemical Engineering Transactions 45:121–126.
Nguyen T-V, Fülöp TG, Breuhaus P, Elmegaard B (2014) Life performance of oil and gas platforms: Site integration and thermodynamic evaluation. Energy 73:282–301.
Nie XR, Zhu XX (1999) Heat exchanger network retrofit considering pressure drop and heat-transfer enhancement. AIChE Journal 45 (6):1239–1254.
Papoulias SA, Grossmann IE (1983) A structural optimization approach in process synthesis—II: Heat recovery networks. Computers & Chemical Engineering 7 (6):707–721.

Peng F, Cui G (2015) Efficient simultaneous synthesis for heat exchanger network with simulated annealing algorithm. Applied Thermal Engineering 78:136–149.
Pereira HM, Navarro LM, Martins IS (2012) Global Biodiversity Change: The Bad, the Good, and the Unknown. Annual Review of Environment and Resources 37 (1):25–50.
Perry S, Klemeš J, Bulatov I (2008) Integrating waste and renewable energy to reduce the carbon footprint of locally integrated energy sectors. Energy 33 (10):1489–1497.
PIL (2012) i-SteamTM. www.processint.com/chemical-industrial-software/i-steam. Accessed 15. 8.2015.
Rao SS (2009) Engineering Optimization: Theory and Practice, 4th Edition. John Wiley & Sons, Inc., Hoboken, New Yersey, USA.
Rodera H, Bagajewicz MJ (2001) Multipurpose Heat-Exchanger Networks for Heat Integration Across Plants. Industrial & Engineering Chemistry Research 40 (23):5585–5603.
Shang Z, Kokossis A (2004) A transhipment model for the optimisation of steam levels of total site utility system for multiperiod operation. Computers & Chemical Engineering 28 (9):1673–1688.
Shenoy UV (1995) Heat Exchanger Network Synthesis: Process Optimization by Energy and Resource Analysis. Gulf Publishing Company, Houston, Texas, US.
Smith R (1995) Chemical Process Design. McGraw-Hill, New York, USA.
Smith R, Jobson M, Chen L (2010a) Recent development in the retrofit of heat exchanger networks. Applied Thermal Engineering 30 (16):2281–2289.
Smith R, Jobson M, Chen L (2010b) Recent development in the retrofit of heat exchanger networks. Chemical Engineering Transactions 18:27–32.
Smith RM (2005) Chemical Process: Design and Integration. John Wiley & Sons, Ltd., Chichester, West Sussex, England.
Soršak A, Kravanja Z (2002) Simultaneous MINLP synthesis of heat exchanger networks comprising different exchanger types. Computers & Chemical Engineering 26 (4–5):599–615.
Soršak A, Kravanja Z (2004) MINLP retrofit of heat exchanger networks comprising different exchanger types. Computers & Chemical Engineering 28 (1–2):235–251.
Sreepathi BK, Rangaiah GP (2014) Review of Heat Exchanger Network Retrofitting Methodologies and Their Applications. Industrial & Engineering Chemistry Research 53 (28):11205–11220.
STAR (2015). www.ceas.manchester.ac.uk/media/eps/schoolofchemicalengineeringandanalyticalscience/content/researchall/centres/processintegration/STAR.pdf. Accessed 14.8.2015.
Tantimuratha L, Kokossis AC, Müller FU (2000) The heat exchanger network design as a paradigm of technology integration. Applied Thermal Engineering 20 (15–16):1589–1605.
TUHH (2015) Enigma Module. demo.efenis.alexandra.dk/public-portal/tuhh.hta. Accessed 15.8. 2015.
U.S. EIA (2014) How much energy is consumed in the world by each sector? www.eia.gov/tools/faqs/faq.cfm?id=447&t=1. Accessed 19.7.2015.
Van Reisen JLB (2008) A structured approach to heat exchanger network retrofit design. PhD thesis, TU Delft, Delft University of Technology, Delft, The Netherlands. repository.tudelft.nl/view/ir/uuid%3Ac3ce9716-dfa4-478e-a49f-2b61a76e202b/. Accessed 19.8.2015.
Varbanov P, Perry S, Klemeš J, Smith R (2005) Synthesis of industrial utility systems: cost-effective de-carbonisation. Applied Thermal Engineering 25 (7):985–1001.
Varbanov PS (2013) Basic Process Integration Terminology. In: Klemeš JJ (ed) Handbook of process integration (PI): minimisation of energy and water use, waste and emissions. Woodhead Publishing Limited, Cambridge, UK.
Varbanov PS, Klemeš JJ (2011) Integration and management of renewables into Total Sites with variable supply and demand. Computers & Chemical Engineering 35 (9):1815–1826.
Vaskan P, Guillén-Gosálbez G, Jiménez L (2012) Multi-objective design of heat-exchanger networks considering several life cycle impacts using a rigorous MILP-based dimensionality reduction technique. Applied Energy 98:149–161.

Vocciante M, Mantelli V, Aloi N, Dovì V, Reverberi A (2014) Process data reconciliation in the presence of non-uniform measurement errors. Clean Technologies and Environmental Policy 16 (7):1287–1294.

Vujanović A, Pahor B, Nemet A, Kravanja Z (2015) Pinch analysis method for Potential Utilisation of Low-Temperature Heat– An Industrial Case Study Paper presented at the 10th Conference on Sustainable Development of Energy, Water and Environment Systems– SDEWES Dubrovnik, Croatia, 27 September - 3 October, 2015.

Wang Y, Chang C, Feng X (2015) A systematic framework for multi-plants Heat Integration combining Direct and Indirect Heat Integration methods. Energy:doi: 10.106/j.energy.2015.2004.2015.

Wang Y, Pan M, Bulatov I, Smith R, Kim J-K (2012) Application of intensified heat transfer for the retrofit of heat exchanger network. Applied Energy 89 (1):45–59.

Wang Y, Wang W, Feng X (2013) Heat integration across plants considering distance factor. Chemical Engineering Transactions 35:25–30.

XRG Simulation GmbH (2014) SyntHEX. www.xrg-simulation.de/en/products/xrg-applications/synthex. Accessed 14.8.2015.

Yee TF, Grossmann IE (1990) Simultaneous optimization models for heat integration—II. Heat exchanger network synthesis. Computers & Chemical Engineering 14 (10):1165–1184.

Yee TF, Grossmann IE (1991) A screening and optimization approach for the retrofit of heat-exchanger networks. Industrial & Engineering Chemistry Research 30 (1):146–162.

Yee TF, Grossmann IE, Kravanja Z (1990) Simultaneous optimization models for heat integration—III. Process and heat exchanger network optimization. Computers & Chemical Engineering 14 (11):1185–1200.

Improving Energy Efficiency in Batch Plants Through Direct Heat Integration

Pedro M. Castro

Abstract Increasing energy efficiency is an effective way to combat climate change. System optimization methods can help identify energy saving potentials in industry and have been widely applied to continuous plants. In batch plants, however, energy sources and sinks need to coexist in time, further complicating the problem. The main novelty of this paper is to propose a general mixed-integer linear programming formulation for solving the scheduling and heat integration problem simultaneously. It is derived from Generalized Disjunctive Programming and generates the optimal set of heat exchanger matches that minimize utility consumption. Through the solution of a real-life problem from a vegetable oil refinery, we show that there is a trade-off between makespan and utility consumption and that savings up to 39 % can be obtained compared to no heat integration.

Keywords Process integration · Single stage batch plants · Mathematical modeling · Multi-objective optimization · Pareto frontier

1 Introduction

Both the strategic plan of the United States Department of Energy and the European Union Framework Programme for Research and Innovation (Horizon 2020) aim to improve energy productivity by increasing energy efficiency in industry (U.S. Department of Energy 2014; European Commission 2014). In particular, activities should focus on organizational innovation and consider total-site energy management schemes and system optimization methodologies to obtain savings of at least 13 %.

The systems approach for energy efficiency is known as process integration. It includes powerful utility targeting and network design methods that have been widely applied to continuous plants and are either based on pinch analysis (Linnhoff

P.M. Castro (✉)
Centro de Matemática Aplicações Fundamentais e Investigação Operacional, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisbon, Portugal
e-mail: pmcastro@fc.ul.pt

M. Martín (ed.), *Alternative Energy Sources and Technologies*,
DOI 10.1007/978-3-319-28752-2_12

et al. 1982) or mathematical programming (Biegler et al. 1997). Meeting the heating or cooling requirements of process streams can be done by relying on external utilities such as cooling water and steam, or preferably by exchanging heat between a hot process stream at a higher temperature and a cold process stream at a lower temperature. By doing this, energy consumption of external utilities is reduced twice, which in turn lowers the consumption of fossil fuels or biomass that are needed for the production of steam and of electricity that is needed for refrigeration. Essentially, it is one of the easiest and most cost-effective ways to combat climate change.

Energy integration in continuous plants can occur 24 h/day, or as long as the plant is operating. The heat exchanger network is designed for steady-state operation being assumed that the enthalpies of process streams are constant in time. This is no longer the case in multiproduct batch plants, where: (i) production recipes vary between products; (ii) only subtasks of the recipe have heating and cooling requirements, which occur during short time windows of the production time horizon; (iii) customer orders may change frequently.

Three main challenges arise. In terms of modeling, the scheduling and heat integration problems need to be tackled simultaneously so that heating and cooling subtasks belonging to recipes of different products are synchronized (one can also rely on an auxiliary heat transfer medium for energy storage). In terms of operation, the outcome will be a dynamic heat exchanger network that will be much more difficult to control. In terms of economic viability, energy savings will occur over a fraction of the day making it harder to compensate the investment on heat exchangers and associated equipment. It is thus no surprise that the topic of energy integration in batch plants has received considerably less attention in the literature.

A recent review of energy recovery in batch processes can be found in Fernández et al. (2012). We describe next the methods that are closely related to the one being proposed.

In continuous plants, simple graphical methods like composite curves, heat cascades, or linear programming models can be used for minimum utility targeting. Pinch analysis for batch plants involves decoupling scheduling from heat integration decisions. Kemp and MacDonald (1987) first find the optimal schedule and then divide it into time intervals with a constant number of tasks (streams). Standard pinch analysis was then applied to each interval, with the total utility consumption resulting from the sum over all intervals. After analyzing the solutions, the authors noticed that lower values could be obtained after rescheduling, thus realizing the importance of handling both problems simultaneously (Kemp and Deakin 1989). This is done more effectively and efficiently through mathematical programming approaches.

Heat integration in batch plants can either be direct or indirect. The direct mode requires streams to coexist in time. The indirect mode is more flexible in the sense that an auxiliary heat transfer medium can be used to collect energy from the hot stream, wait in a storage tank, and later be used to supply energy to the cold stream. Further utility savings can thus be achieved.

Taking a multipurpose batch plant as an example, Papageorgiou et al. (1994) explored the combination of heat integration and scheduling for both direct and indirect modes. They extended a general discrete-time State Task Network (STN) based formulation to incorporate direct heat integration, leading to a mixed-integer linear programming problem (MILP). On the other hand, indirect heat integration required a detailed characterization of the mass and energy holdups of the heat transfer medium, leading to a non-convex mixed-integer nonlinear programming problem (MINLP). It should be highlighted that the STN features three-stage production recipes with alternative stand-alone and heat-integrated tasks, fully characterized in terms of duration and heat loads in the different hourly segments of the full duration of the task. It was thus not possible to operate somewhere in between the two extreme cases, for instance with variable temperature profiles to be determined by the optimization and resulting from heat transfer with an external utility and possibly multiple process streams, either in sequence or in parallel.

Vaklieva-Bancheva et al. (1996) tackled direct heat integration in a multipurpose batch plant with a continuous-time formulation. Working with nonredundant campaigns, each featuring a subset of products processed in different equipment units, the heat load of each potential heat integration match was determined. The model binary variables determine the optimal matches and campaigns, with each product being limited to at most one match with another product. Similarly to the previous work, it is assumed that the starting times of heat-integrated tasks are synchronized.

Vaklieva-Bancheva et al. (1996) noted that minimizing utility consumption might lead to substantial production delays, possibly even compromising fulfilling customer orders. For a multistage plant with two hot and two cold streams and a maximum of one match per hot/cold stream, Adonyi et al. (2003) have shown that utility consumption with direct heat integration can be reduced considerably by raising the upper bound on the makespan. However, neither work quantified the trade-off between production time and utility consumption.

Halim and Srinivasan (2009) recognized that this is in essence a bi-objective optimization problem with two conflicting objectives. It was tackled by a specialized two-step optimization algorithm that generates a minimum makespan schedule in the first step followed by an integer-cut based stochastic search to generate alternative, near-optimal solutions. In the second step, the full-time horizon was divided into different time intervals, with minimum utility targets calculated for every slot of every generated schedule. Results for the most studied problem involving multipurpose batch plants, have led to the identification of three sets of solutions (after 1000 iterations) with very similar makespans that strangely, also exhibited minimum utility consumption. This was probably caused by the heuristic solution approach since more recently, Seid and Majozi (2014) have reported a solution featuring improvements in both utility consumption and makespan, thus showing that the previous were in fact dominated. However, Seid and Majozi (2014) did not pursue the generation of the Pareto front (Ehrgott 2005).

The model of Seid and Majozi (2014) is perhaps the most general in the sense that it can address both direct and indirect heat integration for the most complex scheduling topology. Majozi (2006), Chen and Chang (2009), and Stamp and Majozi (2011) proposed closely related formulations. In the most recent model, heat integration between tasks can take place during multiple intervals and no longer involves constant temperatures. It is, however, not possible for a stream linked to a production task to be split, thus eliminating parallel heat integration matches from the solution space.

The complexity of a scheduling problem is very much related to the underlying plant topology. It is well known that problems from multipurpose batch plants are in general more difficult to solve than those from single or multistage batch plants (Harjunkoski et al. 2014). General formulations can address both, but tend to perform rather poorly in the latter case, as recently shown by Castro et al. (2014b) for an industrial problem involving sharing of a limiting heating utility (Castro et al. 2002). The challenge is to incorporate difficult constraints concerning the interaction of production tasks into an efficient continuous-time formulation. Fortunately, the modeling burden can be greatly reduced with the help of Generalized Disjunctive Programming (Raman and Grossmann 1994; Grossmann and Trespalacios 2013).

It should be highlighted at this point that computational efficiency is even more important in the context of a multi-objective approach, since multiple problems will have to be solved to generate the Pareto front and quantify the trade-offs between makespan and utility consumption.

Castro et al. (2015) have proposed a continuous-time formulation for tackling the simultaneous scheduling and heat integration problem of a single stage batch plant corresponding to the neutralization stage of a vegetable oil refinery. Two temperature-changing stages are postulated for every hot/cold process stream, being possible for a stream to exchange energy in parallel with streams belonging to other equipment units as well as with the external utility. The bi-objective optimization problem was solved rigorously, with the Pareto optimal solutions showing major differences in product sequencing and heat exchanger networks. A clear trade-off between makespan and utility consumption was observed, with potential savings from heat integration ranging from roughly 16 % for the shortest production time and 18 streams, to above 40 % for the 46 streams problem. Larger savings are, however, linked to significantly longer production times (in one case it more than doubled).

The novelty in this paper is to generalize the model by Castro et al. (2015) from two to any given number of temperature-changing stages per stream. The new model is in fact a relaxation of the previous one and is thus capable of finding better solutions to the problem. The disadvantage is the major increase in computational complexity already for three stages, which might compromise the generation of the Pareto front in a reasonable time and hence may not be translated into a practical advantage with current hardware and software.

The rest of the paper is structured as follows. Section 2 provides the problem definition and is followed by the two modules of the mathematical formulation.

Section 3 focuses on the heat integration part, whereas Sect. 4 deals with the basic scheduling formulation for single stage plants. The two sets of starting and ending time variables are then linked in Sect. 5. The rigorous optimization algorithm for generating the Pareto optimal solutions is given in Sect. 6. Section 7 studies the influence of the number of temperature-changing stages in solution quality and computational time for the single objective, minimum utility problem. The results and comparison of the Pareto optimal solutions are then the subject of Sect. 8, before the conclusions in Sect. 9.

2 Problem Definition

We are given a multiproduct batch plant with a sequential production environment. The production recipe of a particular product $i \in I$ involves several tasks $k \in K_i$ of known duration $p_{i,k}$ [h] to be performed inside a specific equipment unit $m \in M_i$. Associated to some of these tasks is the use of an external heating or cooling utility (e.g., steam, cooling water) to raise or lower the temperature of the vessel's contents (sensible heat) being assumed that no phase change (latent heat) is involved. This is the case of the tasks producing states K2, K4, K6, K8, K10 and K11, in the State Task Network representation of the production recipe of product I1 in Fig. 1.

For the purpose of heat integration, let a temperature change subtask (i, k) be called a process stream ps (there is a 1:1 correspondence and so the indices will be used interchangeably whenever convenient). Process streams can be either hot H or cold C. Given are their initial $t_{\mathrm{ps}}^{\mathrm{in}}$, target temperatures $t_{\mathrm{ps}}^{\mathrm{out}}$ [K] and heat capacity cp_{ps} [J/K]. The goal of heat integration is to replace part of the energy supplied/removed by utilities by exchanging energy directly between a hot and a cold stream. To ensure a good driving force for heat transfer, the minimum temperature difference between a hot and cold stream must be greater than Δt[K].

The main challenge of direct heat integration in batch plants is for the tasks linked to the hot and cold streams to be executed simultaneously. Then, one needs to ensure that the heating/cooling process does not occur too fast or too slow so as to not compromise product quality. Let $\mathrm{ramp}_{\mathrm{ps}} = \left| t_{\mathrm{ps}}^{\mathrm{in}} - t_{\mathrm{ps}}^{\mathrm{out}} \right| / p_{\mathrm{ps}}$ [K/h] be the

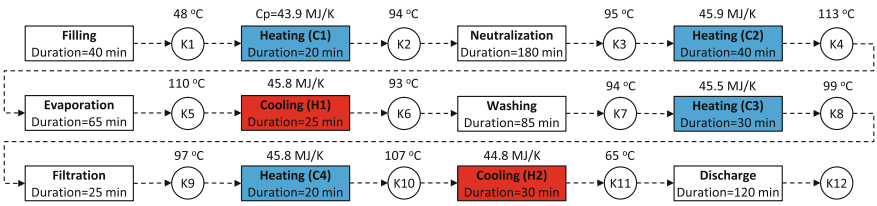


Fig. 1 Production recipe of product I1 showing the subtasks corresponding to hot and cold streams

temperature-changing rate when using external utilities. We assume $\text{ramp}_{\text{ps}}^{\min} = 0.5 \cdot \text{ramp}_{\text{ps}}$ and $\text{ramp}_{\text{ps}}^{\max} = 2 \cdot \text{ramp}_{\text{ps}}$.

Improving energy efficiency in batch plants is in essence a bi-objective optimization problem with two conflicting objectives: (i) minimum energy consumption; (ii) minimum production time (makespan). This is because the optimal heat exchange matches will typically involve major production delays to ensure that the critical hot and cold streams coexist in time. It is thus better to generate the Pareto optimal solutions and then decide on the best trade-off.

3 Heat Integration Model

The model for direct heat integration in batch plants given below adds the timing element to the well-known formulation of Yee et al. (1990) for continuous plants. It assumes that any hot/cold stream goes through a sequence of cooling/heating stages to reach the target temperature. Within each stage, a stream can exchange energy with multiple streams and also with the external utility. The condition of isothermal mixing (all parallel matches start and end at the same temperature) is, however, assumed to avoid the bilinear terms in the energy balances, which would lead to a mixed-integer nonlinear model. Since any number of stages can be postulated, the proposed model can be seen as a generalization of the two-stage model of Castro et al. (2015).

Let $t \in T$ represent a temperature-changing stage of stream ps, taking the temperature from $T_{\text{ps},t-1}$ to $T_{\text{ps},t}$, starting at time $Ti_{\text{ps},t-1}$ and ending at time $Ti_{\text{ps},t}$. Since there is a duration associated to the temperature-changing stage, to be determined by the optimization, it can also be viewed as a time slot. Note that not all variables need to be defined since both the temperature at the start of the first slot ($t_{\text{ps}}^{\text{in}}$) and at the end of the last slot ($t_{\text{ps}}^{\text{out}}$) are known, see Fig. 2. The variables related to the starting time of the first slot (Ts_{ps}) and ending time of the last slot (Te_{ps}) are also named differently to make the connection to the scheduling model easier.

Figure 3 shows the superstructure associated to the cooling of stream H1 belonging to the production recipe of product I1 (recall Fig. 1). Parallel heat exchange matches are possible in every slot t, either with streams belonging to other products (this is the reason why we start with cold stream C5) or the external utility.

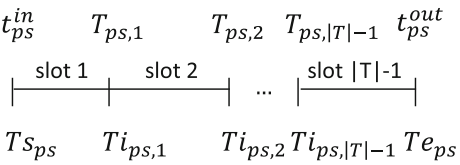


Fig. 2 The heating/cooling stage of a stream is divided into |T| temperature-changing stages that can also be viewed as time slots

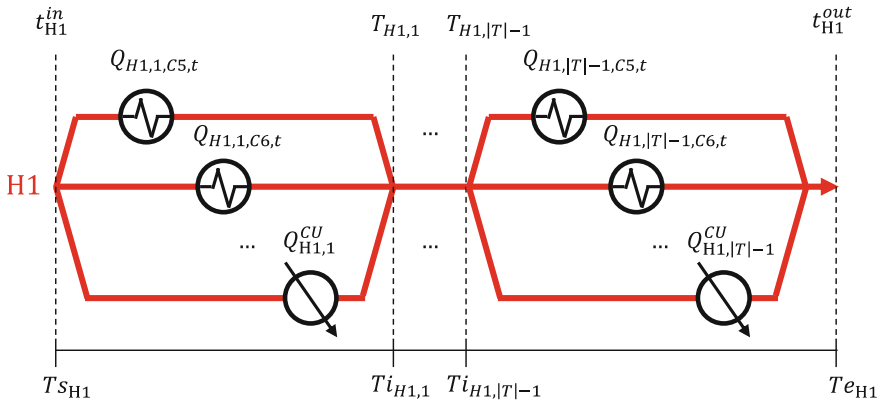


Fig. 3 Superstructure for heat integration between hot stream H1, cold streams not belonging to the same production recipe and external cooling utility

The energy associated to each heat-integrated match is represented by variables $Q_{h,t,c,t'}$ [J], whereas for external cold and hot utilities we use $Q_{h,t}^{\mathrm{CU}}$ and $Q_{c,t}^{\mathrm{HU}}$. Notice that in addition to the two stream indices, two slot indices are required for heat-integrated variables. In a particular slot t, stream H1 can be cooled with any cold stream c on one of its heating stages t'. However, matches between a given pair of hot and cold streams can occur at most once.

3.1 Constraints Featuring Binary Variables

We are now in a position to write as a disjunction (Balas 1979; Raman and Grossmann 1994; Castro and Grossmann 2012) the constraints ensuring: (i) synchronization of hot and cold streams; (ii) temperature driving force above the minimum. Let binary variable $Y_{h,t,c,t'}$ be equal to one if there is a match between hot stream h and cold stream c and $Y_{h,c}^{\mathrm{no}} = 1$ otherwise. These two alternatives are reflected in the two terms in Eq. (1).

If there is a match between streams h and c, it can occur in every postulated pair of slots but at most once, hence the exclusive OR operator over slots t and t' on the left-hand side (LHS). We then need to ensure that the slots' starting and ending time variables are the same. The actual variables involved change from $Ti_{\mathrm{ps},t}$ to Ts_{ps} or Te_{ps} when dealing with the first or last slot. The same can be said for the third constraint (fourth row) that ensures that the lowest temperature difference between the streams, occurring at the end of slots t and t', is higher than the minimum approach temperature Δt. Now, however, the change is from variables $T_{\mathrm{ps},t}$ to parameters $t_{\mathrm{ps}}^{\mathrm{out}}$. On the fifth row, we have the upper bound on energy transfer,

which is the minimum over external utility requirements for h and c, see Eq. (2). Finally, the sixth row sets to zero the energy transfer for all other pairs of slots.

$$\bigvee_{t}\bigvee_{t'}\begin{bmatrix} Y_{h,t,c,t'} \\ Ts_h\big|_{t=1} + Ti_{h,t-1}\big|_{t>1} = Ts_c\big|_{t'=1} + Ti_{c,t'-1}\big|_{t'>1} \\ Ti_{h,t}\big|_{t<|T|} + Te_h\big|_{t=|T|} = Ti_{c,t'}\big|_{t'<|T|} + Te_c\big|_{t'=|T|} \\ T_{h,t}\big|_{t<|T|} + t_h^{\text{out}}\big|_{t=|T|} \geq T_{c,t'}\big|_{t'<|T|} + t_c^{\text{out}}\big|_{t'=|T|} + \Delta t \\ Q_{h,t,c,t'} \leq q_{h,c} \\ Q_{h,t'',c,t'''} = 0 \quad \forall t'' \neq t \vee t''' \neq t' \end{bmatrix}$$

$$\bigvee\begin{bmatrix} Y_{h,c}^{\text{no}} \\ Ts_h \geq 0 \\ Ts_c \geq 0 \\ Q_{h,t,c,t'} = 0 \quad \forall t, t' \end{bmatrix} \quad \forall h, c \tag{1}$$

$$q_{h,c} = \min\left[cp_h\left(t_h^{\text{in}} - t_h^{\text{out}}\right),\ cp_c\left(t_c^{\text{out}} - t_c^{\text{in}}\right)\right] \quad \forall h,\ c \tag{2}$$

The right-hand side (RHS) of Eq. (1) models the no integration case between streams h and c. No constraints apply to the timing (and temperature) variables, whereas all heat exchange variables $Q_{h,t,c,t'}$ become zero. Variables $Y_{h,c}^{\text{no}}$ do not actually need to be a part of the mixed-integer linear programming model, being implicit in Eq. (3) that corresponds to the exclusive OR in Eq. (1).

$$\sum_{t}\sum_{t'} Y_{h,t,c,t'} \leq 1 \quad \forall h,\ c \tag{3}$$

The next step is to reformulate the constraints inside the disjunctive terms of Eq. (1). It can be done using the convex hull and big-M relaxation methods (Balas 1985; Castro and Grossmann 2012). More specifically, we apply the convex hull reformulation for constraints involving the energy variables, leading to Eq. (4), and the big-M reformulation for the others.

$$Q_{h,t,c,t'} \leq q_{h,c} Y_{h,t,c,t'} \quad \forall h,\ t,\ c,\ t' \tag{4}$$

The big-M parameters ensure that the constraint is relaxed if the binary variable is not active. Using values as low as possible can potentially improve the quality of the linear relaxation. In this case, the tightest big-M parameter for the temperature driving force constraints between streams h and c is $bm_{h,c} = -t_h^{\text{out}} + t_c^{\text{out}} + \Delta t$. Since last slot interactions involve only parameters, the MILP constraint can simply be

obtained by multiplying the constraint by the binary variable, Eq. (5). Otherwise, we need to use Eq. (6).

$$Y_{h,|T|,c,|T|}\left(t_h^{\text{out}} - t_c^{\text{out}} - \Delta t\right) \geq 0 \quad \forall h,\ c \tag{5}$$

$$T_{h,t}\big|_{t<|T|} + t_h^{\text{out}}\big|_{t=|T|} \geq T_{c,t'}\big|_{t'<|T|} + t_c^{\text{out}}\big|_{t'=|T|} + \Delta t - bm_{h,c}\left(1 - Y_{h,t,c,t'}\right)$$
$$\forall h,\ c,\ (t<|T| \vee t'<|T|) \tag{6}$$

The timing variables have more freedom than the temperature variables and so a global big-M parameter is used instead. The upper bound on the makespan can be seen as the time horizon of the scheduling problem with a highly conservative value being $hz = \sum_i \sum_k p_{i,k}$. Each equality constraint is converted into two inequality constraints. Furthermore, as noted by Castro et al. (2015), constraints of type $Ti_{h,t} = Ti_{c,t'}$ appear in disjunctive terms (h, t, c, t') and $(h,\ t+1,\ c,\ t'+1)$, which is the reason why two binary variables appear in Eqs. (11)–(12), see also Castro et al. (2014).

$$Ts_h\big|_{t=1} + Ti_{h,t-1}\big|_{t>1} \geq Ts_c\big|_{t'=1} + Ti_{c,t'-1}\big|_{t'>1} - hz\left(1 - Y_{h,t,c,t'}\right)$$
$$\forall h,\ c,\ (t=1 \vee t'=1) \tag{7}$$

$$Ts_h\big|_{t=1} + Ti_{h,t-1}\big|_{t>1} \leq Ts_c\big|_{t'=1} + Ti_{c,t'-1}\big|_{t'>1} + hz\left(1 - Y_{h,t,c,t'}\right)$$
$$\forall h,\ c,\ (t=1 \vee t'=1) \tag{8}$$

$$Ti_{h,t}\big|_{t<|T|} + Te_h\big|_{t=|T|} \geq Ti_{c,t'}\big|_{t'<|T|} + Te_c\big|_{t'=|T|} - hz\left(1 - Y_{h,t,c,t'}\right)$$
$$\forall h,\ c,\ (t=|T| \vee t'=|T|) \tag{9}$$

$$Ti_{h,t}\big|_{t<|T|} + Te_h\big|_{t=|T|} \leq Ti_{c,t'}\big|_{t'<|T|} + Te_c\big|_{t'=|T|} + hz\left(1 - Y_{h,t,c,t'}\right)$$
$$\forall h,\ c,\ (t=|T| \vee t'=|T|) \tag{10}$$

$$Ti_{h,t} \geq Ti_{c,t'} - hz\left(1 - Y_{h,t,c,t'} - Y_{h,t+1,c,t'+1}\right) \quad \forall h,\ c,\ t<|T|,\ t'<|T|) \tag{11}$$

$$Ti_{h,t} \leq Ti_{c,t'} + hz\left(1 - Y_{h,t,c,t'} - Y_{h,t+1,c,t'+1}\right) \quad \forall h,\ c,\ t<|T|,\ t'<|T|) \tag{12}$$

3.2 Linear Constraints

The remaining set of constraints does not involve binary variables and so it is not directly dependent on the particular heat transfer decisions. Let $\text{Temp}_{\text{ps},t}$ and $\text{Time}_{\text{ps},t}$, respectively, represent the temperature and time difference of stream *ps* in slot *t*, see Eqs. (13)–(15). Notice the change in sign on the RHS of Eqs. (13)–(14), which reflects the decrease in temperature of the hot stream and the temperature increase of the cold stream.

$$\text{Temp}_{h,t} = t_h^{\text{in}}\big|_{t=1} + T_{h,t-1}\big|_{t>1} - T_{h,t}\big|_{t<|T|} - t_h^{\text{out}}\big|_{t=|T|} \quad \forall h,\ t \tag{13}$$

$$\text{Temp}_{c,t} = T_{c,t'}\big|_{t'<|T|} + t_c^{\text{out}}\big|_{t'=|T|} - T_{c,t'-1}\big|_{t'>1} - t_c^{\text{in}}\big|_{t'=1} \quad \forall c,\ t \tag{14}$$

$$\text{Time}_{\text{ps},t} = Ti_{h,t}\big|_{t<|T|} + Te_h\big|_{t=|T|} - Ts_h\big|_{t=1} - Ti_{h,t-1}\big|_{t>1} \quad \forall ps,\ t \tag{15}$$

First, we need to verify the energy balances for every stream ps in every slot t. The energy removed from hot stream h is the sum of the energy removed by all cold streams c in their active slots t' plus the energy removed by the external cold utility. On the LHS of Eq. (13), the product of the heat capacity by the slot temperature difference gives the energy removed, whereas the first term on the RHS features the energy variables from Eq. (4). The cold streams energy balances in Eq. (14) are very similar, the difference being that an external hot utility is involved.

$$cp_h \text{Temp}_{h,t} = \sum_c \sum_{t'} Q_{h,t,c,t'} + Q_{h,t}^{\text{CU}} \quad \forall h, t \tag{13}$$

$$cp_c \text{Temp}_{c,t} = \sum_h \sum_t Q_{h,t,c,t'} + Q_{c,t'}^{\text{HU}} \quad \forall c, t' \tag{14}$$

The total utility consumption UT is computed by Eq. (15).

$$\text{UT} = \sum_h \sum_t Q_{c,t}^{\text{CU}} + \sum_c \sum_{t'} Q_{c,t'}^{\text{HU}} \tag{15}$$

Then, the rate of temperature change must be within given upper and lower bounds, Eq. (16).

$$\text{ramp}_{\text{ps}}^{\min} \text{Time}_{\text{ps},t} \le \text{Temp}_{\text{ps},t} \le \text{ramp}_{\text{ps}}^{\max} \text{Time}_{\text{ps},t} \quad \forall ps,\ t \tag{16}$$

Finally, we have the bounds for the intermediate temperature and all timing variables, Eqs. (17)–(20).

$$t_h^{\text{out}} \le T_{h,t} \le t_h^{\text{in}} \quad \forall h,\ t < |T| \tag{17}$$

$$t_c^{\text{in}} \le T_{c,t} \le t_c^{\text{out}} \quad \forall c,\ t < |T| \tag{18}$$

$$Ts_{\text{ps}} \le Ti_{\text{ps},t} \le Te_{\text{ps}} \quad \forall \text{ps},\ t < |T| \tag{19}$$

$$Te_{\text{ps}} \le hz \quad \forall \text{ps} \tag{20}$$

4 Scheduling Model

The heat integration model just presented features timing constraints concerning the interaction of tasks belonging to different products. We now need to sequence products allocated to the same equipment unit and determine the timing of all tasks of the production recipe.

The production facility can be classified as a sequential production environment featuring a single stage batch plant with parallel units and so the recommended models are of the continuous type (Harjunkoski et al. 2014). Amongst the available alternatives, the general precedence model of Méndez et al. (2001) is perhaps the easiest to understand and features timing variables with task and product indices that can be linked straightforwardly to the heat integration module (Castro et al. 2015). It is thus chosen for product scheduling.

Given that product-unit assignments are known a priori, the 0–1 decisions are limited to sequencing. Let $X_{i,i'} = 1$ if product i is produced before $i' > i$. More specifically, if the sequence is I1–I3–I2, then I1 generally precedes I2 and I3, while I3 precedes I2, i.e., $X_{1,2} = 1$, $X_{1,3} = 1$ and $X_{2,3} = 0$. Now, let $S_{i,k}$ and $E_{i,k}$ be nonnegative continuous variables representing the starting and ending time of execution of task k for product i. Production of i includes execution of all tasks ranging from $k = 1$ to the last task of the recipe $k = |K_i|$. Thus, if $X_{i,i'} = 1$, then the starting time of the first task of i' must be greater than the ending time of the last task of i, see Fig. 4. Alternatively, if $X_{i,i'} = 0$, the reverse is true, leading to the disjunction in Eq. (21).

$$\begin{bmatrix} X_{i,i'} \\ E_{i,|K_i|} \le S_{i',1} \end{bmatrix} \veebar \begin{bmatrix} \neg X_{i,i'} \\ E_{i',|K_{i'}|} \le S_{i,1} \end{bmatrix} \quad \forall i,\ i' > i,\ m \in M_i \cap M_{i'} \tag{21}$$

The big-M reformulation of Eq. (21) leads to two sets of timing constraints.

$$E_{i,|K_i|} \le S_{i',1} + hz(1 - X_{i,i'}) \quad \forall i, i' > i, m \in M_i \cap M_{i'} \tag{22}$$

$$E_{i',|K_{i'}|} \le S_{i,1} + hz \cdot X_{i,i'} \quad \forall i, i' > i, m \in M_i \cap M_{i'} \tag{23}$$

It remains to relate the timing variables of tasks belonging to the same production recipe. The ending time is equal to the starting time plus the duration,

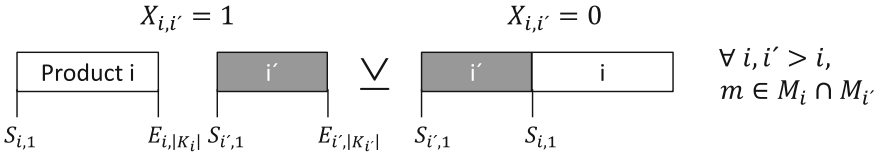


Fig. 4 Given any two products produced on the same equipment unit, one generally precedes the other

Eq. (24). Then, Eq. (25) states that the starting time of task $k + 1$ is equal to the ending time of k. It should be highlighted that the use of the inequality provides more flexibility for heat integration but it is not considered since excessive waiting times would lead to temperature decrease (heat losses), which would be unaccounted for. The maximum over the ending times of all products defines the makespan MK.

$$E_{i,k} = S_{i,k} + p_{i,k} \quad \forall i,\ k \in K_i \tag{24}$$

$$S_{i,k+1} = E_{i,k} \quad \forall i,\ k \in K_i \backslash |K_i| \tag{25}$$

$$\mathrm{MK} \geq E_{i,|K_i|} \quad \forall i \tag{26}$$

5 Linking the Two Models

The separation in heat integration and scheduling is a result of the wide variety of production environments. As it is, instead of working with a highly complex model, all we need is to relate the timing variables of streams *ps* with the timing variables of their corresponding (*i, k*) tasks. For the general precedence model described in the previous section, the two sets of equations needed are simply:

$$Ts_{\mathrm{ps}} = S_{\mathrm{ps}} \quad \forall \mathrm{ps} \tag{27}$$

$$Te_{\mathrm{ps}} = E_{\mathrm{ps}} \quad \forall \mathrm{ps} \tag{28}$$

The two modules can also be viewed as belonging to different time scales. On a higher level, we have the scheduling layer defining the time window for executing a

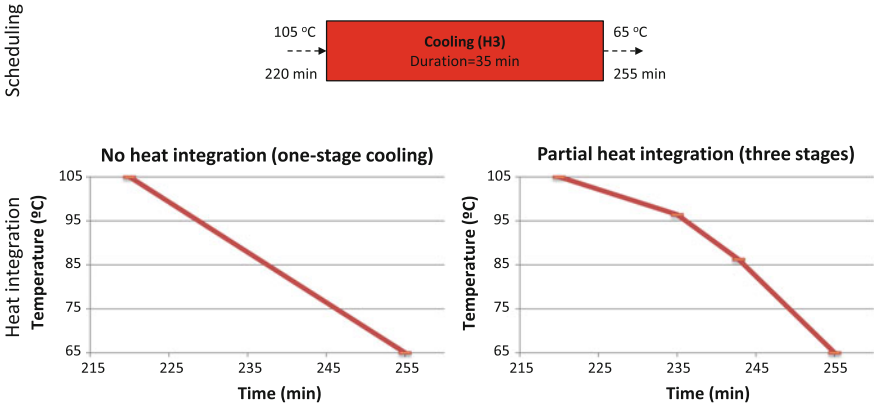


Fig. 5 The complete model is divided into scheduling and heat integration layers. Given a product's schedule, the lower level heat integration layer can select different temperature profiles

heating/cooling task. On a lower layer, we have heat integration, potentially dividing such window in a number of shorter, temperature-changing slots, see Fig. 5. In a way, it can be viewed as determining the set points of a temperature control system that ensures that the task is executed as planned. The proposed framework thus falls under the scope of the integration of scheduling and control (Harjunkoski et al. 2009).

6 Multi-objective Optimization

The main goal of heat integration is to reduce energy consumption from external utilities. If one is dealing with process design for the steady-state operation of a continuous plant, the design of the heat exchanger network is traditionally undertaken after making the decisions for production rates (Douglas 1988; Smith 1995; Biegler et al. 1997). In batch plants, however, a batch involves the production of a certain amount of material, with the production rate over a given time period being dependent on the time it takes to complete all products on the different equipment units. Since direct heat integration in batch plants requires the synchronization of production tasks, a heating/cooling task might have to wait for its counterpart, possibly leading to major delays in the case of minimum utility consumption. At the other extreme, one can still have heat integration without compromising the makespan.

The trade-off between utility consumption UT and production time MK can be captured through the following bi-objective optimization problem (**BO**).

$$\begin{aligned} &\min(\mathrm{UT},\ \mathrm{MK}) \\ &\text{subject to Eqs. } (2)-(20),\ (22)-(28) \end{aligned} \tag{BO}$$

The solution of (**BO**) is given by a set of Pareto optimal points that can be generated by the ε-constraint method (Ehrgott 2005; Guillén-Gosálbez et al. 2010). Due to the discrete nature of the processing times, we choose to associate parameter ε with the makespan variable UT, acting as an upper bound in the single objective optimization problem (**SO**).

$$\begin{aligned} \min \quad & \mathrm{UT} \\ \text{s.t.} \quad & \mathrm{MK} \leq \varepsilon \\ & \text{Eqs. } (2)-(20), (22)-(28) \end{aligned} \tag{SO}$$

The set of Pareto optimal points can be determined by varying ε over a range of values $[\underline{\varepsilon}, \bar{\varepsilon}]$ and checking utility consumption. As we increase ε, it can either remain the same or decrease. The former case corresponds to a dominated solution, being worse than the previous due to a longer makespan. The latter case belongs to the Pareto set and is the one that we are interested in.

The challenge of generating the Pareto set with a continuous-time scheduling formulation allowing for multiple heat integration stages, is that there are regions where minimum utility consumption varies continuously with ε. Interestingly, and even though no exhaustive search was performed, this behavior was not observed when tackling the current industrial problem with the two-stage model (Castro et al. 2015). Since it would be impractical to rely on a continuous domain for ε, we stick with our recent approach and consider a discrete domain instead. More precisely, $\varepsilon \in \{\underline{\varepsilon}, \underline{\varepsilon}+\delta, \underline{\varepsilon}+2\delta, \ldots, \bar{\varepsilon}\}$ with δ being chosen as the greatest common factor of the processing times, 5 min. As a consequence, the outcome of solving multiple instances of (**SO**) is a subset of the Pareto optimal solutions.

It still remains to determine the values of parameters $\underline{\varepsilon}$ and $\bar{\varepsilon}$. The lower bound for ε is the minimum makespan that can be obtained by solving optimization problem (**LEPS**). Notice that this is a much smaller problem than (**SO**) since the constraints associated with heat integration, and linking heat integration and scheduling, are not needed for feasibility and do not affect the makespan.

$$\begin{aligned} \underline{\varepsilon} := \min \mathrm{MK} \\ \text{s.t. Eqs. } (22)-(26) \end{aligned} \tag{LEPS}$$

The upper bound is then the makespan obtained after solving the minimum utility problem (**SO**) without the first constraint. More specifically, $\bar{\varepsilon} := \mathrm{MK}(x^*, y^*)$, where (x^*, y^*) is the set of points of the given argument in terms of continuous variables x and binary variables y, for which the utility function attains its minimum value (**UEPS**).

$$\begin{aligned} (x^*, y^*) := \arg\min_{x,y} \mathrm{UT}(x, y) \\ \text{s.t. Eqs. } (2)-(20), (22)-(28) \end{aligned} \tag{UEPS}$$

Overall, bi-objective problem (**BO**) is tackled by solving a sequence of single objective optimization problems. The total computational time needed to generate the Pareto optimal solutions includes the time to solve (**LEPS**) + (**UEPS**) plus $\frac{\bar{\varepsilon}-\underline{\varepsilon}}{\delta}+1$ (**SO**) problems.

It should be noted that the left-most point on the Pareto front is obtained by solving the first (**SO**) problem, with $\varepsilon = \underline{\varepsilon}$, giving the lowest utility consumption for the minimum makespan. An alternative would be to solve min UT s.t. Eqs. (2)–(20) with fixed timing variables Ts_{ps} and Te_{ps}, following the solution of (**LEPS**). This exactly corresponds to the sequential approach is which one first solves the scheduling and subsequently the heat integration problem.

6.1 Computational Environment

The mixed-integer linear programming mathematical formulations and the algorithm for generating the Pareto set were implemented in GAMS. The MILP problems were solved in CPLEX 12.6 running in parallel deterministic mode using up to 8 threads and a relative optimality tolerance = 10^{-6}. The hardware consisted of a desktop with an Intel i7-4790 CPU running at 3.6 GHz, with 8 GB of RAM, solid-state drive and running Windows 7, 64-bit operating system.

7 Choosing the Number of Temperature-Changing Stages

The number of temperature-changing stages (time slots) acts as a hidden model constraint. The larger the number of postulated stages $|T|$, the larger the number of heat integration possibilities and the lower the utility consumption. In other words, solution quality has the potential to improve with every single increment in $|T|$, a well-known characteristic of slot-based scheduling formulations (Harjunkoski et al. 2014). The drawback is that problem size, and hence computational effort, also increases with $|T|$, making it unlikely to go beyond a few slots. To overcome this challenge, we can rely on the classic heuristic search procedure that iterates over $|T|$ until no improvement is observed in the objective function or the problem becomes intractable.

Consider the smallest problem from Castro et al. (2015) as illustrative example (P18S). It features 18 streams, 5 hot and 13 cold, and 4 products to be produced in 3 parallel equipment units. Data for the other products is similar to the one given in the production recipe for product I1 in Fig. 1. It shows that the oils need to be heated from roughly 50 to 110 °C and cooled from 110 to 65 °C in different phases. Due to a similar temperature range and the temperature driving force constraint in Eq. (5), making $|T| = 1$ is not enough to allow for heat integration (UT = 13715.3 MJ). Solving the makespan-unconstrained minimum utility problem (**UEPS**) for $|T| = 2$ leads to utility savings of 33.6 % that increase slightly to 34.0 % for $|T| = 3$. No further improvements are observed for $|T| = 4$ and so the search is stopped. Notice in Table 1 the steep increase in computational time with the number of stages.

Problem P26S features 6 products in the same 3 units, with recipes leading to the definition of 7 hot and 19 cold streams (Castro et al. 2015). The results in Table 2 show that the relative utility savings are higher and so are the benefits of using three

Table 1 Influence of number of temperature-changing stages in utility savings due to heat integration and computational time for problem P18S

Stages $\|T\|$	1	2	3	4
Utility savings (%)	0	33.6	34.0	34.0
Computational time (CPUs)	0.12	0.73	53.0	1558

Table 2 Influence of number of temperature-changing stages for problem P26S (UT = 16061.8 MJ for no integration)

Stages $\|T\|$	1	2	3
Utility savings (%)	0	37.7	39.2
Computational time (CPUs)	0.11	18.4	3600[a]

[a]Maximum computational time

rather than two stages. The main disadvantage is that the optimization problem is considerably more difficult to solve, being the MILP solver unable to prove optimality for $|T| = 3$ up to one hour of computational time (optimality gap = 30.7 %), whereas $|T| = 2$ can be solved in under a minute.

The makespan-constrained problem (**SO**) can be solved much faster than its unconstrained counterpart (**UEPS**). Making $\varepsilon = \overline{\varepsilon} = 805$ min, leads to a computational time of just 1.73 CPUs for P26S and $|T| = 3$, corresponding to savings of 18.2 % versus 15.5 % for $|T| = 2$. The optimal schedules and heat exchange matches for these two cases are given in Fig. 6. The schedules are essentially the same, with production sequences of I2–I3 in unit M2 and I4–I5–I6 in M3. Still, unit M3 starts earlier on the right to extend the time C14 is exchanging energy with H3, a stream belonging to product I2 that is made in bottleneck unit M2. For that to be possible, hot stream H3 needs to be cooled with an external utility both before (from 220 to 235 min) and after (242.8–255 min) heat integration. Additional

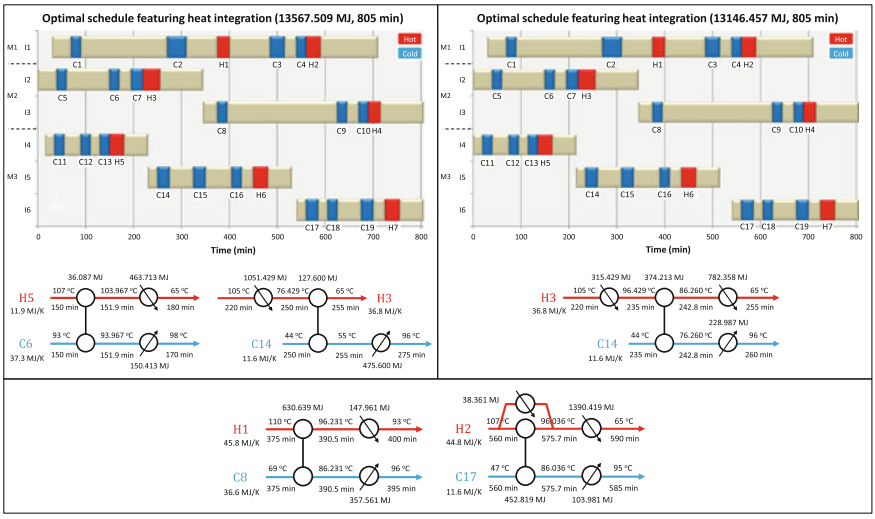


Fig. 6 Optimal solutions for problem P26S assuming minimum makespan and either two (on the *left*) or three temperature-changing stages (on the *right*). Matches H1–C8 and H2–C17 are shared by the two networks while non heat-integrated streams are omitted for simplification

savings of 210.526 MJ are thus obtained using the exact same number of heat exchangers. Stream H3 is actually the only one taking advantage of the three temperature-changing stages. Notice also that stream H2 features two parallel matches in the first stage, one with stream C17 and another with the external utility, and that the temperature driving force of 10 °C is active in all heat-integrated matches.

8 Makespan Versus Utility Consumption

The set of Pareto optimal points provides solutions to the bi-objective problem that achieve different trade-offs between productivity and energy efficiency concerns. The left-most point corresponds to production as fast as possible, for which important utility savings can still be achieved. Delaying processing tasks so as to ensure a perfect synchronization between hot and cold process streams can further increase utility savings, ultimately leading to minimum utility consumption in the right-most point. This solution may, however, be of limited practical interest in cases of significantly longer production times. It is thus expected that the decision-maker will select a point located in the left region of the Pareto chart.

Let us consider the Pareto optimal solutions for problem P18S for illustration purposes (Fig. 7). For the minimum makespan of 805 min, the minimum utility consumption compared to no integration is equal to 84.4 %. Interestingly, with just an half an hour delay in production time, utility consumption can be reduced to 77.1 %, making this a potentially better solution when considering three stages. This value then remains constant until the next Pareto optimal solution at 940 min (74.4 %). The next interesting solution is perhaps (955 min; 70.6 %), with the fourth and last being the one corresponding to the maximum savings of 34 % at 1140 min.

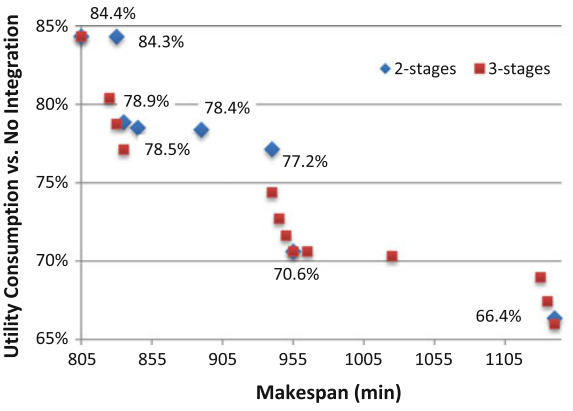


Fig. 7 Pareto optimal solutions for problem P18S considering two (labels shown) and three temperature-changing stages

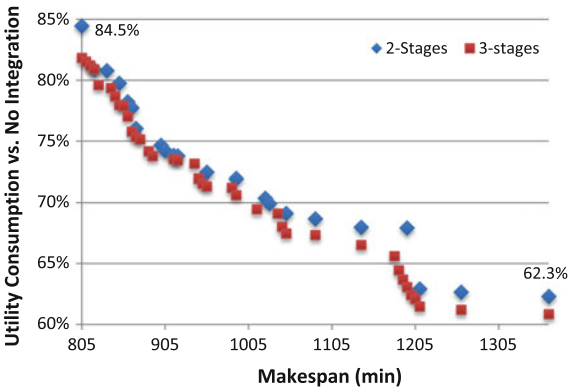


Fig. 8 Pareto optimal solutions for problem P26S

Table 3 Total computational time (CPUs) required for generating the Pareto set

Problem/Stages $\|T\|$	2	3
P18S	29.5	927
P26S	463	202,652
P33S	171,971	–

As mentioned earlier, problems resulting from the assumption of three temperature-changing stages are a relaxation of their two-stage counterparts. Utility consumption is thus never larger provided that the MILP problems are solved to optimality. What is interesting is that only two Pareto points are shared, (805; 84.4 %) and (955; 70.6 %), with the three-stage approach leading to a total of 13 versus 8 points.

The Pareto frontier for the 26 streams problem in Fig. 8 is somewhat smoother. The curves for two and three stages are closer together and feature significant more optimal points (22 and 38, respectively). The most important differences occur for the minimum makespan of 805 min (84.5 vs. 81.8 %) and for 1195 min (67.9 % vs. 63.1 %). Overall, since the trend is essentially the same and it takes much longer to generate the Pareto set for three stages (the total computational time is actually four times longer than the minimum makespan, meaning that one would not have the time to wait for the solution, see Table 3), it is perhaps more reasonable to rely on two stages. Furthermore, increasing the number of stages may potentially create more low duration matches (e.g., 1 min), which by being difficult to implement in practice, decreases the value of the solution for decision making.

8.1 Influence on Optimal Schedule and Heat Exchanger Network

A Pareto optimal solution gives the makespan and utility consumption but also: (i) product sequencing within an equipment unit; (ii) timing of each subtask of the

production recipe; (iii) heat exchange matches between pairs of hot and cold process streams and between process streams and utilities. A detailed analysis of the actual schedule and heat exchanger network provides additional valuable information that is hidden in the Pareto frontier of Fig. 8, which may help deciding about the best trade-off between makespan and energy efficiency.

In Fig. 9, we show the schedule and network for two interesting solutions of the bi-objective optimization problem, corresponding to (890 min, 73.8 %) and (1050 min, 67.5 %). These are points after which the slope of the Pareto curve decreases significantly. Note that the minimum makespan solution under the assumption of three heating/cooling stages (Fig. 6 on the right) was already discussed.

Binary variables $X_{i,i'}$ with $i' > i$ define the product sequence. The 26-streams problem features three equipment units with six preassigned products, I1 in M1, {I2, I3} in M2 and {I4, I5, I6} in M3. As can be seen in the left schedule of Fig. 9, the optimal sequences are I3–I2 in M2 and I5–I6–I4 in M3, corresponding to $X_{I2,I3} = 0$, $X_{I4,I5} = 0$, $X_{I4,I6} = 0$ and $X_{I5,I6} = 1$. Concerning the network, one hot (H7) and two cold streams (C5 and C11) take advantage of the three postulated stages.

Two aspects are worth highlighting. The first is the very complex subnetwork involving streams H7, C5 and H1 that requires C5 to be synchronized with H7 in the first stage and with H1 in the second stage, both executed in parallel with hot utility matches that continue in stage three. The second aspect is the very low duration (0.8 min) of the last heating stage of C11, which increases the temperature from 94.068 to 95 °C. Thus, such solution would be very difficult to implement in

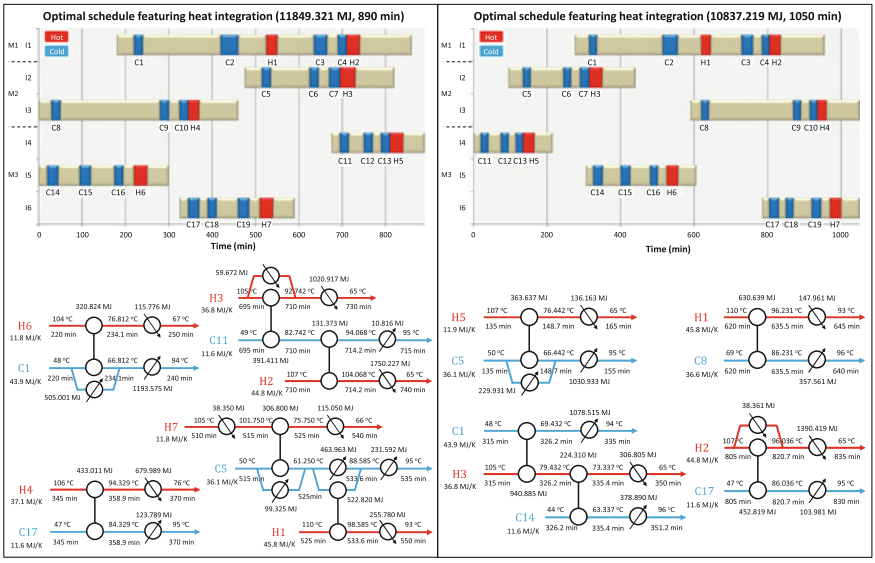


Fig. 9 Two interesting Pareto optimal solutions for problem P26S

practice, much more so than the solutions for 805 and 1050 min, given respectively in Fig. 6 and on the right of Fig. 9.

Increasing the makespan to 1050 min (30 % longer than the fastest production time) leads to the exact same product sequence as for 805 min, I2–I3 in M2 and I4–I5–I6 in M3, see Fig. 9. Active $X_{i,i'}$ variables for unit M3 are $X_{\mathrm{I4,I5}} = 1$, $X_{\mathrm{I4,I6}} = 1$ and $X_{\mathrm{I5,I6}} = 1$. Notice that since these are general precedence variables and so I4 is produced before I6 despite I5 being made in between. It should also be highlighted that solution degeneracy is one the reasons why the three-stage model is considerable more difficult to solve. As an example, stream H1 requires two cooling stages, with stages two and three being active in the optimal solution ($Y_{\mathrm{H1,2,C8,2}} = 1$, $Q^{\mathrm{CU}}_{\mathrm{H1,3}} = 147.961\,\mathrm{MJ}$). However, the exact same value of the objective function would be obtained with cooling in stages one and two or one and three. The same is true for its interacting cold stream C8 that ends its heating in stage three with $Q^{\mathrm{HU}}_{\mathrm{C8,3}} = 357.561\,\mathrm{MJ}$. As the number of streams increase so does the number of combinations.

9 Conclusions

This paper has proposed a framework for improving energy efficiency in batch plants. Utility consumption is reduced by means of direct heat integration between hot and cold streams linked to specific subtasks of the production recipes for the different products. The challenge has been to derive a mathematical model capable of synchronizing integrated streams while considering the most relevant operational alternatives. Compared to our recent work, the novelty has been to postulate for each stream a superstructure with multiple, rather than two, temperature-changing stages, resulting in better solutions. In each heating/cooling stage, a cold/hot stream can exchange energy in parallel with every other hot/cold stream (not belonging to the same production recipe) and the external hot/cold utility. Simplifying assumptions include isothermal mixing in parallel matches to avoid generating a MINLP problem and at most one match between any given pair of hot and cold process streams.

Through the solution of a real industrial case study from a vegetable oil refinery, we have seen that the minimum utility targeting and design problem in batch plants is in essence a bi-objective optimization problem. While potential savings of 39 % can be obtained compared to no heat integration, major delays are required to achieve the optimal synchronization of production tasks. If the aim is to not compromise productivity, savings of 15.6 % are still possible for the smallest problem considered. In between these two extreme cases there were a variety of Pareto optimal solutions that could be selected by the decision maker.

Finally, it should be emphasized that the proposed model for simultaneous scheduling and direct heat integration will need to be adapted for production environments other than single stage batch plants. The way the complete model was

presented, divided into heat integration and scheduling modules with linking constraints, makes it easier to replace the scheduling module with one more appropriate for a network topology, which tend to be not as efficient computationally. Future work will also look at further improving energy efficiency by means of indirect (nonsynchronous) heat integration, which involves energy storage.

Acknowledgments The author wishes to acknowledge Bruno Custódio for collecting and providing the industrial plant data as well as financial support from Fundação para a Ciência e Tecnologia (FCT) through the Investigador FCT 2013 program and project PTDC/EQU-ESI/118253/2010.

References

Adonyi, R., Romero, J., Puigjaner, L., Friedler, F. (2003). Incorporating Heat Integration in Batch Process Scheduling. Applied Thermal Engineering 23, 1743–1762.

Balas, E. (1979). Disjunctive Programming. Annals of Discrete Mathematics 5, 3–51.

Balas, E. (1985). Disjunctive Programming and a Hierarchy of Relaxations for Discrete Optimization Problems. SIAM Journal of Algebraic and Discrete Mathematics 6, 466–486.

Biegler, L.T, Grossmann, I.E., Westerber, A.W. (1997). Systematic Methods of Chemical Process Design. Prentice Hall: Upper Saddle River, NY, USA.

Castro, P.M., Custódio, B., Matos, H.A. (2015). Optimal Scheduling of Single Stage Batch Plants with Direct Heat Integration. Comput. Chem. Eng. 82, 172–185.

Castro, P.M., Grossmann, I.E. (2012). Generalized Disjunctive Programming as a Systematic Modeling Framework to Derive Scheduling Formulations. Ind. Eng. Chem. Res. 51, 5781–5792.

Castro, P.M., Grossmann, I.E., Veldhuizen, P., Esplin, D. (2014). Optimal Maintenance Scheduling of a Gas Engine Power Plant using Generalized Disjunctive Programming. AIChE J. 60, 2083–2097.

Castro, P., Matos, H., Barbosa-Póvoa, A.P.F.D. (2002). Dynamic modeling and scheduling of an industrial batch system. Comp. Chem. Eng. 26, 671–686.

Castro, P.M., Rodrigues, D., Matos, H.A. (2014b). Cyclic scheduling of pulp digesters with integrated heating tasks. Ind. Eng. Chem. Res. 53, 17098–17111.

Chen, C.L., Chang, C.-Y. (2009). A resource-task network approach for optimal short-term/periodic scheduling and heat integration in multipurpose batch plants. App. Therm. Eng. 29, 1195–1208.

Douglas, J.M. (1988). Conceptual Design of Chemical Processes. McGraw-Hill: New York, NY, USA.

Ehrgott, M. (2005). Multicretira Optimization. Springer Berlin: Heidelberg, Germany.

European Commission (2014). Horizon 2020: Work Programme 2014–2015. http://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/main/h2020-wp1415-energy_en.pdf (accessed October 20, 2014).

Fernández, I. Renedo, C.J., Pérez, S.F., Ortiz, A., Mañana, M. (2012). A review: Energy recovery in batch processes. Renewable and Sustainable Energy Reviews 16, 2260–2277.

Grossmann, I.E, Trespalacios, F. (2013). Systematic Modeling of Discrete-Continuous Optimization Models through Generalized Disjunctive Programming. AIChE J. 59, 3276–3295.

Guillén-Gosálbez, G., Mele, F.D., Grossmann, I.E. (2010). A Bi-criterion optimization approach for the design and planning of hydrogen supply chains for vehicle use. AIChE J. 56, 650–667.

Halim, I., Srinivasan, R. (2009). Sequential methodology for scheduling of heat-integrated batch plants. Ind. Eng. Chem. Res. 48, 8551–8565.
Harjunkoski, I., Nystrom, R., Horch A. (2009). Integration of scheduling and control- Theory or practice? Comput. Chem. Eng. 33, 1909–1918.
Harjunkoski, I., Maravelias, C., Bongers P., Castro P.M., Engell, S., Grossmann, I., Hooker, J., Méndez, C., Sand, G., Wassick, J. (2014). Scope for Industrial Applications of Production Scheduling Models and Solution Methods. Comput. Chem. Eng. 62, 161–193.
Kemp, I.C., Deakin, A.W. (1989). The cascade analysis for energy process integration of batch process. Part 2. Network design and process scheduling. Chem. Eng. Res. Des. 67, 510–516.
Kemp, I.C., MacDonald, E.K. (1987). Energy and process integration in continuous and batch processes. Innovation in process energy utilization. IChemE Symp Series 105, 185–200.
Linnhoff, B., Townsend, D.W., Boland, D., Hewitt, G.F., Thomas, B.E.A., Guy, A.R., Marsland, R.H. (1982). A User Guide on Process Integration for the Efficient Use of Energy. The Institution of Chemical Engineers: Rugby, England.
Majozi, T. (2006). Heat integration of multipurpose batch plants using a continuous-time framework. Appl. Therm. Eng. 26, 1369–1377.
Méndez, C.A., Henning, G.P., Cerdá, J. (2001). An MILP continuous-time approach to short-term scheduling of resource-constrained multistage flowshop batch facilities. Comp. Chem. Eng. 25, 701.
Papageorgiou, L.G., Shah, N., Pantelides, C.C. (1994). Optimal Scheduling of Heat-Integrated Multipurpose Plants. Ind. Eng. Chem. Res. 33, 3168–3186.
Raman, R., Grossmann, I.E. (1994). Modeling and computational techniques for logic based integer programming. Computers & Chemical Engineering, 18, 563–578.
Seid, E.R., Majozi, T. (2014) Heat integration in multipurpose batch plants using a robust scheduling framework. Energy 71, 302–320.
Smith, R. (1995). Chemical Process Design. McGraw-Hill: New York, NY, USA.
Stamp, J.D., Majozi, T. (2011). Optimum heat storage design for heat integrated multipurpose batch plants. Energy 36, 5119–5131.
U.S. Department of Energy (2014). Strategic Plan: 2014–2018. http://www.energy.gov/sites/prod/files/2014/04/f14/2014_dept_energy_strategic_plan.pdf (accessed October 20, 2014).
Vaklieva-Bancheva, N., Ivanonv, B.B., Shah, N., Pantelides, C.C. (1996). Heat exchanger network design for multipurpose batch plants. Comp. Chem. Eng. 20, 989–1001.
Yee, T.F., Grossmann, I.E., Kravanja, Z. (1990). Simultaneous Optimization Models for Heat Integration-1. Area and Energy Targeting and Modeling of Multistream Heat Exchangers. Comput. Chem. Eng. 14(10), 1165–1184.

Life Cycle Algal Biorefinery Design

Jian Gong and Fengqi You

Abstract Life cycle analysis (LCA) is a trending methodology for evaluating environmental impacts associated with a product through its life cycle. This chapter presents the general model formulation for life cycle optimization, as well as systematic approaches to solve the resulting multi-objective mixed-integer nonlinear programming problems. We discuss the life cycle design of an algal biorefinery for the production of sustainable biofuels and profitable by-products. We review the existing articles for the simulation and optimization of the algal biorefinery processes. We show the importance of superstructure optimization, and introduce the technology alternatives that can be incorporated into an algal process superstructure. Subsequently, we show the optimization results of an algal biorefinery. We further summarize the challenges and possible directions for future life cycle algal biorefinery designs.

1 Introduction

The development of the chemical industry in the last 150 years has greatly shaped our society, but the side effects are adversely influencing the environment we live in and severely degrading the ecological foundation of all activities (Bakshi 2014). The future growth, however, hinges on the effective solutions to the competition with emerging impacts from a variety of sustainable factors, such as materials availability, climate change mitigation, and inherent security. Therefore, it is of significance to consider these impacts throughout the decision-making process via establishing quantitative metrics to measure the level of sustainability for intended processes, and employing efficient approaches to assist the selection of optimal sustainable processes.

J. Gong · F. You (✉)
Department of Chemical and Biological Engineering, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA
e-mail: you@northwestern.edu

M. Martín (ed.), *Alternative Energy Sources and Technologies*,
DOI 10.1007/978-3-319-28752-2_13

This chapter shows the superstructure-based method for sustainable process design and its application to life cycle design of an algal biorefinery. In the following sections, we first introduce a general modeling framework for sustainable process design synthesis. Next, we review the current progress of life cycle design of an algal biorefinery. Later, we briefly describe the methods in process design and synthesis, and provide a large number of technology alternatives that could be potentially considered in the six sections of an algal biorefinery process. Then typical results of a life cycle algal biorefinery design are shown. Finally, we discuss the challenges and opportunities for future development.

2 A Framework for Sustainable Process Design and Synthesis

2.1 Life Cycle Analysis

Based on an emerging methodology named life cycle analysis (LCA), there is a growing number of publications addressing the design of sustainable processes and supply chains (Yue et al. 2014b; Garcia and You). LCA can be used to estimate and assess the environmental impacts attributable to the life cycle of a product or a service (Rebitzer et al. 2004). Early work of LCA can date back to the late 1990s and the consensus of LCA results in the ISO 14000 series (Institution 2006). After years of improvement, LCA becomes a comprehensive methodology and considers all attributes of natural environment, human health, and resources. As an evolving methodology, the recent development in LCA involves attributional and consequential LCA, input–output LCA, and hybrid LCA (Finnveden et al. 2009). Although effort is made in developing social LCA (Grießhammer et al. 2006), most attention has been put on thriving and applying environmental LCA.

An LCA study encompasses four phases (Institution 2006): *goal and scope definition*, *life cycle inventory (LCI) analysis*, *life cycle impact assessment*, and *life cycle interpretation*. *Goal and scope definition* elucidates essential information of the study, including the purpose of the study, the intended application and audience, the system boundary of the analysis, and the functional unit to quantify the function of the final product or service. In the next phase, LCI is investigated, which specifies the amounts of input and output materials, emissions, and energy with respect to the functional unit. In the impact assessment phase, the LCI is converted to various indicators, which provide the basis for analyzing the contributions of individual LCI entries to a number of environmental impacts. In the last phase, the results of inventory analysis and impact assessment are explained, and uncertainty analysis and sensitivity analysis can be conducted. And conclusions are made after all the targets are met.

Integrating LCA with process models and techno-economic analysis, researchers are now able to develop sustainable processes and supply chains with respect to a variety of environmental impact indicators of interest (You and Wang 2011).

2.2 *Life Cycle Optimization*

Most process design problems in the past decades were focused solely on improving the overall economic performance. Minimizing the total annualized cost and maximizing the net present value are prevalent objective functions in the mathematical models if the superstructure optimization approach is employed. A typical structure of such mathematical models contains mass balance constraints, energy balance constraints, and economic evaluation constraints. Accordingly, only one optimal solution will be returned if the problem is solved. However, sustainability is made up of three pillars: economy, society, and environment (Finnveden et al. 2009). The economically optimal solution may render undesirable social and environmental performance. From a sustainability point of view, such optimal solutions from single-objective formulations are unsurprisingly unfavorable, because significant impacts on other aspects of sustainability may cause irreversible consequences for the future development of mankind and the environment we are dwelling in.

LCA is a systematic tool for evaluating the environmental impacts of a fixed process, but such analyses require pre-defined processes and systems, therefore lacking the ability to automatically generate feasible process designs from a process superstructure and identify the best process performance. To bridge this research gap, researchers develop a life cycle optimization framework to integrate multi-objective optimization techniques with LCA and techno-economic analysis (You and Wang 2011). The combination of rigorous mathematical programming and LCA enables systematic generation and evaluation of design candidates. With more than one objective functions allowing for simultaneous optimization of the environmental and economic performances, we are able to obtain more sustainable solutions for a great number of applications (Yue et al. 2013; Zhang et al. 2014).

In life cycle optimization, instead of including the whole package of LCA into optimization models, *goal and scope definition* and *LCI analysis* are performed following the development of a superstructure optimization model, the constraints introduced in *life cycle impact assessment* are added to the original superstructure optimization model, and one or more impact indicators are also included as objective functions in the new model. In most applications, we selectively focus on one environmental impact, such as the greenhouse gas emissions. The motivation for such simplification is twofold: (1) compared to other environmental impacts, climate change receives the most attention worldwide; (2) from the perspective of optimization, including more objective functions leads to computationally demanding optimization problems, especially when the process synthesis problem itself contains complex nonlinear terms.

$$\begin{aligned} \min_{x,y} \quad & TAC \\ \min_{x,y} \quad & GHG \\ \text{s.t.} \quad & h(x,y) = 0 \\ & g(x,y) \leq 0 \\ & x \in \Re, y \in \{0,1\} \end{aligned} \tag{P1}$$

A typical model formulation is shown above. This model has two objectives: (1) minimizing total annualized cost (TAC), and (2) minimizing greenhouse gas emissions (GHG). x denotes continuous variables, such as the mass flow rates and energy consumption rates, while y represents binary variables and determines the superstructure configuration. There are four types of constraints in the model: mass balance constraints, energy balance constraints, economic evaluation constraints, and environmental impact analysis constraints. Mass balance and energy balance are based on mass and energy conservation relationships between input and output materials of each equipment unit. Coefficients are extracted from experiments and simulations results, and nonlinear terms are necessarily employed. For example, we need bilinear terms as the upper bounds of specific mass flowrates (Gong and You 2014b); we employ power functions to estimate equipment capital costs (Gebreslassie et al. 2013b); we formulate fractional objective functions to capture the optimal economic and environmental values per functional unit (Gong and You 2014a).

As a result, the combination of nonlinear terms and the integer variables for superstructure configuration leads to a bi-criteria mixed-integer nonlinear programming (MINLP) problem, which cannot be solved directly by off-the-shelf solvers. There are several methods proposed to handle the multi-objective optimization problem, which are weighted-sum method, ε-constraint method, and evolutionary algorithms (Deb 2014). The ε-constraint method is widely used due to its efficiency and simplicity. The key step of the ε-constraint involves transforming one of the objective functions, such as the environmental objective, into an additional bounding constraint and generates an auxiliary MINLP problem (P2).

$$\begin{aligned} \min_{x,y} \quad & TAC \\ \text{s.t.} \quad & h(x,y) = 0 \\ & g(x,y) \leq 0 \\ & x \in \Re,\ y \in \{0,1\} \\ & GHG \leq \varepsilon \end{aligned} \tag{P2}$$

The upper bound of ε is defined as the evaluated value of the environmental objective when only the economic objective is maximized in problem (P1), while the lower bound of ε is the optimal objective function value when we optimize only the environmental objective in problem (P1). Next, the interval between the upper bound and lower bound is equally partitioned. For each partitioned value of ε, we solve the corresponding problem (P2) and obtain an optimal solution. The set of

optimal solutions can be plotted into a Pareto-optimal curve, which clearly reveals the tradeoff between two objective functions (You and Wang 2011). In addition to the traditional ε-constraint method, an improved method is also examined (Garcia and You 2015).

Attractive alternatives to the objective functions in problem (P1) are unit objective functions associated with the functional units defined in the goal and scope definition. Optimizing unit objective functions considers the effects of both total behavior and the product quantity, which enables direct and constructive comparison with real market values. If the nonlinear terms in the mathematical model appear only when evaluating equipment capital costs with sizing equations, the corresponding bi-criteria mixed-integer fractional programing problem can be effectively optimized with a global optimization strategy (Gong and You 2014a).

3 Process Design and Synthesis of Algal Biorefineries

The recent years have witnessed the thriving of renewable energy as one of the solutions to climate change and energy uncertainties (Lal 2008). Readers are referred to chapter "Biomass as Source for Chemicals, Power and Fuels" that focuses on general biomass-based process design and synthesis. In this chapter, we concentrate on microalgae and algal biorefineries, which have received increasing attention for developing advanced biofuels. Microalgae refer to a diverse category of microorganisms which, similar to most green plants, rely on photosynthesis to thrive. Since carbon dioxide is captured and sequestered into biomass during photosynthesis, green plants are naturally one of the most efficient contributors to carbon mitigation. As an ancient creature on earth, microalgae have a much higher lipid yield compared to traditional energy crops (Ferrell and Sarisky-Reed 2010). Furthermore, microalgae cultivation does not require valuable arable land, thus minimizing the competition with crops and the threat to food supplies. Given that lipids are the precursor to biofuels, algal biofuels will be potentially affordable and sustainable if the downstream processing is efficient and reliable.

Regarding the process design and synthesis of an algal biorefinery, which integrates both algae production and algal biofuel production, there has been an increasing awareness of the importance to develop sustainable algal systems in the recent decade. *National algal biofuel technology roadmap* (Ferrell and Sarisky-Reed 2010), released by the U.S. Department of Energy, is a critical report that summarizes the progress made by the Biomass Program. It states the challenges in commercializing algal biorefineries. It is believed that the realization of commercially viable and environmentally benign algal biofuels requires the cooperation among scientists, engineers, and investors across various disciplines. In a different report from national laboratories (Davis and Aden 2012), specific attention was paid on harmonizing the results for techno-economic analysis (Davis et al. 2011), life cycle analysis (Frank et al. 2011), and resource assessment (Wigmosta et al. 2011) with respect to the production of algal renewable diesel. Other thorough evaluations

examined algal biorefineries from different angles (Alabi et al. 2009; Jones et al. 2014; Davis et al. 2014).

In addition to the official reports, a plethora of academic articles is concerned with the process design and synthesis of algal biorefineries. Reviews with general concepts and promising technologies of algal biorefinery processes were performed by Chisti (2007), Mata et al. (2010), Brennan and Owende (2010), Chen et al. (2011a), Gouveia (2011), and Kim et al. (2013). From a systematic point of view, several attempts have been made to assess the algal biofuel production processes using state-of-the-art simulation methods. Pokoo-Aikins et al. addressed the design and techno-economic analysis of the production of biodiesel from algal oil using ASPEN PLUS simulation and ICARUS cost estimation (Pokoo-Aikins et al. 2010). Lundquist et al. evaluated the economics of microalgae biofuels production through an analysis of five production scenarios and provided recommendations for large-scale algae biofuel production (Lundquist et al. 2010). Delrue et al. developed a model for process evaluation based on energetic, economic, and environmental criteria and conclude that a wet-lipid-extraction route performed better, but this route still has an unfavorable production cost (Delrue et al. 2012). The same authors subsequently evaluated the application of hydrothermal liquefaction (HTL), oil secretion, and alkane secretion in the algal processes (Delrue et al. 2013). Furthermore, Silva et al. simulated an algae-to-biodiesel process with detailed kinetics models, resulting in an attractive price $4.34/gallon (Silva et al. 2014; Dunlop et al. 2013).

Beyond the use of simulation tools to establish feasible algal processes with extensive details, superstructure optimization not only offers sustainable insights into the performance of technology alternatives, but also generates optimal processes with the most favorable biofuel products. Gebreslassie et al. proposed a detailed superstructure of an algal biorefinery focusing on biofuel production (Gebreslassie et al. 2013b). In order to explore the potential of algal biorefinery processes for biological carbon sequestration and utilization, Gong et al. improved the overall performance of the optimal algal biorefinery by incorporating more technology alternatives into the superstructure (Gong and You 2014b). Later, the same authors constructed an algae processing networking considering 7,800 processing routes, highlighting the performance of HTL in the most environmentally sustainable processing route (Gong and You 2014a). In a recent study, they developed the most comprehensive superstructure to demonstrate the importance of coproduction of algal biodiesel and bioproducts (Gong and You 2015). Martín and Grossmann optimized the production of biodiesel from waste cooking oil and algae oil with the consideration of five transesterification pathways (Martín and Grossmann 2012). In a recent article, the authors explored the simultaneous production of algal substitutes, including ethanol, biodiesel, and glycerol ethers (Martín and Grossmann 2014). Rizwan et al. proposed a superstructure to quickly scan through the algal processing pathways and identified the optimal routes with respect to various objective functions (Rizwan et al. 2013; Quaglia et al. 2012). Additionally, much effort was also made on polygeneration systems (Chen et al. 2011b, 2012; Liu et al. 2010), thermochemical conversion system (Baliban et al. 2013; Wang et al. 2013),

and sustainability (Zhang et al. 2014; Gong and You 2014a; Gebreslassie et al. 2013a; Lira-Barragán et al.), offering useful ideas for the integration of energy systems.

4 Superstructure of an Algal Biorefinery

The life cycle design of an algal biorefinery lays its foundation on superstructure optimization. This chapter briefly discusses the advantages of superstructure optimization and provides a thorough overview for the generation of a superstructure for algal biorefinery processes.

4.1 The Role of Superstructure Optimization

Process synthesis determines the type, design, and interconnection of processing units from an enormous number of alternatives after the task of a process is specified with the input and output materials (Nishida et al. 1981). There are a variety of systematic methods to evaluate and search among design alternatives (Biegler et al. 1997). The first method is total enumeration of an explicit space, which is done by simply evaluating every alternative and selecting the one with the best evaluation results. This method seems straightforward in principle but demands large amounts of computation or human resources, unless the number of alternatives is considerably small. The second is referred to as evolutionary methods, which begin with a feasible base-case design and make incremental improvements to part of the design in each iteration. The most significant drawback of this method is that it is unable to locate and reach the globally optimal process design within the available design space. In the third option, we could propose a superstructure containing all available technology alternatives. Based on the superstructure, we then develop a rigorous mathematical programming model which offers an opportunity to simultaneously evaluate the performance of all alternatives and automatically generate the optimal process design. This method is limited by the complexity of the superstructure and the corresponding mathematical formulation. In the early stage of process design, superstructure optimization can be substantially useful to provide initial process configuration from a number of promising alternatives.

The above methods have their pros and cons depending on the specific problems to be solved. Traditional goal for process design and synthesis is to develop sophisticated designs based on the knowledge of existing ones and experimental results. In concert with the developments in computing, process system engineering has made rapid progress over the past 50 years. In order to pursue higher value from a process, the goal, however, has been shifted to finding best designs from numbers of possible options (Westerberg 2004). Assisted directly by the computation

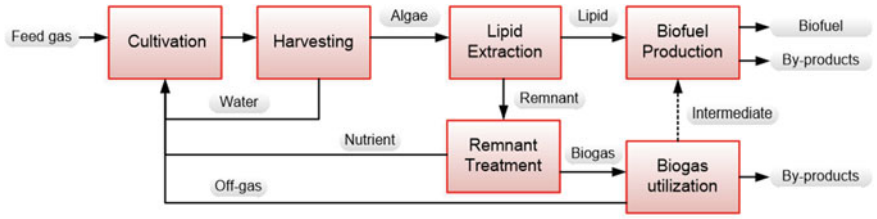


Fig. 1 The typical structure of an algal biorefinery

techniques, superstructure optimization is able to identify the globally optimal process designs with adequate modeling details.

In superstructure-based process synthesis, the first step is to develop a superstructure to include process/technology alternatives across multiple unit operations from the feed to the products. A mathematical model is then developed to optimize the superstructure using computational tools. The major decisions of superstructure optimization involve choosing the optimal process network design, configuration and equipment sizes, as well as the mass and energy balance of each unit operation. The corresponding mathematical programming models are usually formulated as MINLP problems, which can be solved by off-the-shelf global MINLP solvers, such as BARON and SCIP. If, in some circumstances, the resulting MINLP models are equipped with specific nonlinear functions, such as fractional functions, researchers exploit the mathematical properties of these functions and develop tailored algorithms to enhance the computation efficiency (Yue et al. 2013; Gong and You 2014a).

4.2 Technology Alternatives of an Algal Biorefinery

Before any methods are applied to superstructure optimization, the first step in life cycle process design is to generate a superstructure with all possible technology alternatives.

As shown in Fig. 1, an algal biorefinery can be divided into six sections, namely, cultivation, harvesting, lipid extraction, remnant treatment, biofuel production, and biogas utilization. We briefly describe the technology alternatives in the following sections.

5 Cultivation

For autotrophic or phototrophic algae, the feed gas is usually a carbon dioxide-rich source, but the corresponding carbon dioxide concentration should not be too high to inhibit the growth (Hanagata et al. 1992). A widely used feed source is flue gas,

the waste gas from coal-fired power plants. A common concern about using flue gas for cultivation is the toxicity of the minor constitutes, such as SO_x and NO_x, to algae (Wang et al. 2008). However, further discussion about the algal tolerance to these chemicals exceeds the scope of this paper. We assume the upstream power plants have installed effective pollution control units, including acid gas scrubbers which are capable of reducing the concentrations of toxic species to the level that their adverse effects to algae are negligible. We also assume the microalgae strain in our discussion, *Chlorella vulgaris*, contains 25 wt% of the lipid materials in its dry weight and the algal system is immune to the disturbance of weather and other external factors. Although ambitious from a practical point of view, these assumptions are significantly important in simplifying the design problem via excluding the dynamic nature and inherent uncertainties from the system. Therefore, these assumptions are widely used in the articles reviewed in the previous section.

There are two major types of algal bioreactors: open systems, or open ponds, and closed systems, or photobioreactors (PBRs) (Razzak et al. 2013). Open ponds include unstirred ponds, paddle-wheel raceway ponds, and circular ponds. PBRs are more diverse, including tubular PBRs, flat plate PBRs, bubble column PBRs, and plastic bag PBRs. Open ponds are economically favorable to set up and operate, but the productivity is poor. Contrarily, PBRs are reliable and efficient, but they suffer from much higher capital costs. Further idea of a hybrid system has also been mentioned in order to take into account the inoculation into the system (Gouveia 2011). However, in most process synthesis problems, inoculation is omitted due to its small contribution to the total cost.

6 Harvesting

The direct product from the cultivation reactors contains only 0.014 wt% of mature biomass for open ponds and 0.4 wt% of mature biomass for PBRs (Chisti 2007). However, the downstream lipid extraction and HTL equipment require the algae concentration higher than 30 and 5 wt%, respectively (Valdez et al. 2012). Therefore, we need a harvesting section to remove the undesired water content. A reasonable harvesting design involves three steps (Davis et al. 2011). The first step takes place in a sedimentation tank/basin, where a large proportion of the surplus water is separated via autoflocculation assisted with dissolved air flotation. This step achieves an algae slurry of 10 wt% (Uduman et al. 2010). There are a variety of flocculants to be used for gravitational separation, including poly-electrolyte, sodium hydroxide, poly-aluminum chloride, aluminum sulfate, chitosan, and poly-γ-glutamic acid (Rizwan et al. 2013). They differ from each other by the concentrations for effective separation, and the percentages of algal biomass to the next harvesting step. The impact of flocculants on the downstream processing is assumed negligible. The second step, or the dewatering step, involves the use of pressure filtration or centrifugation to thicken the algae slurry to 30 wt%

(Wang et al. 2007; Uduman et al. 2010). Pressure filtration consumes less energy, but a centrifuge has a lower capital cost. In the last step, or the drying step, the algae paste is dehydrated to over 85 wt% using freeze-drying or thermal-drying equipment. This step is only necessary if the lipid extraction requires algae powder as feed materials. In order to reduce both the operating cost and the environmental impact, the surplus water from the first two harvesting steps is recycled and reused in cultivation. 5 wt% of the separated water is purged to prevent the buildup of salts.

Continuous operation is a crucial feature to produce bulk commodities in today's chemical plants. As a result, shutdowns and startups appear only during regular maintenance, unexpected failures, and facilities upgrading, since frequent discontinuities not only jeopardize the mechanical stability of devices, but also incur large quantity of non-profitable off-spec products during transition. Constrained by the availability of sunlight, the equipments in the cultivation and harvesting sections have to stop operating at night. Therefore, a biomass storage tank would be substantially important at the end of the harvesting section. The storage tank can be designed to reserve part of the biomass product during daytime operation and release it to the downstream at night, so that the downstream processes receive biomass feed continuously throughout the day. Another approach is to provide artificial light to the cultivation system during nighttime production. However, this idea proves both economically and environmentally unfavorable (Gong and You 2014b).

7 Lipid Extraction

The goal of this section is to separate the upgradable lipid materials from algae cells. There are two major methods. In the first group, we take advantage of the solvent extraction methods that have already been implemented in large-scale biodiesel production from older generations of biomass. Nevertheless, the existence of algal biological features like the cell wall and the cell membrane physically prevents the contact of solvent with intracellular lipid materials. One solution is to employ cell disruption technologies. Possible options for cell disruption include sonication, high pressure homogenization, bead beating, freeze-drying, microwaving, autoclaving, and hydrodynamic cavitation (Lee et al. 2012). Although different in terms of the cell disruption efficiency, these methods share a common drawback, which is high energy consumption. Hence, it is worthwhile to consider the option of bypassing the cell disruption in a superstructure, which lowers the lipid content available in solvent extraction, but avoids the exceedingly high energy consumption.

A rough classification of extraction technologies with respect to algal biomass shows a dry route and a wet route (Xu et al. 2011). The dry route aims at separating lipid materials from algal powders, while the wet route does not require complete drying in the harvesting section, but more energy is consumed in solvent recovery

(Gong and You 2015). Quite a few publications attempt to find the most appropriate extraction method at the laboratory scale (Halim et al. 2012; Bollmeier et al. 1989). Multiple solvents and solvent mixtures have been examined, including water, methanol, chloroform, butanol, hexane, ethanol, isopropanol, supercritical carbon dioxide, etc. but these experimental results are less attractive in the context of a complete process. Therefore, when we set up a superstructure involving solvent extraction, it is crucial to consider the entire extraction–recovery cycle no matter what routes and what solvents are included as alternatives (Gong and You 2015).

Another promising alternative to solvent extraction is HTL, which is able to utilize algal slurry directly and convert it to biogas, lipids, and remnant. In addition to the flexible requirement with respect to feed concentration, another advantage of this technology is its high lipid yield. This is achieved by converting a portion of the algal protein and carbohydrates into lipids (Valdez et al. 2012). However, the operating condition of technology is close to the supercritical condition of water, leading to both high capital cost and high energy consumption.

8 Remnant Treatment

After lipid extraction, a significant amount of energy captured in algal photosynthesis still remains in the remnant. If this amount of biomass is discarded or sold with low prices, the algal biorefinery will lose the potential for harvesting the most value from algae. Two technologies are mentioned in the literature for utilizing algal remnant: anaerobic digestion and catalytic gasification.

Anaerobic digestion is the most widely used technology in the remnant treatment, resulting in a biogas stream with about 40 vol% of carbon dioxide and 60 vol% of methane, and an aqueous stream rich in soluble nutrients for algae cultivation and solid precipitates (Collet et al. 2011). Methane can be carefully purified from the biogas stream through a pressure-swing-adsorption system. Simultaneously, solid precipitates are separated by centrifugation and the liquid product is reused in cultivation. Alternatively, according to Elliott et al. (2012), catalytic gasification generates four products, including a biogas product consisting of 0.5 % ethane, 2.1 % hydrogen, 38 % carbon dioxide, and 57.1 % methane, a solid product rich in mineral deposits, an aqueous product with nitrogen and phosphorus nutrition, and an oil product which can enhance the biofuel yield.

9 Biogas Utilization

No matter what pathway is selected in remnant treatment, a biogas product is guaranteed. After further separation using a pressure-swing-adsorption system, the versatile methane can be concentrated and utilized onsite. Three technologies are found worth considering in an algal biorefinery: direct combustion, hydrogen

production, and methanol production (Gong and You 2014b, 2015). Direct combustion process comprised an air compressor, a combustor, a gas turbine, and a heat-recovery-steam-generation system. The chemical energy embedded in methane is partially transformed to heat, and then to electricity through the energy conversion in combustion, shaft work, and heat exchanging. In addition to the energy utilization, the resulting off-gas is rich in carbon dioxide and could be treated as the feed gas in cultivation.

Hydrogen and methanol are possible raw materials in the biofuel production section. Integrating onsite generation of hydrogen and methanol into the algal biofuel processes holds the potential for cost mitigation. Hydrogen could be produced by introducing methane into steam reforming, followed by water gas shift. In contrast, if the syngas from the stream reforming reactor is sent to a demister, followed by a methanol reactor, a second demister, and a distillation column, we would obtain methanol instead (Cheng 1994). These intermediate products can also be sold to the market if the production exceeds the demand in biofuel production.

10 Biofuel Production

Two major biofuel products can be derived from algal lipids: biodiesel and renewable diesel (Albrecht and Hallen 2011). The former is formally referred to as fatty acid methyl esters, or FAMEs, which are readily produced from transesterification reactions. On the other hand, triacylglycerides, free fatty acids, and FAMEs can undergo deoxygenation reactions and become saturated hydrocarbons, usually known as renewable diesel or green diesel. Renewable diesel has essentially low oxygen content, thus compatible for blending with petroleum-based diesel and used in the existing infrastructures.

From lab-scale experiments to process simulations, extensive research has been done for transesterification pathways (Apostolakou et al. 2009; Xu and Mi 2011; West et al. 2008; Sotoft et al. 2010; Meher et al. 2006). Available reaction conditions include alkaline in situ transesterification, acidic in situ transesterification, enzymatic in situ transesterification, sodium-methoxide-catalyzed transesterification, supercritical methanol transesterification, enzyme-catalyzed transesterification, and heterogeneous acid-catalyzed transesterification (Rizwan et al. 2013; Gebreslassie et al. 2013b). Comparatively, much more research is needed to achieve reliable and cost-effective conditions for algal renewable diesel. Existing superstructures are based on the deoxygenation reaction conditions for vegetable oil (Marker 2005). In this report, Co-Mo catalyst for a decarboxylation mechanism and Ni–Mo catalyst for a hydrodeoxygenation mechanism can be introduced as two alternatives in a superstructure. HZSM-5 catalyst is a tailored option for hydroprocessing the oil product from HTL (Duan and Savage 2011).

Further attention is also paid to the utilization of by-products. For example, transesterification reactions generate roughly one ton of glycerol per every ten tons of biodiesel. When the biodiesel industry continually expands, the price of surplus

glycerol will be pushed to a very low level and the conversion of glycerol into value-added derivatives will undoubtedly benefit the economics of the algal biorefinery. Potential chemicals from glycerol include hydrogen, propylene glycol, glycerol-*tert*-butyl ether, and poly-3-hydroxybutyrate (Gong and You 2015).

11 Life Cycle Design of an Algal Biorefinery

Following the modeling framework introduced in the second section, we are able to develop multi-objective MINLP models to simultaneously optimize the economic and environmental indicators of possible processes in the aforementioned superstructure. Optimal solutions of the multi-objective MINLP problems can be plotted in a Pareto-optimal curve (Gong and You 2015) as in Fig. 2. The points on the Pareto-optimal curve represent optimal solutions with trade-offs in the objective functions, the points to the left side of the curve are infeasible, and the points to the right side of the curve are suboptimal. Point 1 and Point 8 are two extreme points on the curve, representing the most environmentally optimal process design and the most economically optimal process design, respectively. Environmentally friendly process designs favor technologies with less GHG emissions. For example, the optimal process design represented by Point 1 employs pressure filtration, hexane extraction, and enzyme-catalyzed transesterification, which involve less GHG emissions compared to their counterpart technologies.

A good-choice optimal solution represented by Point 4 on the Pareto-optimal curve is shown in Fig. 3. This optimal process employs pressure filtration in the Harvesting section, supercritical carbon dioxide as the lipid extractant, direct combustion for biogas utilization, enzyme-catalyzed transesterification for biofuel

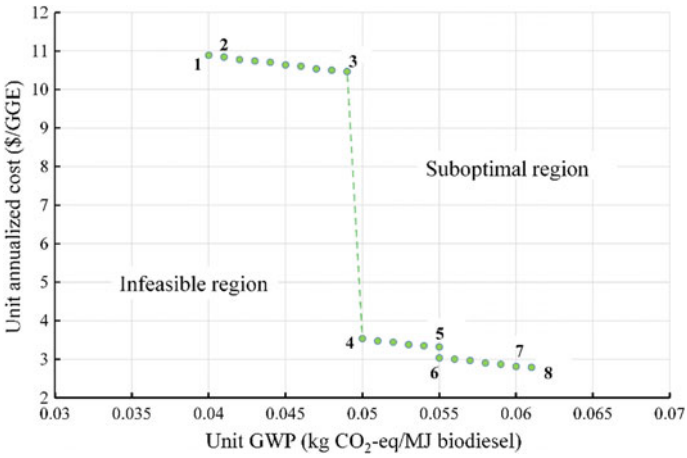


Fig. 2 Pareto-optimal curve (*Source* Gong and You 2015)

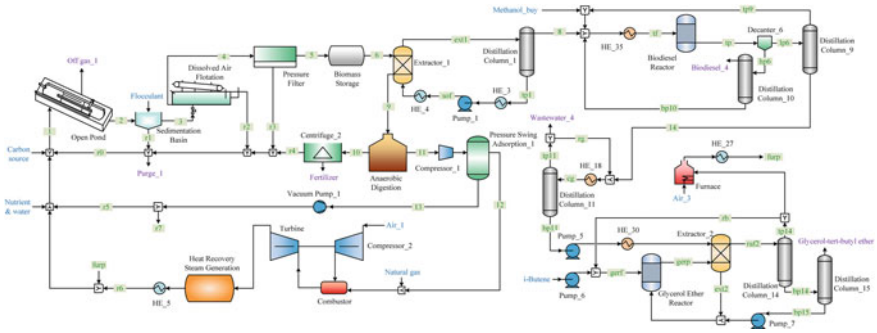


Fig. 3 The optimal process of Point 4 on the Pareto-optimal curve (*Source* Gong and You 2015)

production, and GE synthesis for bioproduct manufacturing. The Pareto-optimal curve provides a flexible way of selecting optimal algal biorefinery designs that complies with environmental regulations, and investigates the optimal economic and environmental performance of an algal biorefinery.

12 Future Directions

12.1 Modeling Details

When we develop the superstructure of an algal biorefinery, several aggressive assumptions regarding algae cultivation are made in order to simplify the mathematical models. However, if more accurate results are desired, these assumptions should be revised to reflect more details in real situations. For example, it is a difficult task to conduct stringent simulation for algal biorefineries as opposed to petroleum refineries, because a large number of species are involved in the bioprocess and the physical properties of specific macromolecules are scarce. An improved model considering more species and their reactions necessitates abundant experimental data in every step of the algal biorefinery. Additionally, the lipid ratio and the productivity of the algae strain is assumed immune to uncertain accidents and external contamination. In order to address this assumption, we can impose productivity uncertainties on the algae cultivation in the design problem.

In the current stage, operating conditions of most technology alternatives tend to be fixed. Ideally, a process synthesis problem should be able to determine the optimal solutions of a wide array of design variables, including separation sequence, equipment sizes, temperatures, and pressures, as well as utility consumption. However, in the case of an algal biorefinery, few fundamental models are reported for correlations between the lipid reaction rate and temperature and pressure. Even though they are available, modeling a biochemical process with all the design variables leads to large numbers of nonlinear relationships. These

nonlinear functions unfortunately contribute to a computationally demanding mathematical problem. Therefore, the challenge of superstructure optimization for algal biorefineries is to develop solvable high-fidelity models for current and future technology alternatives, and integrate them with the superstructure configuration models.

12.2 Life Cycle Considerations

As mentioned earlier, the current life cycle designs of algal biorefineries focus primarily on the economic and environmental performance. However, social considerations have been overlooked from the definition of sustainability (You et al. 2012; Yue et al. 2014a). Also, more environmental impact categories, such as acidification and ecotoxicity, are worth investigating when we intend to implement an algal biorefinery in a specific area.

In terms of climate change evaluation, it has been reported that the widely used metric GWP(100), or global warming potential based on 100-year time horizon, generates significant overshoots of radiative forcing stabilization targets when it is used to quantify the CO_2-equivalent mass of GHGs (Edwards and Trancik 2014). Therefore, improving the accuracy of the environmental evaluation results of a life cycle process design problem can be achieved via switching to better metrics.

13 Conclusions

This chapter discusses the framework of life cycle process design and synthesis, and its application to algal biorefineries. We summarize a large number of technology alternatives that can be incorporated into an algal superstructure and present methods for solving the corresponding multi-objective MINLP problems. Several potential improvements are identified, including developing reliable cultivation models based on experimental results, imposing productivity uncertainties to better describe the influence of external conditions, and integrating social considerations, more environmental impact categories, and better equivalency metric for climate change evaluation into the life cycle optimization framework.

Acknowledgement We gratefully acknowledge the financial support from the Institute for Sustainability and Energy at Northwestern University (ISEN).

References

Alabi AO, Bibeau E, Tampier M, Council BCI (2009) Microalgae Technologies & Processes for Biofuels-bioenergy Production in British Columbia: Current Technology, Suitability & Barriers to Implementation: Final Report. British Columbia Innovation Council,

Albrecht KO, Hallen RT (2011) A Brief Literature Overview of Various Routes to Biorenewable Fuels from Lipids for the National Alliance for Advanced Biofuels and Bio-products (NAABB) Consortium.

Apostolakou AA, Kookos IK, Marazioti C, Angelopoulos KC (2009) Techno-economic analysis of a biodiesel production process from vegetable oils. Fuel Process Technol 90 (7–8):1023-1031. doi:DOI 10.1016/j.fuproc.2009.04.017

Bakshi BR (2014) Methods and tools for sustainable process design. Current Opinion in Chemical Engineering 6 (0):69–74. doi:http://dx.doi.org/10.1016/j.coche.2014.09.005

Baliban RC, Elia JA, Floudas CA, Xiao X, Zhang ZJ, Li J, Cao HB, Ma J, Qiao Y, Hu XT (2013) Thermochemical Conversion of Duckweed Biomass to Gasoline, Diesel, and Jet Fuel: Process Synthesis and Global Optimization. Industrial & Engineering Chemistry Research 52 (33):11436–11450. doi:Doi 10.1021/Ie3034703

Biegler LT, Grossmann IE, Westerberg AW, Kravanja Z (1997) Systematic methods of chemical process design, vol 796. Prentice Hall PTR Upper Saddle River, NJ,

Bollmeier W, Sprague S, No F (1989) Aquatic Species Program. Solar Energy Research Inst., Golden, CO (USA),

Brennan L, Owende P (2010) Biofuels from microalgae-A review of technologies for production, processing, and extractions of biofuels and co-products. Renew Sust Energ Rev 14 (2):557–577. doi:DOI 10.1016/j.rser.2009.10.009

Chen CY, Yeh KL, Aisyah R, Lee DJ, Chang JS (2011a) Cultivation, photobioreactor design and harvesting of microalgae for biodiesel production: A critical review. Bioresource Technol 102 (1):71–81. doi:DOI 10.1016/j.biortech.2010.06.159

Chen Y, Adams TA, Barton PI (2011b) Optimal Design and Operation of Flexible Energy Polygeneration Systems. Industrial & Engineering Chemistry Research 50 (8):4553–4566. doi:Doi 10.1021/Ie1021267

Chen Y, Li X, Adams TA, Barton PI (2012) Decomposition strategy for the global optimization of flexible energy polygeneration systems. AIChE 58 (10):3080–3095. doi:Doi 10.1002/Aic.13708

Cheng W-H (1994) Methanol production and use. CRC Press,

Chisti Y (2007) Biodiesel from microalgae. Biotechnol Adv 25 (3):294–306. doi:DOI 10.1016/j.biotechadv.2007.02.001

Collet P, Helias A, Lardon L, Ras M, Goy RA, Steyer JP (2011) Life-cycle assessment of microalgae culture coupled to biogas production. Bioresource Technol 102 (1):207–214. doi:DOI 10.1016/j.biortech.2010.06.154

Davis R, Aden A (2012) Renewable Diesel from Algal Lipids: An Integrated Baseline for Cost, Emissions, and Resource Potential from a Harmonized Model. Argonne National Laboratory,

Davis R, Aden A, Pienkos PT (2011) Techno-economic analysis of autotrophic microalgae for fuel production. Appl Energ 88 (10):3524–3531. doi:DOI 10.1016/j.apenergy.2011.04.018

Davis R, Kinchin C, Markham J, Tan ECD, Laurens LML, Sexton D, Knorr D, Schoen P, Lukas J (2014) Process Design and Economics for the Conversion of Algal Biomass to Biofuels: Algal Biomass Fractionation to Lipid-and Carbohydrate-Derived Fuel Products.

Deb K (2014) Multi-objective optimization. In: Search methodologies. Springer, pp 403–449

Delrue F, Seiter PA, Sahut C, Cournac L, Roubaud A, Peltier G, Froment AK (2012) An economic, sustainability, and energetic model of biodiesel production from microalgae. Bioresource Technol 111:191–200. doi:DOI 10.1016/j.biortech.2012.02.020

Delrue F, Li-Beisson Y, Setier PA, Sahut C, Roubaud A, Froment AK, Peltier G (2013) Comparison of various microalgae liquid biofuel production pathways based on energetic,

economic and environmental criteria. Bioresource Technol 136:205–212. doi:DOI 10.1016/j.biortech.2013.02.091

Duan PG, Savage PE (2011) Catalytic hydrotreatment of crude algal bio-oil in supercritical water. Appl Catal B-Environ 104 (1–2):136-143. doi:DOI 10.1016/j.apcatb.2011.02.020

Dunlop EH, Coaldrake AK, Silva CS, Seider WD (2013) An Energy-limited Model of Algal Biofuel Production: Toward the Next Generation of Advanced Biofuels. AIChE 59 (12):4641–4654. doi:Doi 10.1002/Aic.14251

Edwards MR, Trancik JE (2014) Climate impacts of energy technologies depend on emissions timing. Nat Clim Change 4 (5):347–352

Elliott DC, Hart TR, Neuenschwander GG, Rotness LJ, Qlarte MV, Zacher AH (2012) Chemical Processing in High-Pressure Aqueous Environments. 9. Process Development for Catalytic Gasification of Algae Feedstocks. Industrial & Engineering Chemistry Research 51 (33):10768–10777. doi:Doi 10.1021/Ie300933w

Ferrell J, Sarisky-Reed V (2010) National algal biofuels technology roadmap (trans: Energy EER). US Department of Energy College Park, MD

Finnveden G, Hauschild MZ, Ekvall T, Guinee J, Heijungs R, Hellweg S, Koehler A, Pennington D, Suh S (2009) Recent developments in Life Cycle Assessment. J Environ Manage 91 (1):1–21

Frank E, Han J, Palou-Rivera I, Elgowainy A, Wang M (2011) Life-cycle analysis of algal lipid fuels with the greet model. Center for Transportation Research, Energy Systems Division, Argonne National Laboratory, Oak Ridge

Garcia DJ, You F Supply chain design and optimization: Challenges and opportunities. Computers & Chemical Engineering: DOI: http://dx.doi.org/10.1016/j.compchemeng.2015.1003.1015. doi:http://dx.doi.org/10.1016/j.compchemeng.2015.03.015

Garcia DJ, You F (2015) Multiobjective optimization of product and process networks: General modeling framework, efficient global optimization algorithm, and case studies on bioconversion. AIChE 61 (2):530–554. doi:10.1002/aic.14666

Gebreslassie BH, Slivinsky M, Wang BL, You FQ (2013a) Life cycle optimization for sustainable design and operations of hydrocarbon biorefinery via fast pyrolysis, hydrotreating and hydrocracking. Computers & Chemical Engineering 50:71–91. doi:DOI 10.1016/j.compchemeng.2012.10.013

Gebreslassie BH, Waymire R, You F (2013b) Sustainable design and synthesis of algae-based biorefinery for simultaneous hydrocarbon biofuel production and carbon sequestration. AIChE 59 (5):1599–1621. doi:10.1002/aic.14075

Gong J, You F (2015) Value-Added Chemicals from Microalgae: Greener, More Economical, or Both? Acs Sustain Chem Eng 3 (1):82–96. doi:10.1021/sc500683w

Gong J, You FQ (2014a) Global Optimization for Sustainable Design and Synthesis of Algae Processing Network for CO2 Mitigation and Biofuel Production Using Life Cycle Optimization. AIChE 60 (9):3195–3210. doi:Doi 10.1002/Aic.14504

Gong J, You FQ (2014b) Optimal Design and Synthesis of Algal Biorefinery Processes for Biological Carbon Sequestration and Utilization with Zero Direct Greenhouse Gas Emissions: MINLP Model and Global Optimization Algorithm. Industrial & Engineering Chemistry Research 53 (4):1563–1579. doi:Doi 10.1021/Ie403459m

Gouveia L (2011) Microalgae as a Feedstock for Biofuels. Springerbrief Micro:1–69. doi:Doi 10.1007/978-3-642-17997-6_1

Grießhammer R, Benoît C, Dreyer LC, Flysjö A, Manhart A, Mazijn B, Méthot A-L, Weidema B (2006) Feasibility study: integration of social aspects into LCA.

Halim R, Danquah MK, Webley PA (2012) Extraction of oil from microalgae for biodiesel production: A review. Biotechnol Adv 30 (3):709–732. doi:DOI 10.1016/j.biotechadv.2012.01.001

Hanagata N, Takeuchi T, Fukuju Y, Barnes DJ, Karube I (1992) Tolerance of Microalgae to High Co-2 and High-Temperature. Phytochemistry 31 (10):3345–3348. doi:Doi 10.1016/0031-9422(92)83682-O

Institution BS (2006) ISO 14040:2006 Environmental Management–Life Cycle Assessment–Principles and Framework. London

Jones SB, Zhu Y, Anderson DB, Hallen RT, Elliott DC, Schmidt AJ, Albrecht KO, Hart TR, Butcher MG, Drennan C, Snowden-Swan LJ, Davis R, Kinchin C (2014) Process Design and Economics for the Conversion of Algal Biomass to Hydrocarbons: Whole Algae Hydrothermal Liquefaction and Upgrading.

Kim J, Yoo G, Lee H, Lim J, Kim K, Kim CW, Park MS, Yang JW (2013) Methods of downstream processing for the production of biodiesel from microalgae. Biotechnol Adv 31 (6):862–876. doi:DOI 10.1016/j.biotechadv.2013.04.006

Lal R (2008) Carbon sequestration. Philos T R Soc B 363 (1492):815–830. doi:DOI 10.1098/rstb.2007.2185

Lee AK, Lewis DM, Ashman PJ (2012) Disruption of microalgal cells for the extraction of lipids for biofuels: Processes and specific energy requirements. Biomass Bioenerg 46:89–101. doi:DOI 10.1016/j.biombioe.2012.06.034

Lira-Barragán LF, Ponce-Ortega JM, Serna-González M, El-Halwagi MM Optimal design of process energy systems integrating sustainable considerations. Energy. doi:http://dx.doi.org/10.1016/j.energy.2014.04.111

Liu P, Pistikopoulos EN, Li Z (2010) A Multi-Objective Optimization Approach to Polygeneration Energy Systems Design. AIChE 56 (5):1218–1234. doi:Doi 10.1002/Aic.12058

Lundquist TJ, Woertz IC, Quinn N, Benemann JR (2010) A realistic technology and engineering assessment of algae biofuel production. Energy Biosciences Institute:1

Marker T (2005) Opportunities for biorenewables in oil refineries. UOP LLC,

Martín M, Grossmann IE (2012) Simultaneous Optimization and Heat Integration for Biodiesel Production from Cooking Oil and Algae. Industrial & Engineering Chemistry Research 51 (23):7998–8014. doi:Doi 10.1021/Ie2024596

Martín M, Grossmann IE (2014) Simultaneous Optimization and Heat Integration for the Coproduction of Diesel Substitutes: Biodiesel (FAME and FAEE) and Glycerol Ethers from Algae Oil. Industrial & Engineering Chemistry Research 53 (28):11371–11383. doi:Doi 10.1021/Ie402443t

Mata TM, Martins AA, Caetano NS (2010) Microalgae for biodiesel production and other applications: A review. Renew Sust Energ Rev 14 (1):217–232. doi:DOI 10.1016/j.rser.2009.07.020

Meher LC, Sagar DV, Naik SN (2006) Technical aspects of biodiesel production by transesterification - a review. Renew Sust Energ Rev 10 (3):248–268. doi:DOI 10.1016/j.rser.2004.09.002

Nishida N, Stephanopoulos G, Westerberg AW (1981) A Review of Process Synthesis. AIChE 27 (3):321–351

Pokoo-Aikins G, Nadim A, El-Halwagi MM, Mahalec V (2010) Design and analysis of biodiesel production from algae grown through carbon sequestration. Clean Technol Envir 12 (3):239–254. doi:DOI 10.1007/s10098-009-0215-6

Quaglia A, Sarup B, Sin G, Gani R (2012) Integrated business and engineering framework for synthesis and design of enterprise-wide processing networks. Computers & Chemical Engineering 38:213–223. doi:DOI 10.1016/j.compchemeng.2011.12.011

Razzak SA, Hossain MM, Lucky RA, Bassi AS, de Lasa H (2013) Integrated CO2 capture, wastewater treatment and biofuel production by microalgae culturing-A review. Renew Sust Energ Rev 27:622–653. doi:DOI 10.1016/j.rser.2013.05.063

Rebitzer G, Ekvall T, Frischknecht R, Hunkeler D, Norris G, Rydberg T, Schmidt WP, Suh S, Weidema BP, Pennington DW (2004) Life cycle assessment Part 1: Framework, goal and scope definition, inventory analysis, and applications. Environ Int 30 (5):701–720

Rizwan M, Lee JH, Gani R (2013) Optimal processing pathway for the production of biodiesel from microalgal biomass: A superstructure based approach. Computers & Chemical Engineering 58:305–314. doi:DOI 10.1016/j.compchemeng.2013.08.002

Silva C, Soliman E, Cameron G, Fabiano LA, Seider WD, Dunlop EH, Coaldrake AK (2014) Commercial-Scale Biodiesel Production from Algae. Industrial & Engineering Chemistry Research 53 (13):5311–5324. doi:10.1021/ie403273b

Sotoft LF, Rong BG, Christensen KV, Norddahl B (2010) Process simulation and economical evaluation of enzymatic biodiesel production plant. Bioresource Technol 101 (14):5266–5274. doi:DOI 10.1016/j.biortech.2010.01.130

Uduman N, Qi Y, Danquah MK, Forde GM, Hoadley A (2010) Dewatering of microalgal cultures: A major bottleneck to algae-based fuels. J Renew Sustain Ener 2 (1). doi:DOI 10.1063/1.3294480

Valdez PJ, Nelson MC, Wang HY, Lin XXNN, Savage PE (2012) Hydrothermal liquefaction of Nannochloropsis sp.: Systematic study of process variables and analysis of the product fractions. Biomass Bioenerg 46:317–331. doi:DOI 10.1016/j.biombioe.2012.08.009

Wang B, Li YQ, Wu N, Lan CQ (2008) CO(2) bio-mitigation using microalgae. Appl Microbiol Biot 79 (5):707–718. doi:DOI 10.1007/s00253-008-1518-y

Wang B, Gebreslassie BH, You FQ (2013) Sustainable design and synthesis of hydrocarbon biorefinery via gasification pathway: Integrated life cycle assessment and technoeconomic analysis with multiobjective superstructure optimization. Computers & Chemical Engineering 52:55–76. doi:DOI 10.1016/j.compchemeng.2012.12.008

Wang LK, Shammas NK, Hung YT (2007) Biosolids Treatment Processes: Volume 6. Humana Press

West AH, Posarac D, Ellis N (2008) Assessment of four biodiesel production processes using HYSYS. Plant. Bioresource Technol 99 (14):6587–6601. doi:DOI 10.1016/j.biortech.2007.11.046

Westerberg AW (2004) A retrospective on design and process synthesis. Computers & Chemical Engineering 28 (4):447–458

Wigmosta MS, Coleman AM, Skaggs RJ, Huesemann MH, Lane LJ (2011) National microalgae biofuel production potential and resource demand. Water Resour Res 47. doi:Artn W00h04 Doi 10.1029/2010wr009966

Xu LX, Brilman DWF, Withag JAM, Brem G, Kersten S (2011) Assessment of a dry and a wet route for the production of biofuels from microalgae: Energy balance analysis. Bioresource Technol 102 (8):5113–5122. doi:DOI 10.1016/j.biortech.2011.01.066

Xu RY, Mi YL (2011) Simplifying the Process of Microalgal Biodiesel Production Through In Situ Transesterification Technology. J Am Oil Chem Soc 88 (1):91–99. doi:DOI 10.1007/s11746-010-1653-3

You FQ, Wang B (2011) Life Cycle Optimization of Biomass-to-Liquid Supply Chains with Distributed-Centralized Processing Networks. Industrial & Engineering Chemistry Research 50 (17):10102–10127. doi:Doi 10.1021/Ie200850t

You FQ, Tao L, Graziano DJ, Snyder SW (2012) Optimal design of sustainable cellulosic biofuel supply chains: Multiobjective optimization coupled with life cycle assessment and input-output analysis. AIChE 58 (4):1157–1180. doi:Doi 10.1002/Aic.12637

Yue DJ, Kim MA, You FQ (2013) Design of Sustainable Product Systems and Supply Chains with Life Cycle Optimization Based on Functional Unit: General Modeling Framework, Mixed-Integer Nonlinear Programming Algorithms and Case Study on Hydrocarbon Biofuels. Acs Sustain Chem Eng 1 (8):1003–1014. doi:Doi 10.1021/Sc400080x

Yue D, Slivinsky M, Sumpter J, You F (2014a) Sustainable Design and Operation of Cellulosic Bioelectricity Supply Chain Networks with Life Cycle Economic, Environmental, and Social Optimization. Industrial & Engineering Chemistry Research 53 (10):4008–4029. doi:10.1021/ie403882v

Yue DJ, You FQ, Snyder SW (2014b) Biomass-to-bioenergy and biofuel supply chain optimization: Overview, key issues and challenges. Computers & Chemical Engineering 66:36–56. doi:DOI 10.1016/j.compchemeng.2013.11.016

Zhang Q, Gong J, Skwarczek M, Yue DJ, You FQ (2014) Sustainable Process Design and Synthesis of Hydrocarbon Biorefinery through Fast Pyrolysis and Hydroprocessing. AIChE 60 (3):980–994. doi:10.1002/aic.14344

Planning and Scheduling for Industrial Demand Side Management: Advances and Challenges

Qi Zhang and Ignacio E. Grossmann

Abstract In the context of the so-called smart grid, the intelligent management of electricity demand, also referred to as demand side management (DSM), has been recognized as an effective approach to increase power grid performance and consumer benefits. Being large electricity consumers, the power-intensive process industries play a key role in DSM. In particular, planning and scheduling for industrial DSM has emerged as a major area of interest for both researchers and practitioners. In this work, we provide an introduction to DSM and present a comprehensive review of existing works on planning and scheduling for industrial DSM. Four main challenges are identified: (1) accurate modeling of operational flexibility, (2) integration of production and energy management, (3) optimization across multiple time scales, and (4) decision-making under uncertainty. Two real-world case studies are presented to demonstrate the capabilities of state-of-the-art models and solution approaches. Finally, we highlight the research gaps and future opportunities in this area.

1 Introduction

Smart Grid and Demand Side Management The power grid is designed to reliably match electricity supply and demand. This task has become increasingly challenging due to high fluctuations in electricity demand and increasing penetration of intermittent renewable energy into the electricity supply mix. Also, the deregulation of electricity markets and the pressure to reduce greenhouse gas emissions have further amplified the need to improve the efficiency, reliability, and sustainability of the power grid.

Q. Zhang · I.E. Grossmann (✉)
Department of Chemical Engineering, Carnegie Mellon University,
Pittsburgh, PA 15213, USA
e-mail: grossmann@cmu.edu

Q. Zhang
e-mail: qz@andrew.cmu.edu

M. Martín (ed.), *Alternative Energy Sources and Technologies*,
DOI 10.1007/978-3-319-28752-2_14

In recent years, the notion of a *smart grid* has been evolving, which represents the concept of a power grid in which the major operations—electricity generation, transmission, distribution, and consumption—are executed in a coordinated and efficient manner. To establish such a smart grid, three essential components are required: (1) An information and communications infrastructure has to be in place such that data on grid conditions can be collected and exchanged in real-time. (2) Advanced decision-making tools have to be developed that can use the collected information to optimize operations in the grid. (3) With the help of the decision-making tools, the grid operator may know which actions should be taken to achieve optimal performance; however, these actions have to be actually implemented by the different participants in the grid, e.g., generators and consumers. To encourage and to a certain extent control these actions, efficient markets are required, which provide appropriate financial incentives to the participants.

The idea of a smart grid has gained considerable interest in industry, research, and public policy. Since 2009, the U.S. Department of Energy and the electricity industry have jointly invested over $7.9 billion in smart grid projects as part of the Smart Grid Investment Grant Program (DOE 2013), and similar efforts have been undertaken in many other countries.

Traditionally, in power systems engineering, the focus has been on improving the power supply for given electricity demands (loads) in the grid. A major innovation in smart grid is to also include the management of flexible loads, which is generally referred to as demand side management (DSM) since it involves the enhancement of energy systems on the electricity demand side. DSM is expected to play a crucial role in the improvement of grid efficiency and reliability as well as the creation of additional benefits for the consumers (Levy 2006; Strbac 2008; Siano 2014); this has spurred tremendous research efforts across multiple disciplines, such as electrical engineering, civil and environmental engineering, economics, data science, behavior science, and engineering public policy.

Two Perspectives on DSM The opportunities in DSM can be viewed from two distinct perspectives: the grid operator's perspective and the electricity consumer's perspective. On the one hand, the grid operator's main objective is to increase efficiency and ensure stability in the power grid. In this context, DSM is regarded as a means to reduce the overall electricity demand, to flatten the load curve and hence reduce the required peak generation capacity, as well as to provide the flexibility to quickly react to supply–demand mismatch in the grid by adjusting loads. On the other hand, the electricity consumer's objective is simply cost reduction. For electricity consumers, DSM is required in order to adapt to price signals coming from the electricity market, and it is a way to take advantage of new financial incentives specifically created for DSM purposes. Therefore, ideally, DSM leads to a win–win situation in which both the power grid and the consumers benefit.

Figure 1 illustrates the activities on the grid operator and electricity consumer sides, which are connected by operations in the physical power grid and the opportunities in the electricity markets. Until recently, DSM has been analyzed primarily from the grid operator's perspective. In the U.S., for example, the power grid is operated by the various independent system operators (ISOs) and regional

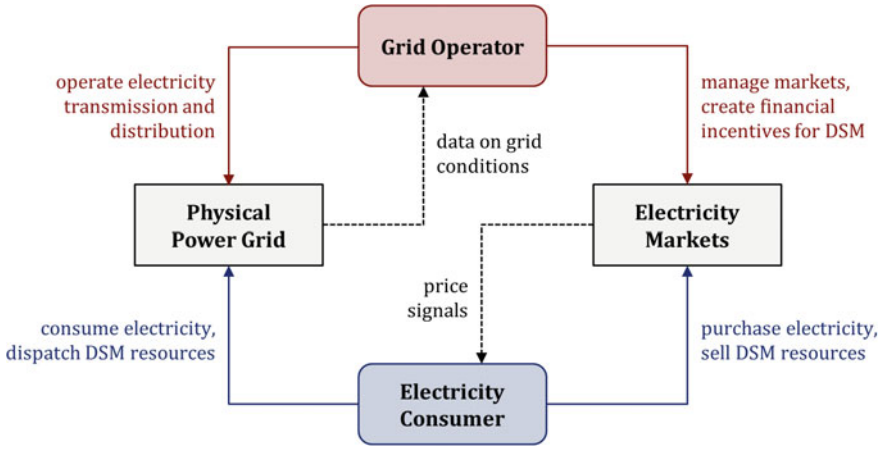


Fig. 1 Two perspectives on DSM, which requires both the grid operator and the consumer to interact through physical grid operations and the electricity markets

transmission organizations (RTOs), which control the transmission networks and operate the electricity markets. In terms of DSM, the main task of the grid operator is to set up DSM programs with appropriate financial incentives such that electricity consumers are encouraged to participate (Walawalkar et al. 2010). In models that are typically used for market design problems, simplistic assumptions are made on consumers' load adjustment capabilities (Albadi and El-Saadany 2008; Mohsenian-Rad et al. 2010); however, such simplified models are often not sufficiently accurate to capture the real DSM potential that each consumer has as well as the costs occurring on the consumer side when implementing different DSM measures.

The electricity consumer's perspective has to be considered in order to achieve a more accurate assessment of individual consumers' DSM potentials. This knowledge will consequently lead to more active participation of consumers and make electricity markets more efficient and competitive (Kirschen 2003). Here, domain knowledge is required since each process has its own operational limitations, costs, safety requirements etc. Also, different preferences in terms of convenience and risk have to be taken into account.

One distinguishes between three consumer sectors: residential, commercial, and industrial. DSM in the first two sectors deals with residential and commercial buildings (Motegi et al. 2007), in which load adjustment is mainly achieved by controlling the HVAC and lighting systems, whereas industrial DSM is concerned with power-intensive industrial processes (Samad and Kiliccote 2012).

Scope and Organization of Review In this work, we take the standpoint of power-intensive industries, which engage in DSM as large electricity consumers. The high potential impact of industrial DSM is widely acknowledged (Paulus and Borggrefe 2011; Samad and Kiliccote 2012) and has been the focus of increased research efforts in recent years. Because of the time-sensitive nature of electricity

prices and DR events, efficient planning, scheduling, and control of plant operations are crucial for enabling effective DSM (Merkert et al. 2014). This is especially true in an industrial setting due to the high standards in terms of product quality and process safety.

The focus of this review is on models and systematic methods for industrial DSM that optimize process operations at the planning and scheduling level. This is still a relatively new research topic; therefore, with this work, we hope to provide a comprehensible introduction to industrial DSM to interested researchers, while reviewing existing works addressing the main challenges that we see in this area.

The remainder of this review is organized as follows. In Sect. 2, a definition of DSM and a classification of the DSM activities are provided. Section 3 describes the special characteristics of DSM in power-intensive process industries. Four main challenges in industrial DSM are identified: (1) accurate modeling of operational flexibility, (2) integration of production and energy management, (3) optimization across multiple time scales, and (4) decision-making under uncertainty. A state-of-the-art review of existing works addressing these four topics is presented in Sect. 4. To demonstrate some of the recently developed models, two case studies are presented in Sect. 5. Future opportunities and challenges that we see in this research area are listed in Sect. 6. Finally, in Sect. 7, we close with some concluding remarks.

2 Definition of Demand Side Management

When it comes to defining DSM, most descriptions represent the grid operator's perspective. For example, in a report released by The World Bank (Charles River Assosicates 2005), DSM is defined as the

> systematic utility and government activities designed to change the amount and/or timing of the customer's use of electricity for the collective benefit of the society, the utility and its customers.

Notice that according to this definition, only activities on the utility and government sides are considered DSM, whereas electricity consumers take a rather passive role and only react to those DSM activities. Although such a definition is perfectly correct and underlines the origin of DSM as a concept proposed by utilities (Gellings 1985), it understates the degree of freedom that consumers have in their decision-making. In fact, only the consumers can actually change their electricity consumption; utilities and governments can only provide incentives that encourage such activities. Hence, a more comprehensive definition of DSM considering both perspectives could be:

> DSM encompasses systematic activities at the interplay between grid operator and electricity consumer aiming at changing the amount and/or timing of the consumer's use of electricity in order to increase grid performance and consumer benefits. DSM activities on

the grid operator side involve the assessment of the need for load adjustment and the creation of financial incentives for the consumer, while the consumer reacts to these financial incentives and performs the actual physical load adjustment operations.

Note that depending on the level of regulation in the electricity market, the grid operator could be an independent organization or the electric utility itself. Depending on the size, an electricity consumer could be one individual consumer or an aggregator that manages many small consumers.

Figure 2 shows a general classification of DSM activities, which consist of various DSM programs introduced by the grid operator (rectangular boxes) and the measures that need to be taken by the consumer (rounded rectangles) in order to participate in these DSM programs. The two main DSM categories are energy efficiency (EE) and demand response (DR) (Charles River Assosicates 2005). The goal of EE is to reduce energy consumption while accomplishing the same tasks, and DR refers to load profile adjustment, such as load shifting and load shedding, driven by market incentives.

In DR, one distinguishes between dispatchable and nondispatchable DR (FERC 2010), which are often also referred to as incentive-based and price-based DR (DOE 2006), respectively. Dispatchable DR refers to load adjustment capacities that consumers provide to the grid operator such that these capacities can be dispatched to maintain grid stability or in times of emergency. The grid operator has control over dispatchable DR resources by either direct load control or by requesting the consumers to reduce their power consumption (interruptible load) when a DR event, e.g., a generator failure, occurs. The various types of dispatchable

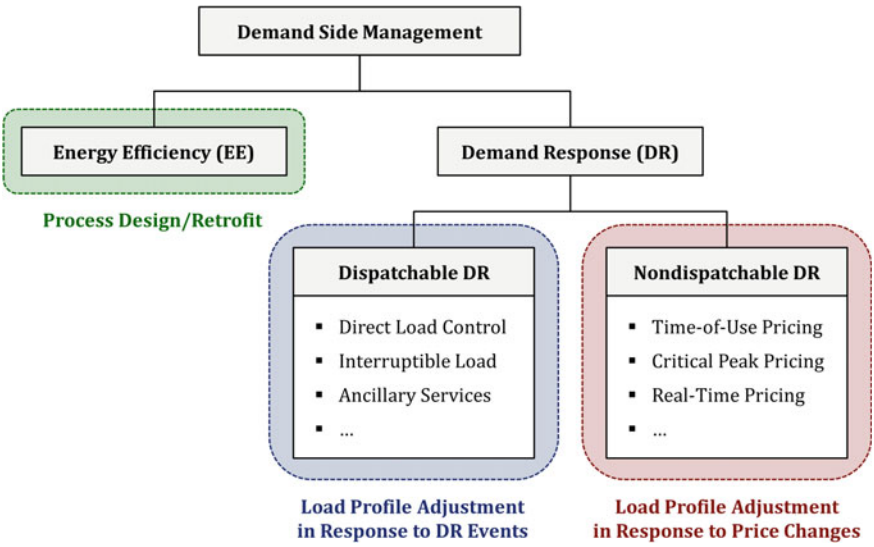


Fig. 2 Classification of DSM activities. *Rectangular boxes* depict DSM programs introduced by the grid operator, whereas *rounded rectangles* indicate measures to be taken by the consumer

DR resources, many of which classified as ancillary services, mainly differ in the amount of time within which the consumer has to respond to DR requests. In general, the faster one can react, the more valuable is the DR service, i.e., the more consumers get paid for providing such DR capacities.

Nondispatchable DR resources are not controlled by the grid operator; instead, consumers choose to adjust their power consumption profiles based on price signals from the electricity market. Time-of-use (TOU), critical peak, and real-time pricing are just three of many pricing schemes designed to encourage consumers to change their load profiles according to the power grid's needs.

EE can be primarily achieved by improved process design or retrofit of the existing process that results in higher efficiency. Also, efficiency can be increased by optimal scheduling and control strategies that maximize the time in which the process runs at its most energy-efficient operating point. Effective planning, scheduling, and control are even more critical in DR. Here, operational flexibility is key. In nondispatchable DR, electricity is treated as any other commodity that can be purchased, but with two distinct characteristics: it is difficult to store electricity, and electricity prices are extremely volatile. Hence, processes have to be flexible in order to react to price changes. In dispatchable DR, consumers are rewarded not so much for the actual dispatch of DR resources but for the capability of quickly reacting to DR events whenever they occur. Providing dispatchable DR requires a very high degree of flexibility in the consumer's process since requests by the grid operator, which typically cannot be anticipated in advance, have to be met while maintaining process feasibility and safety.

3 Characteristics of Industrial DSM

Although there are large untapped DSM potentials in all three—residential, commercial, and industrial—sectors (Gellings et al. 2006), there are some distinguishing features of industrial processes that facilitate the deployment of DSM strategies:

- In the industrial sector, individual power consumption is very high, which motivates and eases participation in DSM programs. For example, aluminum production has an energy intensity of 71 GJ per tonne (Worrell et al. 2008), and a typical aluminum plant produces hundreds of tonnes of aluminum on a daily basis.
- In most cases, advanced metering infrastructure is already in place; therefore, the capital investment required to implement DSM in industry is close to zero.
- Industrial processes operate in isolated environments such that human comfort is usually not an issue; this is in contrast to the residential and commercial sectors, where, e.g., HVAC control is constrained by the maximum decrease in comfort induced by temperature changes.

However, industrial processes are often highly complex and subject to strict safety requirements. In industrial DSM, it is therefore crucial to carefully evaluate the flexibility of each process in order to avoid detrimental disruptions caused by sudden changes in the plant operation.

In the following, we list some distinct characteristics of industrial electricity consumers, which need to be considered in the DSM decision-making process:

- Many manufacturing processes are highly integrated and have critical temporal dependencies, which have to be taken into account when operating the plant; this requires deep process knowledge.
- While direct load control is common in the residential sector, it usually cannot be applied in industry because of safety considerations.
- Electricity is difficult to store; however, most commodity products are not. Product inventory naturally increases the flexibility in plant operations and therefore allows more room for DSM.
- Industrial electricity consumers often enter into power contracts which offer special rates under given conditions.
- Large industrial power-intensive plants often have substantial onsite electricity generation capacities. Some of the generated electricity may even be sold at the market price or transferred offsite.
- Industrial usage data are typically confidential since they could reveal competition-sensitive information on operations strategies and process performance. Hence, all DSM efforts have to be managed within the same company.

According to the EIA (2012), the total net electricity demand by the U.S. industry in 2010 amounted to 850 TWh, with the five most power-intensive sectors —chemicals, primary metals, paper, food, and petroleum and coal products— combined consuming 560 TWh. Highly power-intensive processes include gas compression, electrolysis, and electric heating. Samad and Kiliccote (2012) present five real-world case studies in which industrial DSM has been successfully implemented and has generated considerable cost savings. Among the examples, the most prominent case is probably the one of Alcoa (Todd et al. 2009), which uses the operational flexibility of its aluminum smelting facilities to provide ancillary services through which the company is achieving large economic benefits.

4 Optimization of Planning and Scheduling for Industrial DSM

Because of the strong dependence of electricity price, electricity availability, and DR events on time, effective planning and scheduling tools are essential in DSM, especially in an industrial setting where complex manufacturing processes are involved. Production scheduling has been an active field of research in operations research since the 1950s (Graves 1981), and it started to attract increased attention

in the process systems engineering (PSE) community in the 1970s (Reklaitis 1982). Since then, considerable progress has been made in the modeling of production scheduling problems as well as in the development of efficient methods for solving these models. For recent reviews of works on production scheduling in PSE, we refer to Méndez et al. (2006), Maravelias (2012), and Harjunkoski et al. (2014). Furthermore, Maravelias and Sung (2009) discuss the integration of scheduling and planning, which involves longer time horizons; and Li and Ierapetritou (2008) and Verderame et al. (2010) review approaches proposed for scheduling under uncertainty.

For industrial DSM, we can leverage the tremendous advances in production planning and scheduling made over the last few decades; however, there are unique challenges that require special attention. The four main challenges that we see in planning and scheduling for industrial DSM are the following:

1. *Modeling operational flexibility*
   Electricity prices are extremely time-sensitive. In a typical day-ahead market, the price varies from hour to hour; in the real-time market, it changes every few minutes. In order to capture this time dependence and determine a process' ability to quickly respond to price changes, very detailed scheduling models are required. In these models, the representation of time and the accurate modeling of constraints on transitions between different process operating points are especially critical.
2. *Integration of Production and Energy Management*
   Traditionally, production and energy management are handled separately. For planning and scheduling, this means that first, a production scheduling problem is solved, and once the production schedule is determined, the objective of energy management is to minimize the cost for purchasing the electricity required for this particular production schedule. This sequential approach can easily lead to suboptimal solutions since possible synergies between production and energy management are not taken into account. Hence, an integrated approach that considers both parts simultaneously can be very beneficial, especially when power contracts with complex constraints are applied. Moreover, energy management can be further complicated by the presence of onsite generation and participation in dispatchable DR programs.
3. *Decision-making across multiple time scales*
   To perform DSM scheduling, a detailed model with a fine time representation is required; the typical time horizon for such a scheduling problem is one day or one week. In contrast, in long-term planning, the time horizon may span multiple months or years. In that case, however, we cannot simply apply the same detailed model with an extended time horizon because the resulting model would be computationally intractable, nor can we use an aggregate model with a coarse time representation since then we would not be able to model DSM activities. Hence, computationally efficient planning models have to be developed that can capture both long-term as well as short-term effects.

4. *Optimization under uncertainty*
   The level of uncertainty in the power grid is extremely high due to various factors such as fluctuating electricity demand, deregulation of electricity markets, and increasing penetration of intermittent renewable energy. The most apparent result of this uncertainty is the high volatility in electricity price. Also, all dispatchable DR activities are intrinsically uncertain because the consumer does not know in advance when the grid operator will request the dispatch of those DR services. Furthermore, uncertainties on the production side, e.g., regarding product demand and processing time, also exist. The major challenge lies in the accurate characterization of the relevant sources of uncertainty and optimal decision-making while considering these uncertainties.

Associated with all the above is the challenge of computational efficiency. With the incorporation of new features, the models become more complex and computational tractability becomes an issue. Many large-scale real-world problems cannot be solved by using off-the-shelf tools. Therefore, along with novel modeling approaches, efficient solution methods have to be developed in order to improve the computational performance.

In the following, we present a comprehensive review of existing works addressing the aforementioned four main challenges. All reviewed works are listed in Table 1, presented in chronological order with respect to the publication date. Table 1 shows various features of the different models, and we will refer back to this overview in the next sections.

4.1 Modeling Operational Flexibility

The key to industrial DSM is operational flexibility, which allows load profile adjustment in response to electricity market signals. In this context, a production facility's operational flexibility is mainly defined by its ability to ramp up and down production and its product inventory capacity. In order to assess the potential benefits from DSM for an industrial plant, a detailed scheduling model capturing all relevant process constraints and interactions with electricity markets is required. The development of such scheduling models has been identified as a promising research topic only very recently. From Table 1, one can see that almost all works addressing this subject have been conducted after year 2000, with the vast majority published within the past 5 years.

Relevant Industrial Processes Industrial processes considered in the literature can be grouped into two general categories: continuous production and batch production. For example, air separation and aluminum manufacturing are typical continuous processes, whereas steel production is mainly operated in batch mode. Table 1 shows for each reference whether the proposed model is primarily designed for continuous or batch production processes. Note that some of the models can also be applied to model hybrid (continuous and batch) production environments.

Table 1 Overview of reviewed works

	Continuous Production	Batch Production	Air Separation	Aluminum	Cement	Chlor-Alkali	Flour/Pulp	Machining	Steel	Power Contracts	Onsite Generation	Dispatchable DR	Stochastic Programming	Robust Optimization	Discrete-Time	Continuous-Time	LP	NLP	MILP	MINLP	CP
Daryanian et al (1989)	✓		✓												✓		✓				
Ashok and Banerjee (2001)	✓					✓									✓			✓			
Ierapetritou et al (2002)	✓		✓										✓		✓				✓		
Everett and Philpott (2002)	✓					✓							✓		✓			✓			
Ashok (2006)		✓						✓							✓			✓			
Karwan and Keblis (2007)	✓		✓												✓			✓			
Babu and Ashok (2008)	✓				✓										✓				✓		
Castro et al (2009)	✓			✓											✓	✓		✓			
Yusta et al (2010)	✓						✓								✓		✓				
Nolde and Morari (2010)		✓						✓								✓		✓			
Haït and Artigues (2011)		✓						✓								✓		✓			
Castro et al (2011)	✓			✓											✓	✓		✓			
Fang et al (2011)		✓														✓		✓			
Mitra et al (2012a)	✓		✓	✓											✓			✓			
Mitra et al (2012b)	✓		✓												✓			✓			
Wang et al (2012)		✓						✓			✓				✓			✓			
Vujanic et al (2012)		✓			✓							✓		✓	✓			✓			
Castro et al (2013)		✓						✓							✓			✓			
Artigues et al (2013)		✓														✓		✓			✓
Tan et al (2013)		✓						✓								✓		✓			
Ding et al (2014)	✓		✓												✓			✓			
Wang et al (2014)	✓				✓										✓		✓				
Mitra et al (2014)	✓		✓												✓			✓			
Hadera et al (2014)		✓						✓		✓	✓					✓		✓			
Zhang and Hug (2014)	✓			✓								✓		✓	✓			✓			
Shrouf et al (2014)		✓						✓							✓			✓			
Zhang et al (2015c)	✓		✓								✓	✓		✓	✓			✓			
Zhang and Hug (2015)	✓			✓								✓		✓	✓			✓			
Zhang et al (2016)	✓		✓							✓					✓			✓			
Zhang et al (2015b)	✓		✓							✓				✓	✓			✓			
Hadera et al (2015a)		✓						✓		✓	✓					✓		✓			
Hadera et al (2015b)	✓					✓				✓	✓				✓			✓			
Tan et al (2015)		✓						✓							✓			✓			
Zhang et al (2015e)	✓		✓								✓			✓	✓			✓			
Zhang et al (2015a)	✓		✓							✓			✓		✓			✓			

The table lists the papers in chronological order and shows various features of the applied models. The different features are shown in thematic groups

Also, Table 1 lists particular industrial processes to which each model has been applied. Cryogenic air separation and steel manufacturing, which are arguably the most complex production processes among the ones listed, have been considered more extensively. Cryogenic air separation is power-intensive because of the large amount of compression that is required in the separation and liquefaction processes. In steel manufacturing, the most power-intensive production stages are the melting process in the electric arc furnace and the hot rolling process.

Other power-intensive industrial processes considered in case studies are aluminum, cement, chlor-alkali, flour, and pulp production, and machining. Aluminum and chlor-alkali manufacturing involve electrolysis, while the high power intensities in cement, flour, pulp, and machining processes stem from mechanical operations such as grinding, milling, and turning.

Prevalent Modeling Concepts In one of the first works on DSM scheduling, Daryanian et al. (1989) propose a multiperiod model that merely consists of inventory constraints and bounds on the production rate. However, industrial processes are seldom that simple, and more accurate representations require detailed models involving more complex constraints.

In cases where the scheduling problem is primarily concerned with the sequencing and timing of power-intensive production tasks, typical machine scheduling formulations have been applied in several proposed models (Nolde and Morari 2010; Fang et al. 2011; Wang et al. 2012). Tan et al. (2013) and Hadera et al. (2015a) further include constraints on waiting times between consecutive production stages, which are especially important in steel manufacturing.

Many chemical production environments exhibit a network structure (Maravelias 2012) in which material handling constraints play an essential role. The scheduling of such processes is the focus of the works by Ashok and coworkers (Ashok and Banerjee 2001; Ashok 2006; Babu and Ashok 2008), who emphasize the impact of storage in industrial DSM. For the same purpose, Ding et al. (2014) apply the well-known concept of the state-task network (STN) (Kondili et al. 1993), in which state nodes represent feeds, intermediates, and final products, and task nodes represent processing operations. A similar concept is the one of the resource-task network (RTN) (Pantelides 1994), which forms the basis for several DSM scheduling models proposed by Catro and coworkers (Castro et al. 2009, 2011, 2013).

Another popular modeling concept is based on the notion of operating modes. First proposed by Ierapetritou et al. (2002) and further developed by Karwan and Keblis (2007), it takes into account that production and power consumption characteristics can vary within the same process depending on the configuration or state in which the process is operating. In a mode-based model, the process can only operate in one of the given operating modes, and each mode is defined by a specific feasible region in the product space and a power consumption function with respect to the production rates. Mitra et al. (2012a, 2013) reformulate the model by Karwan and Keblis (2007) to improve the tightness of the formulation, and develop additional constraints to

impose restrictions on the transitions between different modes. The proposed mode transition constraints are as follows:

$$\sum_{m} y_{mt} = 1 \quad \forall t \tag{1a}$$

$$\sum_{m' \in TR^{f}_{m}} z_{m'm,t-1} - \sum_{m' \in TR^{t}_{m}} z_{mm',t-1} = y_{mt} - y_{m,t-1} \quad \forall m,\ t \tag{1b}$$

$$y_{m't} \geq \sum_{k=1}^{\theta_{mm'}} z_{mm',t-k} \quad \forall (m,\ m') \in TR,\ t \tag{1c}$$

$$z_{mm',t-\bar{\theta}_{mm'm''}} = z_{m'm''t} \quad \forall (m,\ m',\ m'') \in SQ,\ t \tag{1d}$$

where the binary variable y_{mt} takes the value 1 if mode m is selected in time period t. Equation (1a) states that one and only one mode has to be selected in each time period. Equation (1b) ensures that the binary variable $z_{mm't}$ takes the value 1 if and only if the plant switches from mode m to mode m' at time point t. Here, TR^{f}_{m} denotes the set of modes from which mode m can be directly reached; similarly, TR^{t}_{m} denotes the set of modes which can be directly reached from mode m. Equation (1c) states that after switching from mode m to mode m', the plant has to remain in mode m' for at least $\theta_{mm'}$ time periods. Moreover, for an allowed mode transition sequence from m to m' to m'', Eq. (1d) can be used to fix the number of time periods that the plant has to remain in mode m' to $\bar{\theta}_{mm'm''}$.

Mode-based models have been applied in some of the most recent works on DSM planning and scheduling (Shrouf et al. 2014; Zhang et al. 2015a, b, c, e, 2016). In the most recent development, Zhang et al. (2016) have further generalized the model such that it can also be used to represent continuous process networks.

Time Representation One important attribute of scheduling models is the representation of time, which is especially critical in DSM applications because of the highly time-sensitive nature of electricity prices. In general, one distinguishes between discrete- and continuous-time models, and there is a large body of work in the literature discussing different formulations and their strengths and limitations (Floudas and Lin 2004; Mendez et al. 2006; Sundaramoorthy and Maravelias 2011).

As shown in Table 1, predominantly discrete-time models have been used in DSM scheduling. In a discrete-time representation, in which the scheduling horizon is divided into time periods of known lengths, it is straightforward to model the time-varying electricity price by simply assigning different price values to different time periods. In most cases, hourly electricity prices are considered, with a scheduling horizon of one day or one week resulting in 24 or 168 time periods, respectively.

Unlike discrete-time models, a continuous time representation allows processing tasks to start at any point in the continuous time domain. In scheduling problems in which tasks can change within small time intervals, continuous-time models can be beneficial since an appropriate discrete-time model may require a very fine time discretization, which could dramatically increase the size of the model. However, with time-varying electricity prices, modeling the electricity cost becomes a challenge in continuous-time formulations. Castro et al. (2009) propose a continuous-time formulation in which different electricity prices e are defined over price periods $p \in P_e$ with starting times L_{ep} and ending times U_{ep}. Also, tasks are disaggregated into tasks executed at different electricity prices, forming the price-dependent sets of tasks $\bar{I}_e$. This time representation is illustrated in Fig. 3, where τ_t denotes the absolute time at event point t, and $\hat{\tau}_t$ denotes the starting time of tasks executed during time interval t. The corresponding timing constraints are:

$$\sum_{e} \sum_{p \in P_e} y_{t\,pe} = 1 \quad \forall t \neq |T| \tag{2a}$$

$$\hat{\tau}_t \geq \sum_{e} \sum_{p \in P_e} L_{ep}\, y_{t\,pe} \quad \forall t \neq |T| \tag{2b}$$

$$\hat{\tau}_t + \sum_{i \in \bar{I}_e} \frac{\bar{\mu}_{ri} \xi_{it}}{\rho_i^{\max}} \leq \sum_{p \in P_e} U_{ep} y_{t\,pe} + M \left(1 - \sum_{i \in \bar{I}_e} \bar{\mu}_{ri}\, z_{it} \right) \quad \forall t \neq |T|,\ e,\ r \in R^{\mathrm{TC}} \tag{2c}$$

where the binary variable $y_{t\ \mathrm{pe}}$ takes the value 1 if during time interval t, tasks are processed within price period $p \in P_e$. Equation (2a) states that there is one active price period during each time interval. Equations (2b) and (2c) ensure that a task starts and finishes within the corresponding price period. Here, $\bar{\mu}_{ri}$ is the discrete interaction of resource r with task i at its end, ξ_{it} is the amount handled by task i during interval t, $\rho_i^{\max}$ is the maximum processing rate of task i, M is a big-M parameter, z_{it} is a binary variable that takes the value 1 if task i is executed in interval t, and R^{TC} denotes the set of equipment resources involved in timing constraints. Note that the disaggregation of a task into tasks executed at different electricity prices with the described method is only valid for continuous processes.

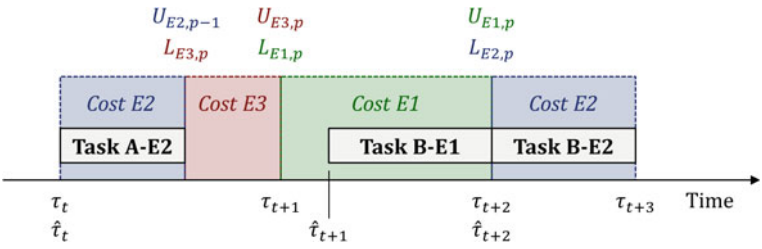


Fig. 3 Illustration of continuous time representation with timing constraints (Castro et al. 2009)

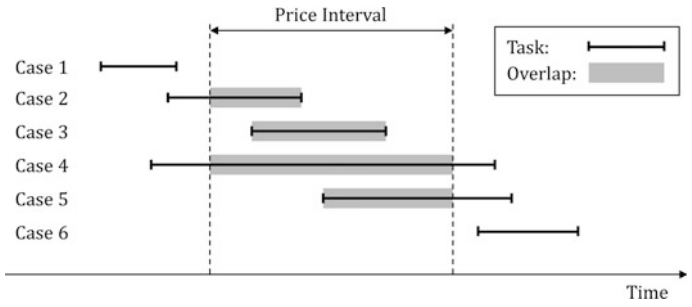


Fig. 4 Six possible task-interval overlap cases (Nolde and Morari 2010)

Nolde and Morari (2010) propose a continuous-time model for load tracking in which the electricity consumption in prespecified load intervals is determined by computing the overlap of tasks with the load intervals. The same concept can be applied to determine electricity cost when electricity prices vary with time (Tan et al. 2013). The idea is illustrated in Fig. 4, which shows six different task-interval overlap cases that need to be considered for each task-interval pair. Nolde and Morari (2010) present a formulation using binary variables and corresponding big-M constraints for each of the six cases. This formulation has been improved by Hait and Artigues (2011) who compute the overlaps by introducing binaries indicating whether a task begins before or during a price interval. This reformulation greatly reduces the number of constraints and binary variables. A continuous-time model incorporating a similar approach has been proposed by Hadera et al. (2015a). While applying the same concept to maintenance scheduling of a gas engine power plant, Castro et al. (2014) propose a further improved reformulation derived by using generalized disjunctive programming techniques Grossmann and Trespalacios, 2013) and by introducing redundant constraints.

Computational performance is often the key criterion when choosing between discrete- and continuous-time scheduling models. An analysis of the reviewed works shows that at this point, discrete-time models generally show better computational performance when applied to large-scale problems. Castro et al. (2009) show that only problems of small size can be handled effectively by the continuous-time model, while problems of industrial significance can be solved efficiently with a comparable discrete-time model. In order to mitigate this limitation, Castro et al. (2011) propose an aggregate discrete-time model that is used in conjunction with a continuous-time model in a rolling horizon framework. The computational results show that the proposed solution approach is considerably more efficient than both full-space traditional discrete- and continuous-time models. However, under restricted power availability, the rolling horizon approach may lead to suboptimal solutions, in which case the full-space discrete-time model becomes the better choice. Hadera et al. (2015a) apply a heuristic bilevel decomposition approach to solve the proposed continuous-time model. The decomposition approach significantly reduces the solution time; however, the problem is still only tractable for a one-day scheduling horizon.

Types of Models From Table 1, we can see that the vast majority of the reviewed models are formulated as mixed-integer linear programs (MILPs). Two very simplistic models (Daryanian et al. 1989; Wang et al. 2014) are formulated as linear programs (LPs). A nonlinear programming (NLP) formulation is proposed by Yusta et al. (2010) for machining process scheduling, where the nonlinearity stems from the equation expressing the lifetime of the cutting tool as a nonlinear function of the cutting speed and the power consumption function.

Ierapetritou et al. (2002) apply a quadratic power consumption function, which gives rise to a mixed-integer nonlinear programming (MINLP) formulation. Generalized Benders decomposition (Geoffrion 1972) and outer approximation (Duran and Grossmann 1986) have been applied to solve the problem. Since the MINLP is convex in the continuous variables, both solution algorithms are guaranteed to obtain the global optimal solution; however, they require considerable computational expense for solving industrial-scale problems. Hence, Ierapetritou et al. (2002) create an approximate MILP by linearizing the power consumption function and solve the MILP instead of the original MINLP. The results show that the MILP model can be solved in significantly less computational time and obtains solutions that are very close to the true optimal solutions of the MINLP. Babu and Ashok (2008) propose an MINLP model with quadratic functions expressing the power factors and efficiencies of each subprocess. The model has been solved with the global optimization algorithm implemented in the LINDO solver.

Clearly, the preference of linear over nonlinear models is a result of the fact that in general, linear models are considerably easier to solve than nonlinear ones. However, many processes exhibit nonlinear behavior that needs to be approximated in linear models. A recent attempt to improve the accuracy of the process representation in MILP models has been proposed by Zhang et al. (2015c, 2016). Here, the idea is to approximate the generally nonconvex feasible region of a process or a certain configuration of a process by a set of polytopes with a linear power consumption function associated with each polytope. Such models are referred to as Convex Region Surrogate (CRS) models (Zhang et al. 2015d). The advantage of a CRS model is that it captures nonlinearities and nonconvexities, yet it can be formulated as a set of mixed-integer linear constraints, which do not increase the computational complexity of the scheduling model if it is already formulated as an MILP.

Moreover, Artigues et al. (2013) show an interesting application of constraint programming (CP) in industrial DSM. Here, the scheduling problem is solved in two steps. In the first step, a job assignment and sequencing problem with fixed job durations is solved with a CP model; then in the second step, an MILP scheduling model is solved with the job assignment and sequencing obtained in the first step. Obviously, this is another attempt to reduce the computational effort by decomposing the problem.

Main Insights from Case Studies A number of case studies have been presented in the reviewed references, some using real-world data from industry. Here, we summarize some of the main insights drawn from these case studies in which scheduling under time-sensitive electricity prices has been considered.

All case study results show the high potential benefit of industrial DSM, accomplished primarily by shifting load toward low-price periods. Cost savings up to 20 % can be achieved compared to scheduling assuming constant electricity prices (Castro et al. 2009). The optimization even shows that in certain cases, it can be beneficial to shut down the plant for a long period of time (Ierapetritou et al. 2002; Mitra et al. 2012a). However, the impact of DSM strongly depends on the level of operational flexibility. In particular, if a plant is highly utilized, i.e., it has to operate at close to full capacity in order to meet the product demand, there will be hardly any room for load shifting (Ashok 2006; Mitra et al. 2012a). Therefore, the benefit of nondispatchable DR usually decreases with increasing level of plant utilization; however, this may be different in the case of dispatchable DR.

4.2 *Integration of Production and Energy Management*

Industrial DSM relies on the integrated optimization of production and energy management. In most existing works on DSM planning and scheduling, the energy management part only consists of purchasing electricity at time-varying prices with possibly an upper bound constraint on the amount of electricity that can be purchased in each time period. As indicated by the overview shown in Table 1, only very recently, more complex energy management activities involving power contracts with various price structures, onsite generation, and dispatchable DR have been considered.

Nolde and Morari (2010) consider a load tracking problem in which the actual electricity consumption of a steel plant is supposed to match a committed load curve; penalties incur for over- and underconsumption. Besides load tracking, the models proposed by Hadera et al. (2014, 2015a) account for multiple electricity sources as well as onsite generation, which generates electricity that can be either used to power the steel plant or sold to the electricity market. A network flow formulation is applied to incorporate the electricity purchase and sales options into the scheduling model. Furthermore, Hadera et al. (2015b) have developed a mean value cross decomposition approach that can help reduce the computational effort for solving such large-scale integrated problems. Results from case studies show that often a substantial amount of electricity is generated onsite, some of which is sold to the market in order to reduce the net electricity cost. Also, the penalties for deviating from the committed load curve can be quite significant and are often in the same order of magnitude as the net electricity cost. Onsite generation is also considered in the model proposed by Wang et al. (2012), which further includes fuel storage constraints and gas emissions in the objective function.

Zhang et al. (2015b) consider the integrated optimization of short-term production scheduling and electricity procurement. In the proposed model, power contracts are assumed to have two price components: a time-dependent and an amount-dependent component. A block contract formulation is applied to model the amount-dependent price component, which allows discount prices depending on the

purchase amount. To also accommodate penalty contracts, Zhang et al. (2016) propose a more general block contract formulation which can be expressed as the following disjunction:

$$\bigvee_{b \in B_c} \begin{bmatrix} X_{cbt} \\ H_{cb't} = H_{cb'}^{\max} \quad \forall b' \in B_c,\ b' < b \\ H_{cb} \leq H_{cb}^{\max} \\ H_{cb't} = 0 \quad \forall b' \in B_c,\ b' > b \end{bmatrix} \quad \forall c,\ t \tag{3a}$$

$$\bigvee_{b \in B_c} X_{cbt} \quad \forall c,\ t \tag{3b}$$

$$X_{cbt} \in \{true,\ false\} \quad \forall c,\ b \in B_c,\ t \tag{3c}$$

where B_c is the set of blocks for contract c, H_{cbt} denotes the amount of electricity purchased in block $b \in B_c$ in time period t, and $H_{cb}^{\max}$ is the amount of electricity that one has to purchase in block $b \in B_c$ before reaching the next block. X_{cbt} is a Boolean variable that is true if block b is the highest block reached for contract c in time period t. Disjunction (3a) states that if X_{cbt} is true, the maximum amount is purchased in all lower blocks $b' < b$, the electricity purchase in block b is bounded by $H_{cb}^{\max}$, and no electricity is purchased in higher blocks $b' > b$. According to logic constraint (3b), one and only one X_{cbt} has to be true. By applying the hull reformulation (Balas 1985), Eqs. (3a)–(3c) can be transformed into a set of mixed-integer linear constraints and incorporated into the MILP scheduling model. Figure 5 illustrates how discount and penalty contracts can be represented as block contracts with appropriate prices assigned to the corresponding blocks. Note that in the example shown in Fig. 5b, penalties incur for both over- and underconsumption.

Vujanic et al. (2012), Zhang and Hug (2014, 2015), and Zhang et al. (2015c, e) develop systematic approaches for the optimization of dispatchable DR activities,

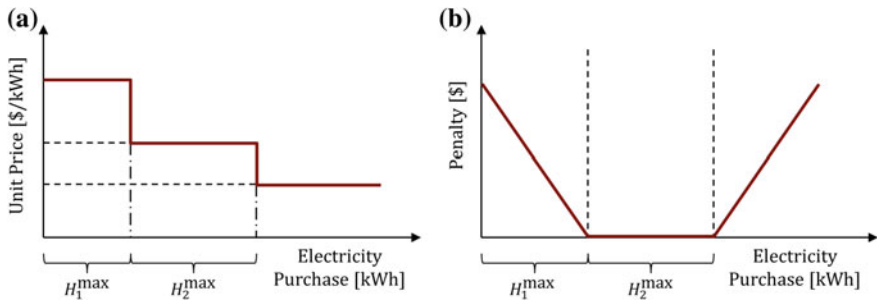


Fig. 5 Discount and penalty contracts can be modeled as block contracts. **a** Discount contract. **b** Penalty contract (Zhang et al. 2016)

e.g., the provision of regulation and operating reserve services. Participation in dispatchable DR programs creates new unconventional revenue streams that can significantly reduce net operating costs; however, the process is associated with high degree of uncertainty due to the unpredictability of DR events. Handling these cases requires advanced stochastic optimization methods; hence, we postpone further detailed discussion of this topic to Sect. 4.4.

4.3 Decision-Making Across Multiple Time Scales

In long-term DSM planning, tactical and strategic decisions may be considered, such as investment decisions for capacity expansion and retrofit, and the selection of long-term power contracts. Since planning problems involve much longer time horizons, they are typically solved using models that are considerably less detailed than short-term scheduling models. However, such aggregate models cannot be applied to industrial DSM problems because they do not capture time-sensitive electricity prices and DR events with sufficient accuracy.

Integrated decision-making across multiple time scales—short-term operational and long-term tactical and strategic DSM decisions—has barely been considered in the literature. Mitra et al. (2014) propose a multiscale capacity planning model for power-intensive continuous processes considering hourly changes in electricity price. The objective is to find the optimal investment strategy for purchasing new equipment, performing equipment upgrades, and installing additional storage facilities over a planning horizon of multiple years. Instead of applying a detailed representation across the entire time horizon, which would be computationally intractable, the model is simplified by leveraging the seasonality of electricity prices. Here, each year is divided into four seasons, and each season is represented by one week, which is repeated cyclically and characterized by a typical electricity price profile that reflects the price's seasonal behavior. An hourly time discretization is applied, which results in 672 time periods representing each year (4 seasons, each with a week divided into 168 time periods). While the number of time periods is rather large, it is considerably smaller compared to the 8,760 time periods that would be required to represent hourly discretization over one year.

Zhang et al. (2015a) apply a similar approach, but represent each season of the year by 2 weeks. While the first week is repeated cyclically, the second is used to impose mass balance constraints between adjacent seasons. In this way, unlike in the model proposed by Mitra et al. (2014), a consistent inventory profile throughout the year is achieved since inventory can be carried over from one season to the next. Using this modeling idea, which is illustrated in Fig. 6, Zhang et al. (2015a) solve a long-term electricity procurement problem under uncertainty over a one-year planning horizon.

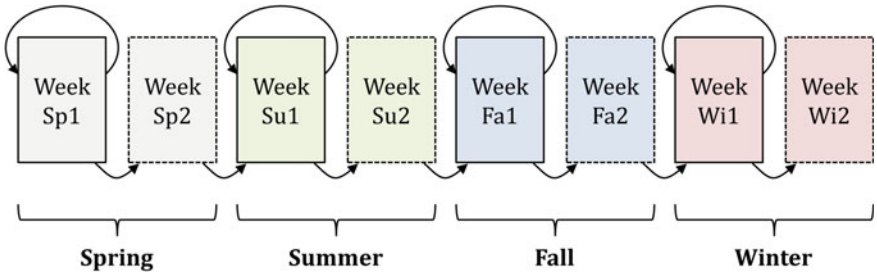


Fig. 6 In the multiscale model proposed by (Zhang et al. 2015a), each of the four seasons of the year is represented by 2 weeks, where the first has a cyclic schedule and the second is noncyclic

4.4 Optimization Under Uncertainty

Most existing planning and scheduling tools for DSM are deterministic, i.e., they assume that all given input parameters are certain, including future electricity prices. However, this assumption is rarely valid, especially in the case of spot electricity prices, which are extremely difficult to forecast (Zareipour et al. 2010). In the light of the high level of uncertainty in DSM problems, optimization models have been developed that consider the characteristics of the uncertain parameters instead of simply assigning to them their expected values. In this way, a more realistic representation of the problem can be achieved, which forms the basis for improved decision-making.

Two major uncertainty modeling approaches have been applied to DSM planning and scheduling problems: stochastic programming (Birge and Louveaux 2011), and robust optimization (Ben-Tal et al. 2009). In stochastic programming, the uncertainty is represented by discrete scenarios with given probabilities, and decisions are made at different stages, which are defined such that realization of uncertainty is observed between two stages. At each stage, actions depending on previous observations are taken; such reactive actions are also referred to as recourse decisions. In robust optimization, the uncertainty is specified in terms of an uncertainty set from which any point is a possible realization of the uncertainty. A robust optimization model is formulated such that it is feasible for the entire uncertainty set and optimizes the worst case.

Modeling Electricity Price Uncertainty Ierapetritou et al. (2002) present a two-stage stochastic programming framework in which the electricity prices for the first 3 days of the scheduling horizon are assumed to be known while the prices for the remaining days are assumed to be stochastic. The uncertain prices are characterized by a set of scenarios with each scenario corresponding to a particular price profile for the time beyond the first 3 days. While the first-stage decisions are related to the first 3 days, the second-stage decisions are related to the remaining days of the scheduling horizon and can be different for each scenario. Given probabilities for each scenario, the objective is to minimize the total expected

operating cost. A similar approach is taken by Everett and Philpott (2002) who assume that all future electricity prices are uncertain; hence, each scenario represents a price profile over the entire scheduling horizon.

Monte Carlo simulation with a stochastic price forecasting model can be used to generate electricity price profiles that are needed in a scenario-based approach. However, in order to accurately characterize the price uncertainty over a longer period of time, many price profiles are required, which leads to a large-scale stochastic programming model that may be computationally intractable. This limitation has motivated Mitra et al. (2012b) to apply a robust optimization approach to model uncertain electricity prices, which features different possible price ranges and can account for correlated data. Here, the problem is formulated as follows:

$$\min_{x} \max_{c,z} \sum_{t} \alpha_t E_t \tag{4a}$$

$$\text{s.t. } \alpha_t = \sum_{t'} \zeta_{tt'} c_{t'} \quad \forall t \tag{4b}$$

$$c_{tk}^{\min} z_{tk} \le \bar{c}_{tk} \le c_{tk}^{\max} z_{tk}, z_{tk} \in \{0,1\} \quad \forall t,\ k \tag{4c}$$

$$c_t = \sum_{k} \bar{c}_{tk}, \sum_{k} z_{tk} = 1 \quad \forall t \tag{4d}$$

$$\sum_{t} z_{tk} \le \Gamma_k \quad \forall k \tag{4e}$$

$$\text{s.t. } x \in X \tag{4f}$$

where x denotes the operational decision variables including E_t, the amount of electricity purchased in time period t. Given the scheduling constraints that are generally represented by Eq. (4f), this bilevel formulation minimizes the cost of purchasing electricity for the worst case within the uncertainty set defined by Eqs. (4b)–(4e). As shown in Eq. (4b), the electricity price in time period t, α_t, is stated as a weighted sum of random variables $c_{t'}$, which allows the modeling of correlations between electricity prices across different time periods. To have a more realistic characterization of the uncertainty as well as to allow adjustment of the level of conservatism, the uncertainty space for the electricity price is split into multiple price ranges (Duzgun and Thiele 2015). Equations (4c)–(4d) ensure that only one price range k is chosen in each time period t, and the parameter Γ_t in Eq. (4e) represents an upper bound on the number of time periods in which the electricity price is within price range k. Finally, by applying duality theory, the inner maximization problem is reformulated into a minimization problem, which transforms the overall bilevel formulation into a tractable single-level minimization problem.

Electricity Price and Product Demand Uncertainty, and Risk Since DSM comprises both production and energy management, only accounting for

uncertainty related to electricity is often insufficient. Uncertainty on the production side may have a different and possibly larger impact on the plant operations, and decision-making in the presence of multiple sources of uncertainty is certainly nontrivial. In their proposed stochastic programming model, Mitra et al. (2014) consider product demand uncertainty, which is a parameter that can have a profound impact on the solution and is often associated with high degree of uncertainty. For different cases, Mitra et al. (2014) compute the value of stochastic solution (VSS), which measures the improvement in the objective function value achieved by solving the stochastic model compared to the solution obtained from the deterministic model. The results show that the VSS can be quite significant, especially for skewed demand distributions with large standard deviations.

Zhang et al. (2015b) consider both electricity price and product demand uncertainty while solving an integrated production scheduling and electricity procurement problem. In the proposed two-stage stochastic programming framework, operating modes are selected and the amount of electricity purchased from power contracts with known prices is determined in the first stage, while recourse decisions in the second stage are the actual production rates and the amount of electricity purchased from the spot market where the prices are uncertain. Multiple modeling and solution strategies have been applied to address real-world problems of relevant size: (1) incorporation of conditional value-at-risk (Rockafellar and Uryasev 2000) as risk measure; (2) generation of electricity price scenarios using ARIMAX models; (3) scenario reduction (Dupacova et al. 2003) to obtain a small subset of scenarios that still represent the uncertainty relatively well; and (4) multicut Benders decomposition (Birge and Loueveaux 1988) to reduce the solution time. The case studies show that there can be substantial differences between the solutions obtained from deterministic, risk-neutral, and risk-averse optimization. Also, an extensive analysis of the VSS shows that in risk-neutral optimization, accounting for electricity price uncertainty does not result in significant additional benefit, whereas in risk-averse optimization, modeling price uncertainty is crucial for obtaining good solutions.

Uncertainty in Dispatchable DR The development of systematic decision-making tools for dispatchable DR has not been attempted until recently. The main challenge lies in the inherent uncertain nature of the problem since the consumer does not know in advance when the dispatch of the provided DR resources will be requested.

Zhang and Hug (2014) apply a stochastic programming approach to optimize the provision of regulation capacity by aluminum smelters. Likewise, by applying a scenario-based approach similar to the one proposed by Conejo et al. (2002) for electricity producers, Zhang and Hug (2015) derive a bidding strategy for aluminum smelters. In the bidding process, participants state how much energy or operating reserve capacity they are willing to sell at which price. In the proposed stochastic programming framework, the price is the uncertain parameter and a scheduling problem is solved for each price scenario. The solution provides price-amount pairs for each scenario, which can be used to construct the bidding curve.

In the above-mentioned stochastic programming approaches, the same probabilities are assumed for all scenarios. This assumption is usually not realistic; in fact, it is very difficult to obtain reasonable probability distributions for dispatchable DR events. Furthermore, when providing operating reserve, dispatch upon request has to be guaranteed since otherwise, one has to pay very high penalties or may not even be allowed to participate in the market. Hence, robust solutions are required. Vujanic et al. (2012) consider uncertainty in the start times of scheduled tasks, which may be caused by load shifting required to meet operating reserve demand. Robust optimization has been applied to ensure feasibility for any changes in task start times within prespecified ranges. Note that the proposed model only identifies the provision of operating reserve as the source of the uncertainty, but does not actually optimize it.

Zhang et al. (2015c) assess the operational benefit of adding cryogenic energy storage capabilities to air separation plants. Among other things, the energy storage can be used to provide operating reserve since it can be quickly ramped up to generate electricity in case of grid contingencies, which is when operating reserves are needed. Here, a robust optimization approach is applied, for which the following uncertainty set U is proposed:

$$U(R) = \left\{ w : \left(D_t = R_t w_t,\ 0 \le w_k \le 1,\ k = 1, \ldots, t,\ \sum_{k=1}^{t} w_k \le \Gamma_t \right) \forall t \right\} \tag{5}$$

where D_t denotes the operating reserve demand in time period t, which is uncertain, and R_t is the reserve capacity provided by the plant in time period t. The normalized reserve demand in time period t, w_t, takes a value between 0 and 1. The parameter Γ_t, which defines the maximum number of time periods in which maximum reserve dispatch can occur up to time t, can be used to adjust the level of conservatism. Note that the uncertainty set depends on the amount of provided reserve capacity.

A major drawback of traditional robust optimization approaches is their inability to account for recourse, which makes the formulation overly conservative. A recent advance in the robust optimization area is the development of so-called adjustable robust formulations (Ben-Tal et al. 2004), in which recourse actions are defined in terms of affine decision rules. Zhang et al. (2015e) apply this approach to optimize the provision of interruptible load, which is operating reserve that large electricity consumers can provide by reducing (interrupting) their load. Here, the production rates, which directly affect the electricity consumption, are the decision variables that need to depend on the uncertainty. Considering the uncertainty set given by Eq. (5), the affine decision rule is stated as follows:

$$Q_{it} = \overline{Q}_{it} + \sum_{k=1}^{t} q_{it\,k}\, w_k \tag{6}$$

where Q_{it} is the actual amount of product i produced in time period t, which consists of the target production rate $\overline{Q}_{it}$ (for the case where reserve demand is zero) and a recourse term. As one can see, Q_{it} is allowed to depend on uncertain parameters

realized in all previous time periods, which makes this a multistage formulation. The coefficients $q_{it\ k}$ are also part of the decision variables. Besides forcing the production rates to decrease when interruptible load is requested, the decision rule also allows the production rates to increase after interruptible load events to make up for lost production. The case study results show that much improved solutions can be achieved by solving the proposed model compared to applying the traditional static robust optimization approach.

5 Case Studies

In the following, we present two case studies in which models proposed to address some of the aforementioned challenges are applied to real-world industrial problems. In both cases, the industrial process under consideration is air separation, for which real plant data have been provided by Praxair. Note, however, that the proposed model formulations are general and can be applied to other similar continuous power-intensive processes. The models are implemented in GAMS 24.4.1 (GAMS Development Corporation 2015) and solved by using the commercial solver CPLEX 12.6.1 on an Intel® Core™ i7-2600 machine at 3.40 GHz with 8 hardware threads and 8 GB RAM running Windows 7 Professional.

5.1 Scheduling of Process Networks with Various Power Contracts

The first case study is taken from Zhang et al. (2016) and deals with the optimal scheduling of continuous power-intensive process networks while considering various power contracts including discount and penalty contracts. Figure 7 shows the process network of the given air separation plant. The gaseous oxygen (GO2), liquid oxygen (LO2), liquid nitrogen (LN2), and liquid argon (LAr) flowing out of the air separation (AS) process can be directly sold whereas gaseous nitrogen (GN2) has to be further compressed before it can be supplied to the customers. Two kinds of GN2 are sold: medium-pressure GN2 (MPGN2) and high-pressure GN2 (HPGN2). GN2 is compressed to MPGN2 through Process LMCompGN2 and can be further compressed to HPGN2 through Process MHCompGN2; it can also be directly converted to HPGN2 by running Process LHCompGN2. Furthermore, GN2 can be liquefied to LN2 by running Process LiqGN2. Overproduced gaseous products can be vented through a venting process, and all liquid products can be converted into the corresponding gaseous products through a so-called driox process.

Figure 8 shows the electricity consumption profiles for each of the processes. The vast majority of the electricity consumption is attributed to the AS process. The GN2 liquefier also consumes a large amount of electricity but is only used five times, each time for a few hours. Compared to the air separation unit and the GN2

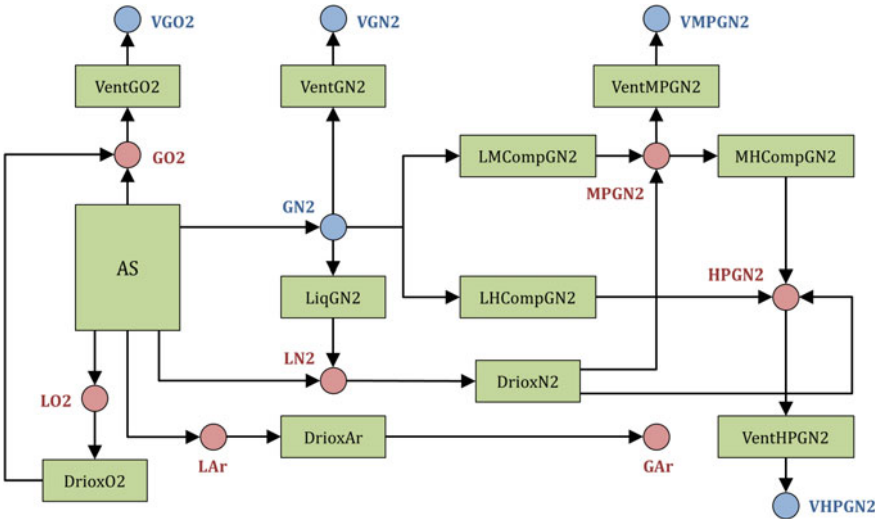


Fig. 7 Process network representing the given air separation plant (Zhang et al. 2016)

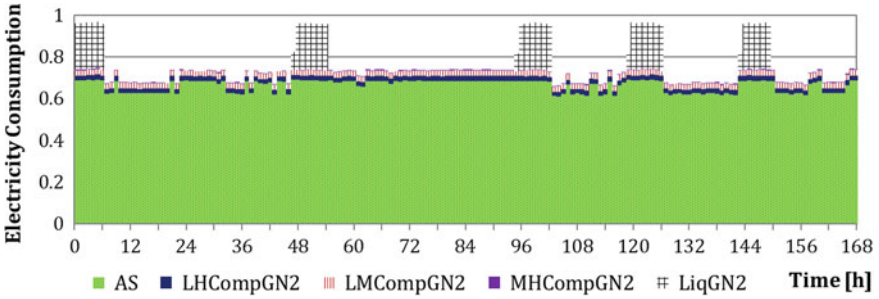


Fig. 8 Amount of electricity consumed by each process of the air separation plant

liquefier, the pipeline compressors contribute relatively little to the total electricity consumption. Significant load shifting can be observed in the schedule; this is mainly realized by operating the liquefier during low-price hours, which allows a fairly constant operation of the other processes.

Three different sources from which electricity can be purchased are considered: a TOU contract, a penalty contract with penalty for underconsumption, and the spot market. The corresponding electricity prices are shown in Fig. 9. The breakdown of the total electricity purchase into the purchases from the three different sources is shown in Fig. 10. One can observe that in each time period, we choose to purchase from the source with the lowest price. Two sources are chosen in the same time period only when the maximum purchase amount is reached for one of the two sources. Also, sufficient amount of electricity is purchased from the penalty contract such that no penalty has to be paid.

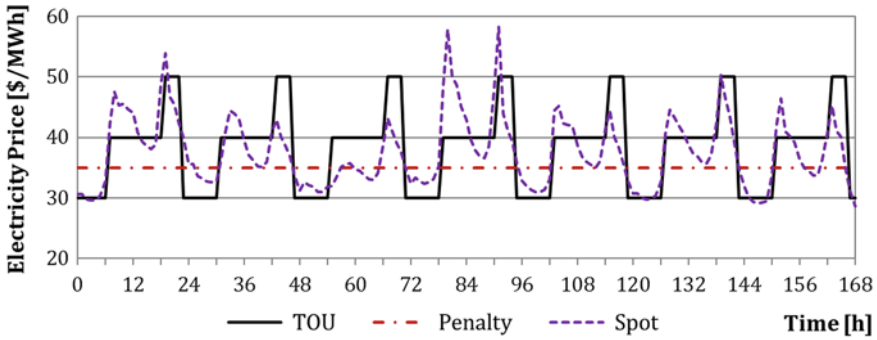


Fig. 9 Electricity prices for the TOU contract, the penalty contract, and the spot market

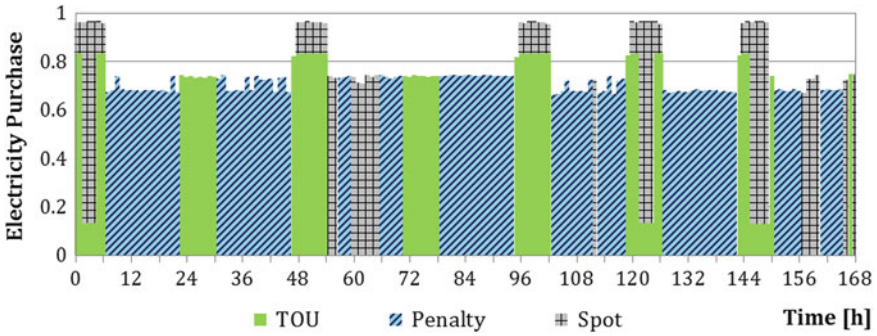


Fig. 10 Breakdown of electricity purchase into purchases from the three difference sources

This case study demonstrates the ability of state-of-the-art optimization models to accurately represent complex process systems and incorporate various power contract structures. One major strength of the proposed discrete-time scheduling model is the computationally efficient formulation. Zhang et al. (2015b) show that large-scale problems with tens of thousands of variables and hundreds of thousands of constraints can be solved to optimality within a few minutes, which allows the use of such a scheduling tool in a real industrial setting. Note that in this case study, a deterministic model has been applied, i.e., no uncertainty in the model parameters has been considered.

5.2 *Risk-Based Integrated Production Scheduling and Electricity Procurement*

The second case study is taken from Zhang et al. (2015b) and considers the integrated production scheduling and electricity procurement under spot electricity

price and product demand uncertainty, with the possibility of purchasing additional products in the case of production shortfall. The proposed two-stage stochastic programming model incorporates CVaR as a measure of risk. Here, the main first-stage decisions are the selection of operating modes and the committed amount of electricity purchase from power contracts; second-stage decisions are the actual production rates, the amount of purchased product, and the amount of electricity purchased from the spot market. For a particular air separation case, Fig. 11 shows the electricity purchase and corresponding price profiles from the deterministic, the risk-neutral stochastic, and the risk-averse stochastic solutions. We highlight the significant differences between these three solutions.

In deterministic optimization, uncertainty is ignored such that decisions are primarily driven by the differences between the power contract prices and the expected spot price. A significant amount of electricity is procured from Contract 2 because the price discount at this purchasing amount makes it less expensive than purchasing from the spot market during on-peak hours. However, there is a very

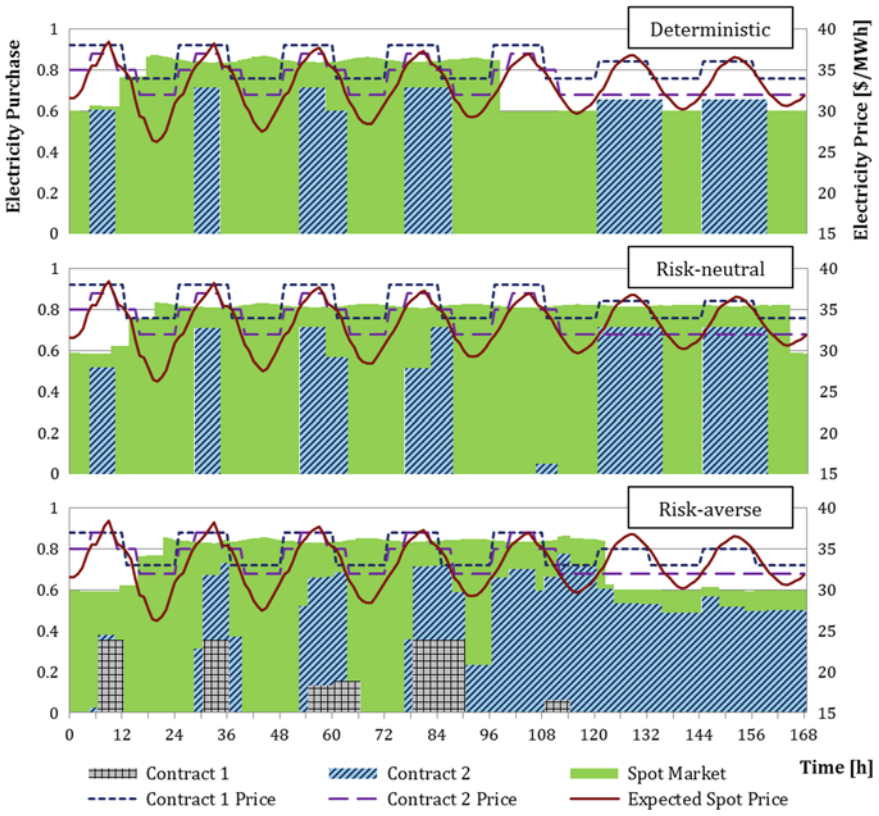


Fig. 11 Electricity purchase profiles obtained from solving the deterministic, the risk-neutral, and the risk-averse optimization models (Zhang et al. 2015b)

high expected cost of purchasing additional products because the selected operating modes do not have sufficient production capacities in the high-demand scenarios.

In risk-neutral stochastic optimization, first-stage decisions are to a great extent driven by the need for flexibility that has to be maintained in the second stage in order to react to different scenarios. Deterministic and risk-neutral optimizations lead to similar schedules for the electricity procurement from power contracts. However, the risk-neutral solution suggests selecting operating modes with higher production capacities in order to be able to accommodate high-demand scenarios. This strategy leads to considerably lower costs for purchasing additional products, especially during the last 3 days of the week. In this case, the relative reduction in total expected cost amounts to 7.2 %.

Unlike the deterministic and risk-neutral solutions, the risk-averse solution suggests to purchase more than half of the required electricity from power contracts. Here, contracts are effectively used to hedge against the risk of low-profit scenarios. In particular, considerable amount of electricity is purchased from contracts toward the end of the week when the level of uncertainty in the spot electricity price is highest. The relative VSS in this case is 8.3 %.

It is worth noting that the stochastic models were solved by applying a multicut Benders decomposition algorithm (Birge and Louveaux 1988). When solving the full-space risk-averse model, for example, CPLEX was not able to find a feasible solution within 2 hours, while the proposed decomposition approach solved the problem to 0.7 % optimality gap within the same time.

6 Future Developments and Challenges

As indicated by the review presented in Sect. 4, planning and scheduling for industrial DSM is a relatively new research area, and there is significant need for further development. In the following, we highlight some remaining challenges and opportunities in this area:

- Power contracts can involve very complex price structures, which cannot be considered in the current models. Also, existing DSM programs are constantly revised and new ones are created. Hence, a stronger focus on the modeling of market mechanisms and the energy management side in general is needed.
- Many DSM problems, including the design of new processes, require long-term planning, which in turn results in multiscale problems as described in Sect. 4.3. There is much potential in the development of models for such problems.
- Synergies are expected from collaboration between multiple DSM participants, which could be a network of multiple plants that share customers but operate in different electricity markets. It could also be a joint effort of two different companies that operate interrelated power-intensive processes; a typical example is the case of a steel plant and an air separation plant that supplies oxygen to the steel plant.

- Following the same line of thought, it is also clear that collaboration between utility and consumer or, to take it a step further, the grid-wide optimization involving all participants can result in significant benefits for everyone. Here, however, besides the size and complexity of the problem, one also has to consider conflicting objectives among all decision makers.
- Decision-making under uncertainty still remains an open area of research in DSM. The accurate and computationally efficient modeling of different kinds of uncertainties is a major challenge. Here, the use of data will help in creating the models as well as in convincingly demonstrating the value of stochastic optimization.
- In general, it is crucial to develop more efficient solution methods, especially for solving multiscale and stochastic optimization models.

7 Concluding Remarks

This work provides an introduction to DSM with an emphasis on the perspective of large industrial electricity consumers, and a state-of-the-art review of existing works on planning and scheduling for industrial DSM. We have identified four major challenges in this emerging research area, and have demonstrated the capabilities of current models with two real-world case studies. We hope that this review will stimulate further research as we still face major challenges, two of which are multiscale (temporal and spatial) optimization and decision-making under uncertainty.

Acknowledgments The authors gratefully acknowledge the financial support from the National Science Foundation under Grant No. 1159443 and from Praxair. The authors would also like to thank Dr. Jose M. Pinto and Dr. Arul Sundaramoorthy from Praxair for the successful collaboration on many projects related to industrial DSM and the real-world data that they have provided for the case studies.

References

Albadi MH, El-Saadany EF (2008) A summary of demand response in electricity markets. Electric Power Systems Research 78(11):1989–1996

Artigues C, Lopez P, Hat A (2013) The energy scheduling problem: Industrial case-study and constraint propagation techniques. International Journal of Production Economics 143(1):13–23

Ashok S (2006) Peak-load management in steel plants. Applied Energy 83(5):413–424

Ashok S, Banerjee R (2001) An optimization mode for industrial load management. IEEE Transactions on Power Systems 16(4):879–884

Babu CA, Ashok S (2008) Peak Load Management in Electrolytic Process Industries. IEEE Transactions on Power Systems 23(2):399–405

Balas E (1985) Disjunctive Programming and a Hierarchy of Relaxations for Discrete Optimization Problems. SIAM Journal on Algebraic Discrete Methods 6(3):466–486

Ben-Tal A, Goryashko A, Guslitzer E, Nemirovski A (2004) Adjustable robust solutions of uncertain linear programs. Mathematical Programming 99(2):351–376

Ben-Tal A, El Ghaoui L, Nemirovski A (2009) Robust Optimization. Princeton University Press, New Jersey

Birge JR, Louveaux F (2011) Introduction to Stochastic Programming, 2nd edn. Springer Science + Business Media

Birge JR, Louveaux FV (1988) A multicut algorithm for two-stage stochastic linear programs. European Journal of Operational Research 34(3):384–392

Castro PM, Harjunkoski I, Grossmann IE (2009) New Continuous-Time Scheduling Formulation for Continuous Plants under Variable Electricity Cost. Industrial & Engineering Chemistry Research 48(14):6701–6714

Castro PM, Harjunkoski I, Grossmann IE (2011) Optimal scheduling of continuous plants with energy constraints. Computers & Chemical Engineering 35(2):372–387

Castro PM, Sun L, Harjunkoski I (2013) Resource-Task Network Formulations for Industrial Demand Side Management of a Steel Plant. Industrial & Engineering Chemistry Research 52:13,046–13,058

Castro PM, Grossmann IE, Veldhuizen P, Esplin D (2014) Optimal Maintenance Scheduling of a Gas Engine Power Plant using Generalized Disjunctive Programming. AIChE Journal 60(6):2083–2097

Charles River Associates (2005) Primer on Demand-Side Management. Tech. Rep. February, The World Bank

Conejo AJ, Nogales FJ, Arroyo JM (2002) Price-Taker Bidding Strategy Under Price Uncertainty. IEEE Transactions on Power Systems 17(4):1081–1088

Daryanian B, Bohn RE, Tabors RD (1989) Optimal Demand-Side Response to Electricity Spot Prices for Storage-Type Customers. IEEE Transactions on Power Systems 4(3):897–903

Ding YM, Hong SH, Li XH (2014) A Demand Response Energy Management Scheme for Industrial Facilities in Smart Grid. IEEE Trans 10(4):2257–2269

DOE (2006) Benefits of demand response in electricity markets and recommendations for achieving them. Tech. rep., U.S. Department of Energy

DOE (2013) Smart Grid Investment Grant Program—Progress Report II. Tech. rep., U.S. Department of Energy

Dupacova J, Gröwe-Kuska N, Römisch W (2003) Scenario reduction in stochastic programming: An approach using probability metrics. Mathematical Programming Ser. A 95:493–511

Duran MA, Grossmann IE (1986) An Outer-Approximation Algorithm for a Class of Mixed-Integer Nonlinear Programs. Mathematical Programming 36:307–339

Düzgün R, Thiele A (2015) Robust Optimization with Multiple Ranges: Theory and Application to Pharmaceutical Project Selection. In: Proceedings of the 14th INFORMS Computing Society Conference, Richmond, pp 103–118

EIA (2012) Manufacturing Energy Consumption Survey Data Table 11.1—Electricity: Components of Net Demand, 2010. Tech. rep., U.S. Energy Information Administration

Everett G, Philpott A (2002) Pulp Mill Electricity Demand Management. In: Proceedings of the 37th Annual Conference of the Operational Research Society of New Zealand, Auckland

Fang K, Uhan N, Zhao F, Sutherland JW (2011) A new approach to scheduling in manufacturing for power consumption and carbon footprint reduction. Journal of Manufacturing Systems 30(4):234–240

FERC (2010) National Action Plan on Demand Response. Tech. rep., Federal Energy Regulatory Commission

Floudas CA, Lin X (2004) Continuous-time versus discrete-time approaches for scheduling of chemical processes: a review. Computers & Chemical Engineering 28(11):2109–2129

GAMS Development Corporation (2015) GAMS version 24.4.1

Gellings CW (1985) The concept of demand-side management for electric utilities. Proceedings of the IEEE 73(10):1468–1470

Gellings CW, Wikler G, Ghosh D (2006) Assessment of U.S. Electric End-Use Energy Efficiency Potential. Electricity Journal 19(9):55–69

Geoffrion A (1972) Generalized Benders Decomposition. Journal of Optimization Theory and Applications 10(4):237–260

Graves SC (1981) A Review of Production Scheduling. Operations Research 29(4):646–675

Grossmann IE, Trespalacios F (2013) Systematic Modeling of Discrete-Continuous Optimization Models through Generalized Disjunctive Programming. AIChE Journal 59(9):3276–3295

Hadera H, Harjunkoski I, Labrik R, Sand G, Engell S (2014) An Improved Energy-awareness Formulation for General Precedence Continuous-time Scheduling Models. Submitted for publication

Hadera H, Harjunkoski I, Sand G, Grossmann IE, Engell S (2015 a) Optimization of Steel Production Scheduling with Complex Time-Sensitive Electricity Cost. Computers & Chemical Engineering 76:117–136

Hadera H, Wide P, Harjunkoski I, Sand G, Engell S (2015 b) A Mean Value Cross Decomposition Strategy for Demand-side Management of a Pulping Process. In: 12th International Symposium on Process Systems Engineering and 25th European Symposium on Computer Aided Process Engineering, pp 1931–1936

Haït A, Artigues C (2011) On electrical load tracking scheduling for a steel plant. Computers and Chemical Engineering 35(12):3044–3047

Harjunkoski I, Maravelias CT, Bongers P, Castro PM, Engell S, Grossmann IE, Hooker J, Méndez C, Sand G, Wassick J (2014) Scope for industrial applications of production scheduling models and solution methods. Computers & Chemical Engineering 62:161–193

Ierapetritou MG, Wu D, Vin J, Sweeney P, Chigirinskiy M (2002) Cost Minimization in an Energy-Intensive Plant Using Mathematical Programming Approaches. Industrial & Engineering Chemistry Research 41(21):5262–5277

Karwan MH, Keblis MF (2007) Operations planning with real time pricing of a primary input. Computers & Operations Research 34(3):848–867

Kirschen DS (2003) Demand-Side View of Electricity Markets. IEEE Transactions on Power Systems 18(2):520–527

Kondili E, Pantelides CC, Sargent RWH (1993) A General Algorithm for Short-Term Scheduling of Batch Operations—I. MILP Formulation. Computers & Chemical Engineering 17(2):211–227

Levy R (2006) A Vision of Demand Response—2016. The Electricity Journal 19(8):12–23

Li Z, Ierapetritou M (2008) Process scheduling under uncertainty: Review and challenges. Computers & Chemical Engineering 32(4–5):715–727

Maravelias CT (2012) General Framework and Modeling Approach Classification for Chemical Production Scheduling. AIChE Journal 58(6):1812–1828

Maravelias CT, Sung C (2009) Integration of production planning and scheduling: Overview, challenges and opportunities. Computers & Chemical Engineering 33(12):1919–1930

Méndez CA, Cerdá J, Grossmann IE, Harjunkoski I, Fahl M (2006) State-of-the-art review of optimization methods for short-term scheduling of batch processes. Computers & Chemical Engineering 30(6–7):913–946

Merkert L, Harjunkoski I, Isaksson A, Säynevirta S, Saarela A, Sand G (2014) Scheduling and energy —Industrial challenges and opportunities. Computers & Chemical Engineering 72:183–198

Mitra S, Grossmann IE, Pinto JM, Arora N (2012a) Optimal production planning under time-sensitive electricity prices for continuous power-intensive processes. Computers & Chemical Engineering 38:171–184

Mitra S, Grossmann IE, Pinto JM, Arora N (2012b) Robust scheduling under time-sensitive electricity prices for continuous power-intensive processes. In: Proceedings of the Foundations of Computer-Aided Process Operations 2012

Mitra S, Sun L, Grossmann IE (2013) Optimal scheduling of industrial combined heat and power plants under time-sensitive electricity prices. Energy 54:194–211

Mitra S, Pinto JM, Grossmann IE (2014) Optimal multi-scale capacity planning for power-intensive continuous processes under time-sensitive electricity prices and demand uncertainty. Part I: Modeling. Computers & Chemical Engineering 65:89–101

Mohsenian-Rad AH, Wong VWS, Jatskevich J, Schober R, Leon-Garcia A (2010) Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid. IEEE Transactions on Smart Grid 1(3):320–331

Motegi N, Piette MA, Watson DS, Kiliccote S, Xu P (2007) Introduction to Commercial Building Control Strategies and Techniques for Demand Response. Tech. rep., Lawrence Berkeley National Laboratory

Nolde K, Morari M (2010) Electrical load tracking scheduling of a steel plant. Computers and Chemical Engineering 34(11):1899–1903

Pantelides CC (1994) Unified Frameworks for Optimal Process Planning and Scheduling. In: Foundations of computer-aided process operations, New York, pp 253–274

Paulus M, Borggrefe F (2011) The potential of demand-side management in energy-intensive industries for electricity markets in Germany. Applied Energy 88(2):432–441

Reklaitis GV (1982) Review of scheduling of process operations. AIChE Symposium Series 78 (214):119–133

Rockafellar RT, Uryasev S (2000) Optimization of Conditional Value-at-Risk. Journal of risk 2:21–42

Samad T, Kiliccote S (2012) Smart grid technologies and applications for the industrial sector. Computers & Chemical Engineering 47:76–84

Shrouf F, Ordieres-Meré J, Garca-Sánchez A, Ortega-Mier M (2014) Optimizing the production scheduling of a single machine to minimize total energy consumption costs. Journal of Cleaner Production 67:197–207

Siano P (2014) Demand response and smart grids—A survey. Renewable and Sustainable Energy Reviews 30:461–478

Strbac G (2008) Demand side management: Benefits and challenges. Energy Policy 36(12): 4419–4426

Sundaramoorthy A, Maravelias CT (2011) Computational Study of Network-Based Mixed-Integer Programming Approaches for Chemical Production Scheduling. Industrial & Engineering Chemistry Research 50:5023–5040

Tan M, Duan B, Su Y, He F (2015) Optimal hot rolling production scheduling for economic load dispatch under time-of-use electricity pricing. Submitted for publication

Tan YY, Huang YL, Liu SX (2013) Two-stage mathematical programming approach for steelmaking process scheduling under variable electricity price. Journal of Iron and Steel Research International 20(7):1–8

Todd D, Helms B, Caufield M, Starke M, Kirby B, Kueck J (2009) Providing Reliability Services through Demand Response: A Preliminary Evaluation of the Demand Response Capabilities of Alcoa Inc. Tech. rep., Alcoa

Verderame PM, Elia JA, Li J, Floudas CA (2010) Planning and Scheduling under Uncertainty: A Review Across Multiple Sectors. Industrial & Engineering Chemistry Research 49(9): 3993–4017

Vujanic R, Mariéthos S, Goulart P, Morari M (2012) Robust Integer Optimization and Scheduling Problems for Large Electricity Consumers. In: Proceedings of the 2012 American Control Conference, pp 3108–3113

Walawalkar R, Fernands S, Thakur N, Chevva KR (2010) Evolution and current status of demand response (DR) in electricity markets: Insights from PJM and NYISO. Energy 35(4):1553–1560

Wang X, Tong C, Palazoglu A, El-farra NH (2014) Energy Management for the Chlor-Alkali Process with Hybrid Renewable Energy Generation using Receding Horizon Optimization. In: Proceedings of the 53rd IEEE Conference on Decision and Control, Los Angeles, pp 4838–4843

Wang Z, Gao F, Zhai Q, Guan X, Liu K, Zhou D (2012) An integrated optimization model for generation and batch production load scheduling in energy intensive enterprise. In: Proceedings of the IEEE Power and Energy Society General Meeting

Worrell E, Price L, Neelis M, Galitsky C, Nan Z (2008) World Best Practice Energy Intensity Values for Selected Industrial Sectors. Tech. rep., Ernest Orlando Lawrence Berkeley National Laboratory

Yusta JM, Torres F, Khodr HM (2010) Optimal methodology for a machining process scheduling in spot electricity markets. Energy Conversion and Management 51(12):2647–2654

Zareipour H, Cañizares CA, Bhattacharya K (2010) Economic Impact of Electricity Market Price Forecasting Errors: A Demand-Side Analysis. IEEE Transactions on Power Systems 25(1):254–262

Zhang Q, Bremen AM, Grossmann IE, Sundaramoorthy A, Pinto JM (2015a) Long-term electricity procurement for large industrial consumers under uncertainty. Working paper

Zhang Q, Cremer JL, Grossmann IE, Sundaramoorthy A, Pinto JM (2015b) Risk-based integrated production scheduling and electricity procurement for continuous power-intensive processes. Computers & Chemical Engineering

Zhang Q, Grossmann IE, Heuberger CF, Sundaramoorthy A, Pinto JM (2015c) Air Separation with Cryogenic Energy Storage: Optimal Scheduling Considering Electric Energy and Reserve Markets. AIChE Journal 61(5):1547–1558

Zhang Q, Grossmann IE, Sundaramoorthy A, Pinto JM (2015d) Data-driven construction of Convex Region Surrogate models. Optimization and Engineering

Zhang Q, Morari MF, Grossmann IE, Sundaramoorthy A, Pinto JM (2015e) An adjustable robust optimization approach to provision of interruptible load by continuous industrial processes. Computers & Chemical Engineering

Zhang Q, Sundaramoorthy A, Grossmann IE, Pinto JM (2016) A discrete-time scheduling model for continuous power-intensive process networks with various power contracts. Computers & Chemical Engineering 84:382–393

Zhang X, Hug G (2014) Optimal Regulation Provision by Aluminum Smelters. In: Proceedings of the IEEE Power and Energy Society General Meeting, National Harbor

Zhang X, Hug G (2015) Bidding Strategy in Energy and Spinning Reserve Markets for Aluminum Smelters' Demand Response. In: Proceedings of the IEEE PES Conference on Innovative Smart Grid Technologies, Washington DC

Industrial Tools and Needs

Chemicals, Stainless Steel, Pulp and Paper

Iiro Harjunkoski and Hubert Hadera

Abstract The topic of demand-side management has grown within one decade from a "buzzword" to a vast academic research field where the multitude of complexities is acknowledged by a large number of potential methods to tackle the problem (see previous chapter). This is just one example of energy-related aspects which are getting attractive for heavy industries. The role of energy has become more important than ever—both because of environmental and cost reasons. Therefore, the time is right to bring the best theories into practice. The industry is both open toward, as well as, looking for solutions that can improve the energy efficiency, sustainability, and productivity simultaneously while reducing the total costs of production. In parallel, the developments within information technology, highlighting the topic of integration and collaboration, support applications that merge earlier separated tasks and aim to maximize the synergies between various industrial solutions. In this chapter, the main focus is on existing tools and how to meet the needs of energy-intensive processes. These will be discussed along real case studies.

1 Background

The main driver for the industry to change its energy-related activities comes from four main factors:

- Fast changing legislation (e.g., national targets for energy efficiency);
- Increasing role of clean technologies, e.g., share of renewable energy sources (especially wind, solar);

I. Harjunkoski (✉)
ABB Corporate Research, Wallstadter Str. 59, 68526 Ladenburg, Germany
e-mail: iiro.harjunkoski@de.abb.com

H. Hadera
BASF SE, Carl-Bosch-Str. 38, 67056 Ludwigshafen, Germany
e-mail: hubert.hadera@basf.com

M. Martín (ed.), *Alternative Energy Sources and Technologies*,
DOI 10.1007/978-3-319-28752-2_15

- Deregulation of electricity markets;
- Changing cost structures and dynamics due to the above.

As all of this has been already discussed and motivated earlier in this book, here the focus is on the realities that industry is facing today. Ultimately, one of the most important energy-related factors is security of energy supply and the fact that energy is a significant cost driver for production. Energy has become more expensive and is thus not anymore something that can be calculated on a fixed basis. The prices can vary overnight significantly with the consequence that an energy-intensive process can turn unprofitable. Also, the pressure to contribute to the "clean" energy is much higher, both economically as well as politically. Many governmental programs such as the US NSF program "Energy for Sustainability" and EU frame program "Horizon 2020" or Energy Efficiency Directive (European Union 2012) highlight the increased use of reusable energy forms and promote the development of new technologies whose final goal is to have carbon-neutral production. Industry is at the same time enforced to actively take part of this through financial sanctions and a stronger public opinion on sustainable energy generation. Some companies commit voluntarily to the use of clean technologies by placing the contribution to environmental sustainability into their vision and mission.

At the same time as the political climate is pushing for new approaches, it also creates uncertainties due to various national agendas and changing short- and mid-term strategies on how to move forward. In Europe, for instance, the decision of the German government to close down the nuclear power plants challenges both the power grid and end users; instead of a stable base supply, fluctuations of energy availability will increase dramatically and the grid of today cannot handle this without governmental investments and infrastructure projects spanning over more than a decade. Being able to forecast the exact energy availability and price using factors such as weather and demand forecasts are becoming more important for maintaining a stable energy supply. Here, a larger responsibility for stabilizing the power grid will be put especially on large electricity consumers who already today have to forecast their consumption. These consumers, especially, will need tools and solutions that maximize the flexibility of their processes and increase the transparency of the total energy needs at each point of time.

Production planning and scheduling have until now commonly been done in an isolated manner, often manually using some type of planning sheets or simple supporting tools. The generated plan has then been submitted to the energy department, whose task has been to find the most economical energy strategy based on a given plan, often leaning on experts' knowledge and experience or energy management systems. Due to this sequential procedure where one solution limits the other one, short-term production planning and energy management may seem to have partially competing goals. However, integrating both in a collaborative way enables the full benefits for the participation on energy markets and stabilization of the power grid. This is also in alignment with the enterprise-wide optimization (EWO) philosophy (Grossmann 2005) that combines various aspects of production. When considering simultaneously aspects from control, manufacturing operations

management and supply chain layers, the roles of existing systems are instrumental. An important aspect is that all participating solutions typically possess some level of optimization capability. If this is the case, a good coordination can avoid the risk of having competing optimization solutions resulting in no improvements.

2 Industrial Processes

Most electricity-intensive processes contain large rotating equipment, e.g., pumps and compressors or electricity-driven heating or cooling devices. Examples of such processes are pulp and paper manufacturing, steel or aluminum production, air separation, electrolysis processes and, to a larger extent, data centers, to name a few. Other processes, typically related to the chemical sector, require large amounts of heat in order to enable chemical reactions to take place. The structure of all of the above vary heavily based on the complexity of the process. Also the process flexibility, i.e., how fast a process can increase, decrease, or completely stop its energy consumption, plays an important role both in the complexity of managing it, as well as, the potential of developing feasible and financially attractive solutions. For a very simple process, probably no optimization is needed and an energy management solution can lead the decision-making. For some of the energy-intensive process industries, the raw material and energy cost can account for up to 60 % of the total production cost. For example in stainless steel manufacturing, where scrap metal is melted in Electric Arc Furnace (EAF), the electricity accounts for about 20–40 % of the total raw material cost (Ali et al. 2007). The daily electricity bill can climb up to hundreds of thousands EUR for such plants. That is why advanced decision-making supporting tools have the potential to bring significant overall cost reductions (see, e.g., Hadera and Harjunkoski 2013; Castro et al. 2013).

In Table 1, selected energy-intensive industrial loads are shown where tools for tackling energy-related aspects are needed. The previously mentioned steel process operates in a batch manner with typically around 80–100 MW intensive EAF stage (Hadera et al. 2015a). In paper making, often parallel pulp refiners require each 10–30 MW for producing the pulp, which is later processed on a paper machine with a typical electricity consumption of around 30–40 MW. The electrolysis processes of aluminum and chloralkali are substantial electricity consumers (15 MWh/t for aluminum and 2.85 MWh/t for chloralkali electrolysis; Hadera 2015) with the special requirement of high process utilization in order to operate on efficient level. Another energy-intensive steel making process using Blast Furnace (BF) is operated for long time periods without interruptions and the entire downstream processing line has an electricity demand of similar to the EAF stage above. Large rotating

Table 1 Example energy-intensive industrial sectors and processes (based on Hadera 2015)

Industrial sector	Process	Energy-intensive stage or equipment
Metals	Stainless-steel making	Energy-intensive Electric Arc Furnace stage, batch process
Pulp and paper	Thermo-mechanical pulping (TMP) and paper making	Energy-intensive refiners and paper machines, storage of pulp
Chemicals	Chloralkali electrolysis	Energy-intensive electrolysis, but high capacity utilization
Metals	Aluminum electrolysis	Energy-intensive electrolysis, but high capacity utilization
Metals	Primary steel making	Energy-intensive Blast Furnace (BF) stage, but high capacity utilization
Cement	Cement milling	Energy-intensive rotary machines
Chemicals	Air separation	Energy-intensive compressors, energy cost is significant raw material cost
Oil and gas	Liquefied Natural Gas (LNG)	Energy-intensive refrigeration
Mining	Winder systems	Energy-intensive winder motors, intermediate storage of transported material
Mining	Belt conveyor systems	Energy-intensive motors for conveyers
Metals	Hot Rolling Mill (HRM)	Energy-intensive motors for the rolling machine, relatively quick batch process
Chemicals	Steam cracking	Energy used to heat up steam flow
Chemicals	Commodity chemicals production	Energy for continuous processes such as distillation and separation

machines are used in cement production for the crushing and grounding phases. The earlier mentioned air separation process for which the main raw material (air) comes for free runs energy-intensive oxygen and air compressors. In LNG plants converting the natural gas into liquid for storage and transportation, a typical process consumes around 5.5–6 kWh per kmol of LNG produced (Zargarzadeh et al. 2007) with about 40 % of operating cost due to energy in the refrigeration section (Hasan et al. 2009). The mining industry contains energy-intensive rock winder systems to transport rock, waste, and machinery between the ground surface and underground levels (Vosloo 2006). Also, belt conveyor systems with large electrical motors are another energy-intensive consumers. Another example process in steel industry is the hot rolling mill. Here, the steel slab thickness is reduced from, for instance, 200–5 mm by applying mechanical pressure on the product. The rolls used for pressing are driven by energy-intensive motors. The process is carried out in batch campaigns of several slabs, during which the rolling of one slab takes only a few minutes, followed by pauses. In contrast, the basic processes in chemical industry, such as steam cracking and production of bulk chemicals are continuous processes which constantly require large amounts of heat.

3 Value of a Better Energy Optimization

Increasing the energy efficiency is a straightforward way to reduce the specific energy consumption and, subsequently, the total energy costs. The goal is to reduce the overall energy consumption without changing the production volume of the site. Some typical actions to increase the energy efficiency are:

- Install more energy-efficient equipment, e.g., LEDs instead of normal lighting;
- Replace fixed speed pumps followed by a valve with pumps with variable speed drives;
- Insulate buildings better in order to decrease heat losses;
- Define better operating strategies leading to fast and easy consumption reductions, e.g., switch off equipment during unproductive times;
- Detect inefficient operations early, e.g., leakage in pipeline systems leading to inefficient pumping and material losses;
- Improve material efficiency by reducing process waste through better planning and control of the production processes.

Several other energy efficiency actions are also possible. However, for implementing energy efficiency actions at industrial sites the potential benefits need to be proven and verified upfront. One very common measure to track energy efficiency is energy consumption per ton of product delivered, but other measures might be more applicable. The actual measures for energy efficiency need to be chosen wisely in order to maximize the impact of the improvement actions.

An energy management system supports a systematic approach for measuring and tracking energy use and in establishing a clear system for planning and evaluation of the benefits of energy efficiency actions. Many companies are striving to improve their energy efficiency by, e.g., deploying the ISO 50001 standard. By 2014 more than 7000 sites have been certified. ISO 50001 is a standard providing the guidelines for energy management systems (ISO 2011). This standard introduces a continuous improvement mechanism; the plan–do–check–act cycle (Fig. 1), according to which each and every action needs not only be planned and implemented, but their effect should also be controlled in order to derive better follow-up actions. The first step to start the continuous improvement process is to install a monitoring system which measures the actual energy consumption. Based on the information gathered with this monitoring possible actions to reduce energy consumption can be evaluated. The system also emphasizes the role of higher management and its commitment to undertake energy-related improvement actions via establishing an energy policy and periodical management reviews. The expected benefit of a certification is around 5–15 % within 3 years. According to Scheihing (2014) a study of nine certified facilities reached annual savings between $87,000 and $984,000. In average, the facilities achieved around 10 % improvement in energy performance within 1.5 years after the certification.

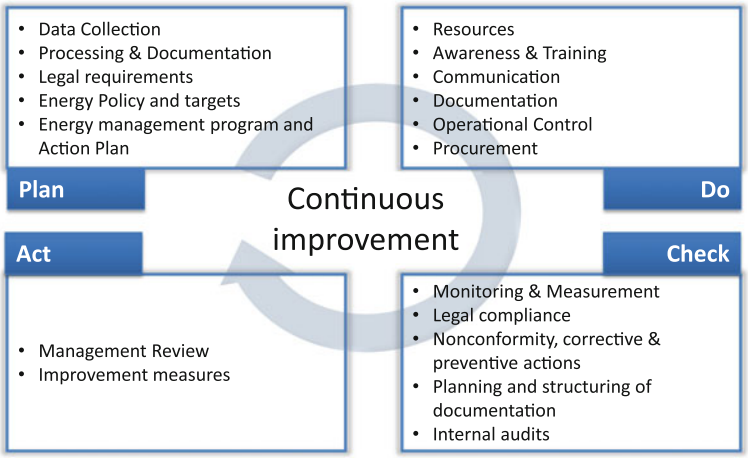


Fig. 1 The plan–do–check–act cycle for continuous improvement with regard to energy-related activities (based on Scheihing 2014)

3.1 Energy Optimization

The above part dealt with well-defined methodologies for improving the energy efficiency. This is often hardware or process equipment specific and therefore standardized processes can bring significant and sustainable improvements. When it comes to optimization of the energy use, which must be done on a systems level the complexity increases significantly due to differences in processes, multitude or equipment involved and dependency on the supporting system environment. In industries such as metals, pulp and paper, cement or air separation, energy is an important cost factor (Merkert et al. 2015). Reducing this cost is of high interest, especially since energy costs are rising due to increasing energy demand and changes in the generation mix and supply. In many countries reduction of CO_2 emissions is on the political agenda (European Union 2011) and recently it was announced that US will further enforce the cutting of future emissions. This is a large concern for many industries as, compared to less regulated countries, they may have a significant competitive drawback since the energy bill will also directly be reflected in product prices. Another serious concern is the diminishing baseload, which may lead to production disruption that may endanger the profitability or in some cases even the operability of a plant. For these reasons it is crucial to identify —and take advantage of the potential flexibility in the process. Hereby, the role of optimization may become significant for energy-intensive industries. For a plant with an annual electricity bill of 100 MEUR, already 1 % savings is equivalent to savings of one million EUR.

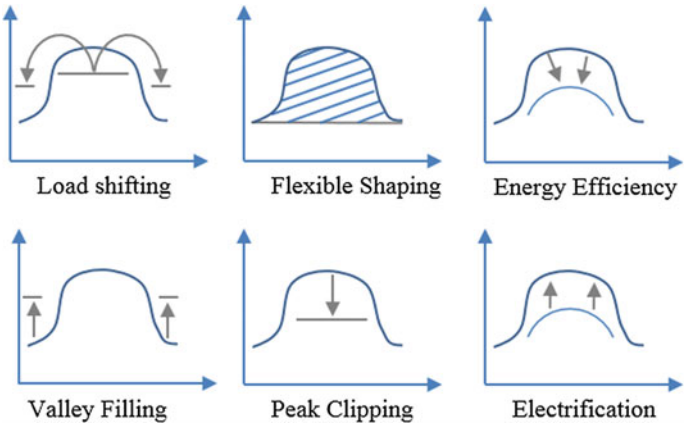


Fig. 2 Load shapes (based on Charles River Associates 2005)

In this context, the keyword for optimization is demand-side management, which has already been discussed in detail in the previous chapter. For individual processing plants this can often be achieved by shaping the load profiles of electricity consumed, as shown in Fig. 2.

Demand-side management is currently mainly applied to electricity but theoretically the same technology can be applied to other forms of energy. There are several ways to influence the demand-side load curve. Typical activity for reducing energy use is to install new more energy-efficient equipment in retrofitting projects in order to reduce overall energy use or introduce advanced maintenance planning in order to increase the operational capabilities (e.g., reduce machine wear). Energy efficiency can be done on equipment level, as well as, on process level as for example process intensification in chemical industry. Nevertheless, demand-side management is also about consuming or not consuming at the right times. The main options for demand-side management are (as shown in Fig. 2):

- Peak clipping or load shedding at peak times;
- Load shifting from peaks to valleys.

From the business point of view, demand-side management is of very high interest to utilities and grid operators as the supply of electricity has to equal consumption within the grid at any point in time and follow load changes, also including the highest consumption peaks for which backup power plants must be available. Maintaining rarely used plants is very expensive, so utilities are interested in reducing the height and occurrence of consumption peaks. To make it work, several business models exist with the most common ones being incentive and price-based programs, see previous chapter for details. Being closer to real-time pricing (electricity prices are defined very short-term) also involves much more frequent interaction between the grid provider, energy producer, and consumer (see Fig. 3). In some cases, the demand-side management can be handled by a control

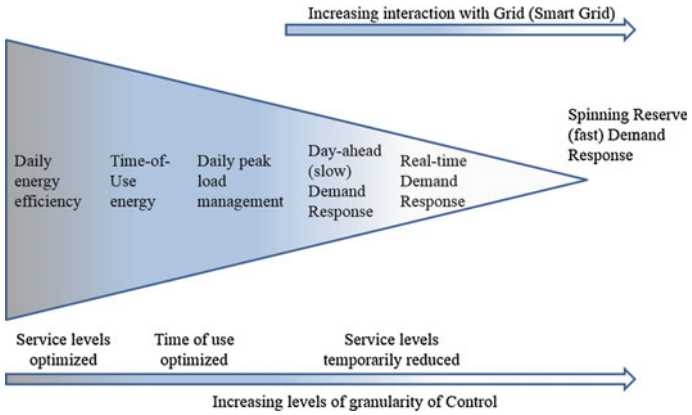


Fig. 3 Different time scales for demand response (based on Demand Response Research Center 2013)

solution, in more complex cases where the production must be replanned and the production scheduling function must be efficiently integrated.

It must, however, be kept in mind that production scheduling is only a part of the decision chain of the production management layer. This is not considered in most theoretical studies, the results of which consequently can only indicate the highest theoretical potential. Commonly, production targets are defined at Enterprise Resource Planning level (ERP—see Fig. 4), and executed in a Manufacturing

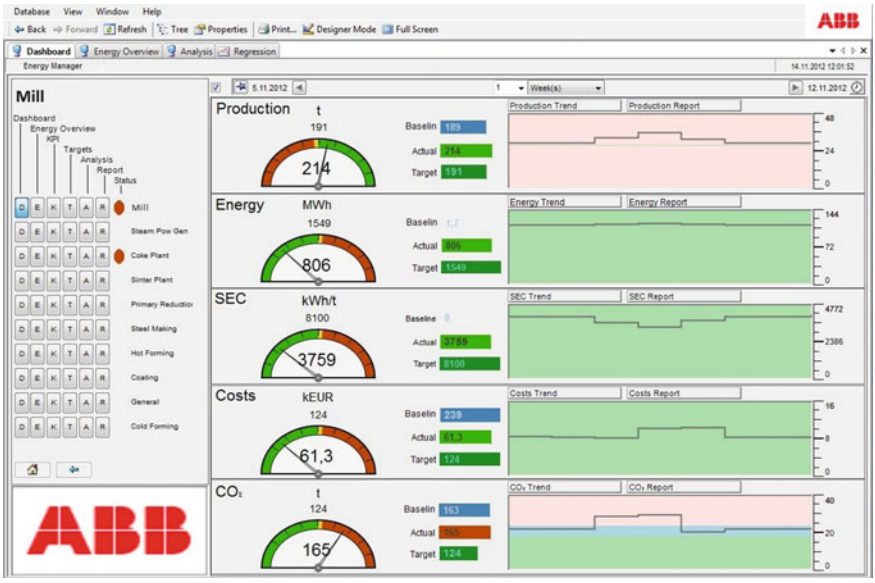


Fig. 4 Dashboard of a monitoring system showing current production trends

Execution System (MES) or Collaborative Production Management (CPM) system, responsible for following the production, taking into account expected and unexpected changes in production (tracking and rescheduling). Depending on the process, various methods can be applied for solving the resulting scheduling problems. Mathematical programming is often the best one as it can guarantee a global optimality, i.e., the best possible solution. However, in practice most optimization methods become eventually slow when the problem size grows and therefore a large number of heuristics and manually supported methods, e.g., through interactive Gantt charts are today more likely to tackle the daily industrial problems.

4 Industrial Tools and Solutions

Regardless of the complexity of the energy-intensive process and decision levels the following factors are common, cross-layer and extremely important for the success of integration of energy-related aspects into industrial operations:

- Measurement and metering;
- Monitoring and visualization;
- Analysis and reporting;
- Forecasting and optimization.

4.1 Measurement and Metering

The absolute prerequisite for anything that can be done in the first place is measurement. Without being able to at least reliably estimate the electricity consumption at each time no other improvements can be performed. Most of the times, e.g., a single pump does not know its electricity consumption (unless controlled by an intelligent drive) but depending on the electrification there are several industrial solutions for power monitoring also with load profiling. The cost of sensors and instrumentation is often the main limiting factor. However, more and more companies realize that smart metering of energy use by equipment and processes is an important enabler of big data technologies, thus cost reduction in long term, i.e., better overview of equipment performance results in more accurate estimation on future state and reducing downtime.

4.2 Monitoring and Visualization

Being able to measure the instantaneous energy use offers countless opportunities for monitoring. Monitoring solutions can also contain calculation engines,

analytics, and various visualizations that can be configured to support operator actions in an optimal way. Today, all types of monitoring solutions are per definition also available on portable devices anytime from anywhere, also allowing remote support and collaboration. An example of such a system can be seen in Fig. 4.

4.3 *Analysis and Reporting*

Using monitoring solutions can help an operator to take the decisions and see the consequences of these close to online. Analysis tools are necessary to support finding trends, identifying problems, and potential for improvement (see Fig. 5). The complexity of analysis tools spans from simple human-driven solutions to larger cloud-based calculation engines. In order to ensure longer term sustainability and be able to track development as well as meet regulatory requirements, reporting is equally important. Reports can span from hours to an entire lifetime of an asset and comprise a single device, entire plant, or enterprise. It is important that the reporting function also follows given standards.

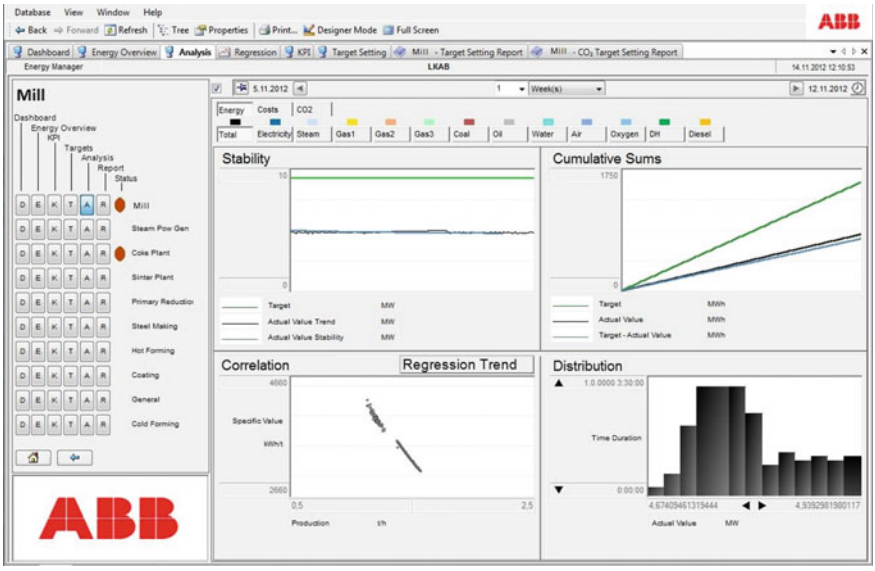


Fig. 5 Analytics screen with a visual summary of results

4.4 Forecasting and Optimization

Apart from the lower level functionalities, in the volatile markets one of the most valuable functions is forecasting. If a reliable forecast for the future electricity demand can be derived, it is possible take better decisions for energy purchase, which can result in significant savings. The challenge of forecasting is that it should ideally not only be able to take into account the future process state and energy consumption but also be able to estimate future energy prices. In the day-ahead and long-term markets the most important business decisions are based on forecasts and therefore poor forecasts can have expensive consequences. Most of the studied optimization solutions in the earlier chapter rely on historical data. This removes one of the big uncertainties from formulating and solving the problem. In reality, any optimization solution depends on forecasting solutions when estimating the problem parameters for the optimization. Figure 6 shows a solution for optimizing the continuous thermomechanical pulping (TMP) lines taking into account electricity price.

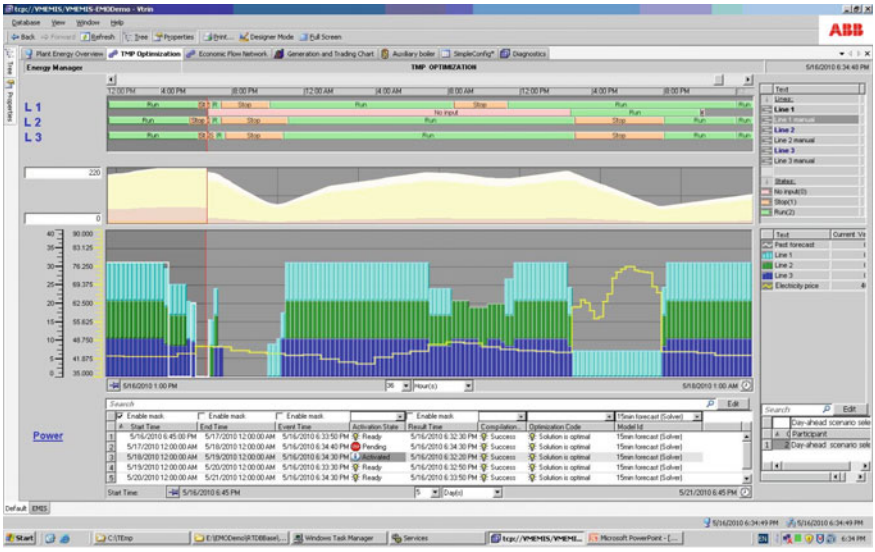


Fig. 6 Optimization of a thermo-mechanical pulp mill production w.r.t. energy

5 Systematical View of Putting All Together

The industrial landscape of tools and solutions for decision-making with regard to energy aspects can be explained using the universal automation pyramid (Fig. 7). The pyramid is also essential for the standard ISA-95 (ANSI/ISA 2005) for which the goal is to define common interfaces between enterprise and control systems which can be applied to any industry and any process.

The above decision model capture the information flow from complex business layer including customers, production demands, and companies external environment (e.g., legislation) up to the production level itself saturated with large amounts of data from sensors, actuators, and devices. Some typical tools and solution scope examples for these levels are:

- Enterprise resource planning (e.g., SAP ERP, SAP Supply Network Planning, ABB Energy Manager, BASF Verbund Simulator)—example activities are energy management on corporate level—goals and targets, security of energy supply (planning and purchase), supply chain networks planning (long-term) in response to seasonal energy availability;
- Manufacturing execution systems (e.g., Siemens Simatic IT, ABB CPM4Metals)—example activities are detailed production scheduling (short-term) considering volatile energy availability (e.g., prices), maintenance planning to increase asset efficiency;

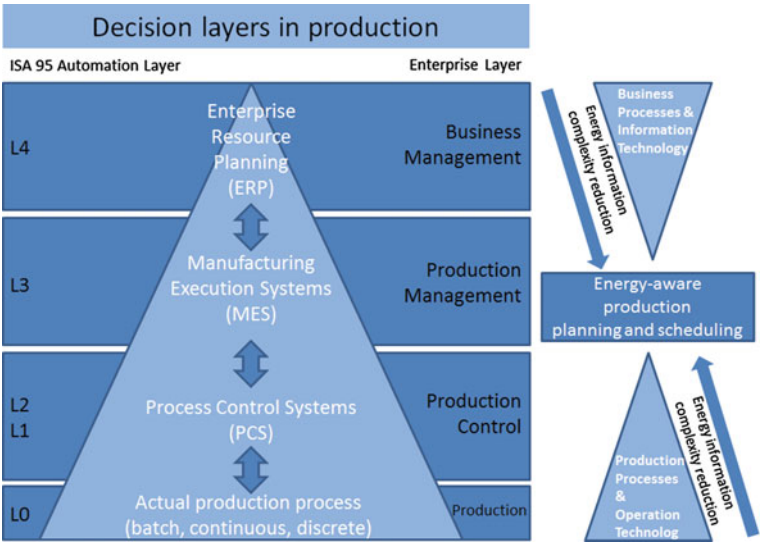


Fig. 7 The automation pyramid (ANSI/ISA 2005) and information overflow reduction layer

- Process control systems (Siemens PCS7, ABB 800xA)—example activities are Optimal process control (Advanced Process Control) for energy-efficient processes, smart sensors and energy metering, frequency and voltage control;
- Actual production process—installation of more energy-efficient equipment, process design, retrofitting.

At the heart of the ISA-95 model there is an information complexity reduction need within the MES layer. The information overflow from the ERP (thousands of suppliers and product orders) and control levels (thousands of measured process parameters) needs to be handled and translated into useful feedback, which can drive the companies near-future activities (Fig. 7). Thus, in this respect one of the main functions of the MES layer is production scheduling and dispatching. The crucial role of energy-aware production scheduling is within the industrial interest and is gaining on attention in recent years also at academia. More detailed real-life examples from this decision layer follow.

In order to make it easier for industry to integrate energy awareness across existing system, the way how energy is combined with production planning matters greatly. As stated earlier, the traditional approach in industry was to schedule the production first without taking energy costs into consideration and then based on a fixed schedule derive cost-optimal strategy for the energy demand satisfaction (left most approach in Fig. 8). This is not system-wide optimal solution because there could be a production schedule not optimal with regard to the production goals, but significantly bring savings on the energy cost side, such that overall system-wide cost is reduced. A common literature approach to tackle this problem is to come up with a single monolithic solution that would embed all the production scheduling constraints with all energy-cost related constraints such that in many works listed in the previous chapter. However, this is often problematic for industrial implementation due to the necessity to install and connect a completely new and complex

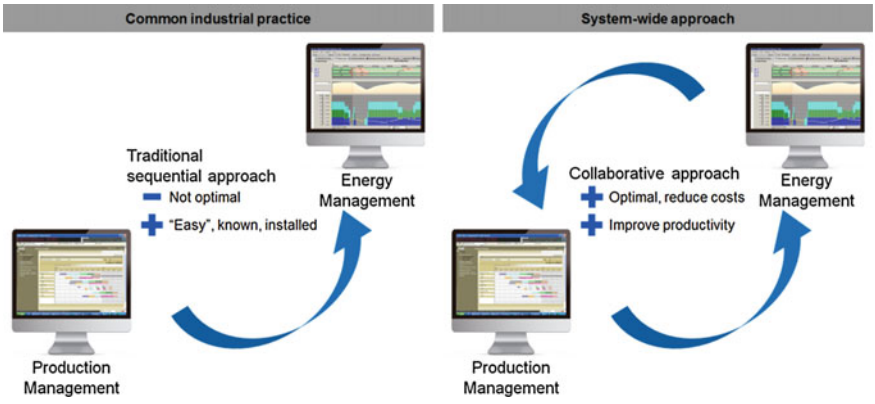


Fig. 8 Common industrial reality is often not matching the system-wide optimal approach for production scheduling with energy considerations (based on Harjunkoski et al. 2015; Hadera et al. 2015b)

solution, whereas some components of it might be potentially existing (for example, production scheduler might already be in use). Therefore, one way to overcome this problem is to use an intelligent information exchange between production scheduling systems and energy cost optimization systems. In a study by Hadera et al. (2015b), theoretical properties of advanced mathematical problem decompositions are studied to conclude that special theoretical features (so-called marginal cost) of mathematical optimization problem can be useful for solving industrially relevant functional separation of production scheduling and energy optimization.

6 Industrial Case Studies

In order to put the earlier methods into an industrial concept, a number of case studies will be discussed. Unlike in the previous chapter, here the focus is not on the theory or mathematics. Instead, the scope is limited to explaining the process together with a short problem statement and discussion on the practical requirements, opportunities, benefits, and limitations.

6.1 Integration of Chemicals Production—BASF's "Verbund" Concept

Efficient integration of material stream flows is in the heart of the production "Verbund" concept of the leading chemical company BASF. At six sites worldwide high level of value chains integration spans from basic chemicals to high value-added products (e.g., crop protection products) and consumer products. The integration can be done on the following levels (BASF 2015a):

- Logistics (shared network of pipelines);
- Infrastructure (shared on-site facilities);
- Energy (linked production and energy demands).

By-products of one plant can be used as raw material in others. Heat available from a production process, which potentially would be released to environment, can be utilized as energy input at other production plant. For example, at a second largest BASF Verbund site—in Antwerp (Belgium)—in 2009 there was about 1 000 ton/h steam of total heat consuming production processes, while heat producing chemical processes accounted for 775 ton/h steam. Therefore, it makes sense to utilize the interconnected needs (see chapter "Improving Energy Efficiency in Batch Plants Through Direct Heat Integration" for heat integration). Even greater potential lies in the Verbund site in Ludwigshafen (Germany), which is more than twice as big in terms of number of production plants—it is one of the world's largest industrial complex. Here, a real case study role model is acrylic acid production plant. In the plant propylene input material is supplied using pipeline from

the steam cracker. The chemical reaction of propylene to acrylic acid results in heat generation, which can be captured into steam. Many of other plants at the site use it, therefore the steam is fed to the stem pipeline network. Two acrylic acid plants in Ludwigshafen are able to cover around 10 % of the steam requirements of the entire Verbund site (BASF 2015a). The use case is shown in Fig. 9, together with another example of acetylene production value chain interconnections. For BASF, in 2014 around 45 % of entire steam supply–demand was covered by waste heat (BASF 2015b).

Large production complexes require large and efficient energy supply system. One profitable way of supplying heat and electricity for the chemical production processes is to use combined heat and power (CHP) cycles. A real case study (Fig. 10) comes from another Verbund site of BASF, Freeport (USA). The idea here is commonly known for the scientific community, however still not widely used in industry. It is more energy efficient to produce electricity and heat in a combined heat and power plant using, for example, a natural gas combustion turbine rather than produce electricity with conventional fuel separately from heat production. In this case around 26 % of total energy efficiency increase is estimated (Geiger 2011). For BASF worldwide by using CHP 70 % of the total electricity demand is satisfied, therefore around 13 million MWh of fossil fuels in 2014 alone was saved (BASF 2015c) compared to conventional systems.

When the industrial ecological system is closely interlinked energy use and cost reduction are achieved. For Ludwigshafen alone avoidance of 7 million metric tons of freight per year is expected (around 280 000 fewer truckloads). It is estimated that the "Energy Verbund" concept brings more than 300 MEUR of savings annually worldwide (BASF 2015a) and about 200 MEUR in Ludwigshafen alone (BASF 2015c). This corresponds to around 30 % of all cost savings of Verbund concept. In addition, only in 2014 the energy savings of around 18 million MWh was achieved, which corresponds to annual reduction of 3.6 million metric tons of carbon emissions (BASF 2015a). At the largest Verbund site more than 50 % of the steam consumption is covered by use of excess process heat and of energetic use of by-products from production. The concept allows to increase the security of energy supply and increase degree of flexibility for energy planning along value chains. Another aspect is increased cross-functional knowledge exchange with regard to energy topics.

6.2 Stainless-Steel Batch Production

Stainless-steel production is one of the most energy-intensive industrial processes. The production starts with the scrap melting phase in an electric arc furnace (EAF) to create a batch of molten heat. The process of melting is carried out by passing large amounts of electricity through electrodes in order to form high-temperature electric arc (up to 3500 °C) that melts scrap metal and additives. In the next stage, the heat is transported to the argon oxygen decarburization

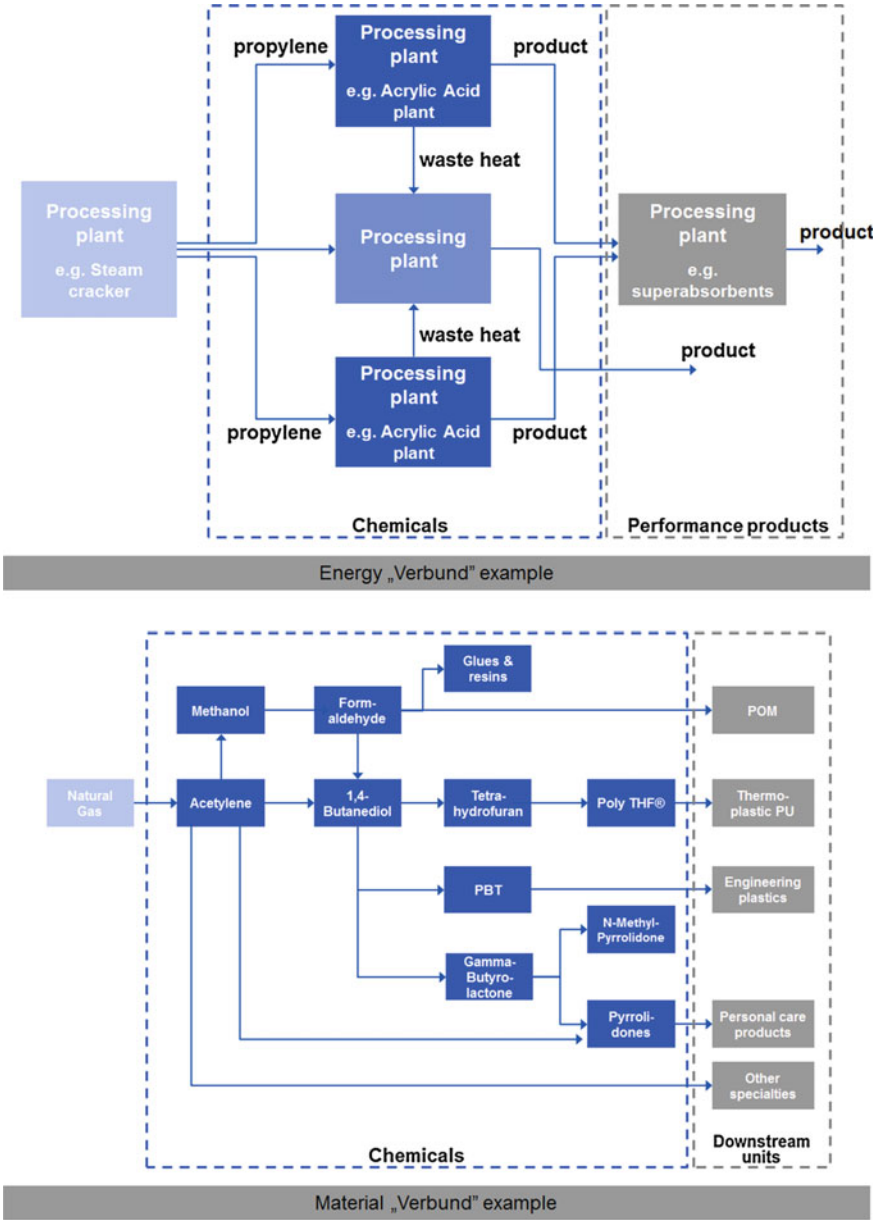


Fig. 9 Integration of energy and material flows in chemicals production—Verbund concept examples of BASF (*Source* BASF 2015a; Engel and Smith 2014)

(AOD), where the carbon content is reduced by injecting a mixture of argon–oxygen gas. After this stage the intermediate product (heat) needs to be heated up in Ladle Furnace (LF) before the final stage of continuous casting (CC), where

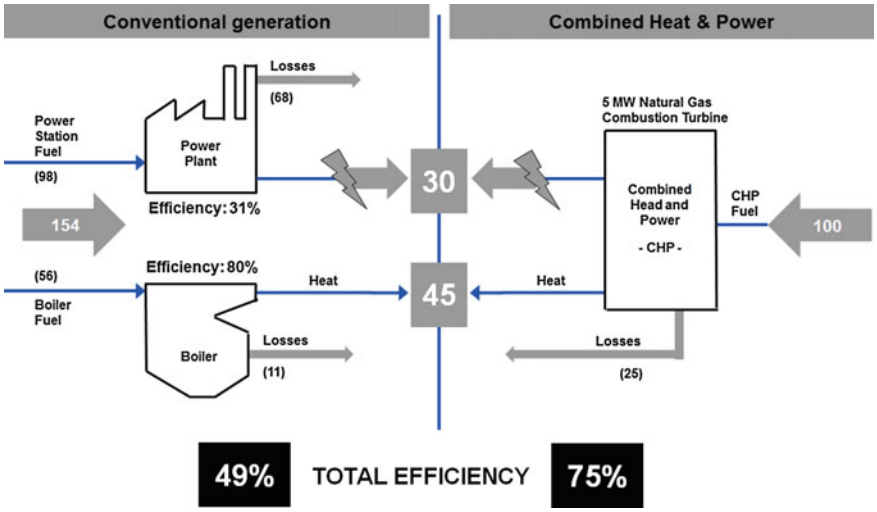


Fig. 10 Increased primary energy efficiency via combined heat and power instead of conventional separate generation of heat and electricity (*Source* Geiger 2011)

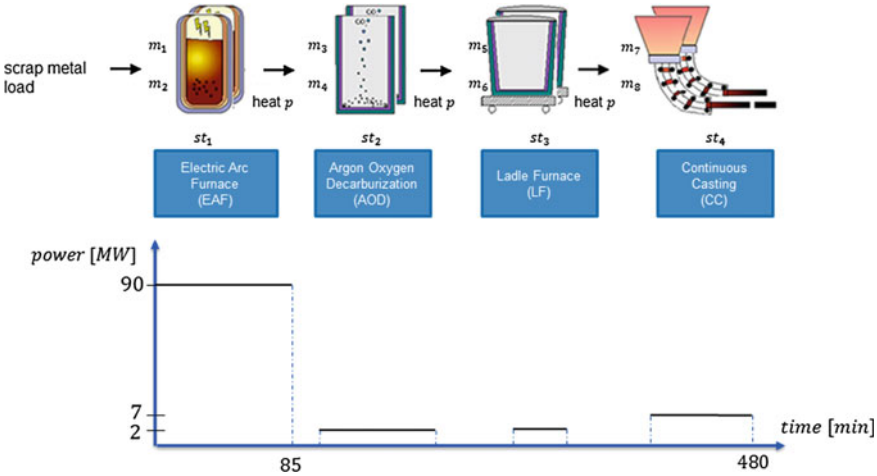


Fig. 11 The stainless-steel production process with simplified power use profile across production stages

smelted steel is formed into solid steel slabs. A process overview is shown in Fig. 11. The electricity need is shown in the below graph for all of the four major processing steps.

Since the production process has a significant electricity-consuming stage which operates in batch mode, it makes sense to intelligently plan its processing times in connection with time-volatile energy prices. Such coordination might have

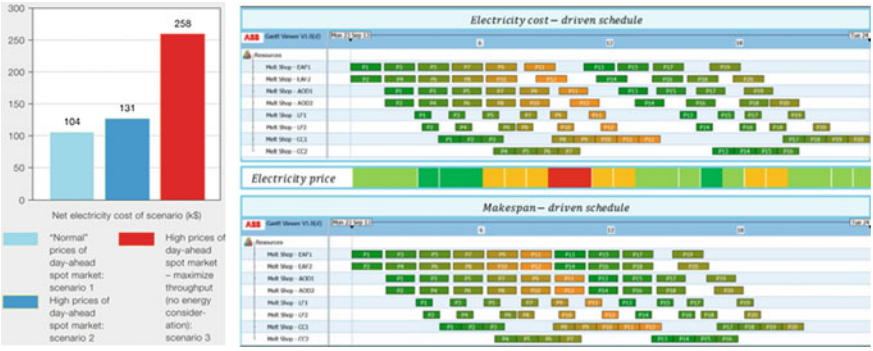


Fig. 12 Real-life example impact of different volatile pricing scenarios on total electricity expenditure (*left* Scenarios 1–3) and different scheduling policies on production schedule (*right* Scenarios 2 and 3) (*Source* Harjunkoski et al. 2015)

significant impact on the actual savings. Some published industrial reports indicate savings at a level of 5 % of total electricity cost (Harjunkoski et al. 2015), which translates into even few hundred thousands of dollars per month. Figure 12 shows real-life example of different energy tariffs and how scheduling policies might impact total electricity cost (Harjunkoski et al. 2015). For a typical 24 h scheduling horizon of a melt shop process with "normal" electricity prices in the volatile markets the net electricity cost is around $105 k. However, when in case there is shortage of energy availability the cost goes up by $27 k. This is just for 1 day of production. Then, for the same day, if one does not perform energy-aware scheduling, but realizes the traditional throughput maximization (or makespan minimization) strategy, the net cost of electricity almost doubles. The impact of the production task timing of two different scheduling strategies, both with high prices in the volatile markets, is shown rightmost in Fig. 12.

Practical implementation aspects of the stainless steel energy-aware scheduling solution are discussed by Gajic et al. (2015). Here, it is pointed out that one important aspect of real-life implementations is the fact that very often complex production plant topologies are difficult to solve in one single mathematical model. As many other production processes, also steel production scheduling usually requires employment of decomposition methods. After inclusion of energy costs optimization aspects the models become even more difficult to solve. A generic decomposition strategy that has been used in many other studies is utilized for energy-aware steel scheduling in continuous-time general precedence approach in Hadera et al. (2015a). In this study, in order to find good-quality solutions two separate problems are solved: an upper level approximation problem and a lower level full scope problem (Fig. 13.). The approximation considers only two production stages, the most energy-intensive one (EAF) and the most constrained one (CC due to semi-continuous production mode). The upper level problem gives an important guidance of how should the two stages be sequenced and assigned to machines. For these difficult decisions with the original full space problem treated

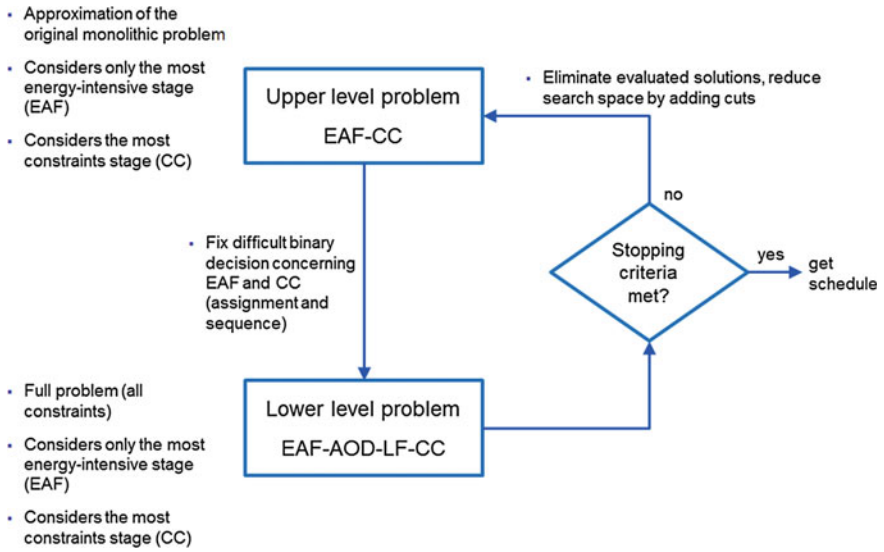


Fig. 13 Example decomposition approach for energy-aware stainless-steel scheduling (based on Hadera et al. 2015a)

as an input the lower level becomes solvable. Without the decomposition, even after 1 h of computation the monolithic problem could only achieve a solution between around 12–26 % far from the optimality. After employing the decomposition scheme, in 10 min the problem can be solved with less than 8–9.8 % far from optimum (Hadera et al. 2015a).

Nevertheless, production planning with energy considerations is just one of many opportunities for improvements related to energy. For melt shop process some examples are:

- Increasing specific energy efficiency by intensification of scrap melting at EAF by magnetic stirring;
- Optimization of scrap decomposition to reduce energy need, while obtaining the same heat composition;
- Reduction of waiting times between production stages to avoid cooling of heat by optimal production scheduling;

Product sequence coordination with hot rolling mill in order to reduce cooling-off of slabs after melt shop by holistic optimal site production planning.

6.3 Pulp and Paper Production

Energy is an important topic in pulp and paper production since some of the process types utilized can actually generate an energy surplus. Therefore, the sector is a

subject of many studies showing potential for using different methods. For example, Santos and Dourado (1999) use a genetic algorithms based approach for an optimal scheduling system for the mass and energy production in a kraft pulp and paper mill. The cost and production rate optimization includes steam and electricity production. In other study an entire supply chain planning of a pulping mill is tackled. The aspects of selling energy products (such as black liquor, liquid rosin, and bark) and energy (steam and electricity) on respective markets are considered to conclude that with the increasing prices of energy the importance of proper planning and coordination of energy management activities increases as well. On the wave of demand-side management studies a real case process of a stand-alone pulping mill under volatile energy prices is investigated by Karagiannopoulos et al. (2014). The considered mill is a member of a so-called balancing group which connects together energy suppliers with energy users.

Another real industrial process that can serve as an example of energy-intensive process is thermo-mechanical pulping (TMP). The input raw materials used in this process are trees cut into logs and chopped at chip mills to form wood chips. Wood chips or soft wood are next fed to the pulping mills, together with waste products (e.g., sawdust, pulp waste, or agricultural residues), to produce different types of pulp. Next, the valuable pulp is stored in large tanks from where it can be sold on the markets as a product or it can be used to produce paper rolls of various sizes using large paper machines. The final product can be newsprint, printing paper, and tissue paper. Apart from the wood, a large amount of energy is used in the form of steam and electricity, therefore the latter is an essential factor of the total production cost (Pulkkinen and Ritala 2008). The process is shown in Fig. 14.

Similarly to the steel example, one could coordinate the production process with time-volatile energy prices. However, in this case the process is in principle a continuous process with storage possibility. Therefore, the pulp product can be produced at energy-intensive refining stage in times where electricity price is cheap

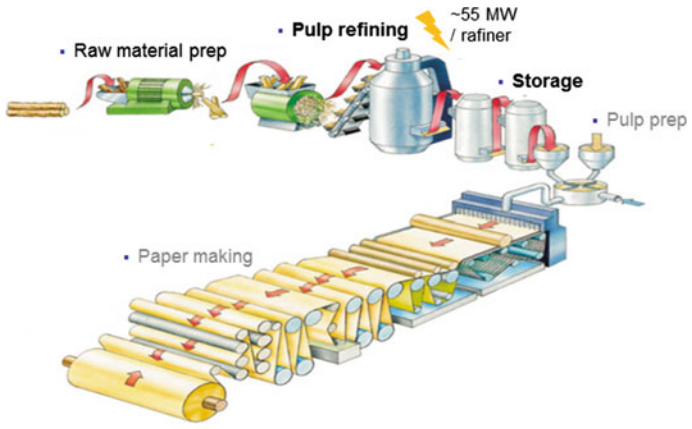


Fig. 14 Thermomechanical pulping process (TMP)

and released from storage tanks to satisfy demand during expensive peaks. Multiple refiners running in parallel cannot change their production mode without performing costly setup tasks. This creates an optimization problem which has been addressed in literature, for example, by functional separation of electricity cost optimization from production scheduling (Hadera et al. 2015b). Some other examples of potential improvements related to energy in pulp and paper industry include:

- Optimal control of processes that require energy, e.g., burning, evaporation;
- Reduction of energy losses in waste paper by optimized paper trimming;
- Precise measurement of product quality (e.g., moisture content and fiber weight) used for optimization paper machine control to achieve energy use reduction;
- Distribution optimization of wood fiber in order to reduce steam use;
- Minimize grade change losses at the paper machine;
- Optimal control of digesters to decrease steam use;
- Optimization of lime kiln temperature profile for minimum energy requirements.

6.4 Other Industrial Success Stories

There are a number of successful implementations of demand side-management in the industry. Many of them relate to the power industries but process industries also have achieved many significant achievements within this area. To make it short they are only mentioned briefly.

- Mining: A large mining company could reduce their energy consumption and CO_2 and SO_2 emissions by installing energy monitoring, and targeting for their headquarters three mines and two harbors. The payback time of the investments was measured in months.
- Steel industry: Another steel company could reduce their purchase of natural gas and electrical power by installing an energy optimization solution that takes into account the process conditions and provides decision support for site-wide gas distribution control and unit commitment for boilers.
- Pulp & Paper: An enterprise-wide centralized energy optimization combining the energy needs of 17 mills could reduce the CO_2 and SO_2 emissions, apart from providing optimal energy procurement and production planning. The payback time was also within months.
- Air separation: A plant producing air products was able to evaluate potential benefits from investments in storage tanks after considering seasonal fluctuation of energy prices. In addition, the plant identified longer time periods in which it was profitable to shut down the plant.

7 Concluding Thoughts

The complexity and the dynamics of production and energy systems require advanced computer-aided planning and optimization capabilities to fully reach the available potential. Well-designed user interfaces are crucial for communicating the relevant aspects of the energy solution to all necessary participants. It is very important that the models must be easily maintainable and understandable to reflect any changes in the operating environment, since the limitations and capabilities affecting an energy system are currently subject to continuous changes. Even if there are already very promising industrial success stories many challenges still remain.

- How to cope with changes within the energy supply structures
- How to deal with the growing problem sizes caused by the integration of additional aspects
- How to mitigate between conflicting or competing operational targets in integrated solutions
- How to efficiently integrate information flows concerning energy throughout all decision layers
- In general manage the targets: a proper balancing of, e.g., make span and energy costs is important
- Few industries have so far tapped the potential: in many cases the basic functions of energy metering, monitoring, and visualization do not exist
- How to select and utilize the relevant unused (not analyzed) data among the overflow of information

On one hand academic push of newly developed methods, on the other hand international standardization and commercial solutions should all be utilized to increase awareness and foster energy-efficient operations. Several approaches for full integration of production planning and energy management work well for academic size problems or specific industry use cases. Nevertheless, extending these to more generic large industrial-size problems is still within the world of research, which also calls for a close collaboration with the industry.

The topic of demand side-management is a true challenge as it combines different, earlier separated, disciplines and industrial targets. Process industries and power industries play here the main roles. It is still unclear how the business models will promote the collaboration, as well as, how political changes affect or steer the development. Demand-side management offers industry a partially unidentified potential to economically cope with the migration toward more renewable energy forms. Simultaneously, it offers the academic world a wonderful opportunity to realize EWO concepts practically without limits and with real industrial needs. The first research results are very promising but not sufficient to support the change. What, however, is clear is that the supply side will not alone be able to balance the supply–demand imbalance in a highly uncertain environment, until there exist cost-efficient and sufficient energy storage technologies. Today, it is possible to find

and define win–win–win opportunities for flexible industrial processes, ensuring that the industry process owner, the grid owner, and the power producer can all benefit and reduce their risks in this volatile business environment.

References

Ali, H., Price, L. K., McKane, A., T., 2007, The State-of-the-Art Clean Technologies (SOACT) for Steelmaking Handbook, Asia Pacific Partnership for Clean Development and Climate.

ANSI/ISA-95.00.03-2005, 2005, Enterprise-Control System Integration. Part 3: Activity models of manufacturing operations management, Research Triangle Park, NC: ISA – The Instrumentation, Systems, and Automation Society.

BASF, 2015a, BASF Factbook 2015 Information for investors and analysts, available online at: https://www.basf.com/en/microsites/factbook-2015.html, accessed 31.07.2015.

BASF, 2015b, Energy and climate protection, Online report 2014, available at: https://report.basf.com/2014/en/managements-report/responsibility-along-the-value-chain/environment/energy-and-climate-protection.html, accessed 31.07.2015.

BASF, 2015c, Sustainability – Energy Efficiency, available online at: https://www.basf.com/en/company/sustainability/environment/energy-andclimate-protection/energy-efficiency.html, accessed 31.07.2015.

Castro, P., Sun, L., Harjunkoski, I., 2013, Resource–Task Network Formulations for Industrial Demand Side Management of a Steel Plant, Ind. Eng. Chem. Res., 52, 36, 13046–13058.

Charles River Associates, 2005, Primer on demand-side management, The World Bank.

Demand Response Research Center, 2013, Demand Response Research Center, available at: http://drrc.lbl.gov/.

Engel, H., Smith, W. T., 2014, At the heart of the Verbund, Investor roundtable Chemicals segment 2014, presentation, available at: https://www.basf.com/documents/corp/en/investor-relations/calendar-and-publications/presentations/2014/140609_BASF_Investor-Roundtable-Chemicals_Boston-NYC.pdf, accessed 31.07.2015.

European Union, 2011, Energy 2020. A strategy for competitive sustainable and secure energy.

European Union, 2012, Directive 2012/27/EU of the European Parliament and of the Council of 25 October 2012 on energy efficiency, amending Directives 2009/125/EC and 2010/30/EU and repealing Directives 2004/8/EC and 2006/32/EC Text with EEA relevance, Official Journal of the European Union.

Gajic, D., Onofri, L., Harjunkoski, I., Hadera, H., Di Gennaro, S., 2015, Integrated production and electricity consumption optimization in a melt shop, submitted to Energy Conversion and Management.

Geiger, T., 2011, Energy Management at BASF Corporation, Center for climate and energy solution, presentation, available at: http://www.c2es.org/docUploads/geiger-austin-ee-workshop.pdf, accessed 31.07.2015.

Grossmann, I. E., 2005, Enterprise-wide optimization: A new frontier in process systems engineering, AIChE Journal, 51, 7, 1547–5905.

Hadera, H. and Harjunkoski, I., 2013, Continuous-time batch scheduling approach for optimizing electricity consumption cost, Computer Aided Chemical Engineering, Elsevier, 32, pp. 403–408.

Hadera, H., Harjunkoski, I., Sand, G., Grossmann, I. E., Engell, S., 2015a, Optimization of steel production scheduling with complex time-sensitive electricity cost, Computers & Chemical Engineering, 76, pp. 117–136.

Hadera, H., 2015, Integrated electricity demand-side management and scheduling of energy-intensive plants: application to stainless-steel and thermo-mechanical pulping industry, PhD thesis, TU Dortmund, Germany.

Hadera, H., Wide, P., Harjunkoski, I., Mäntysaari, J., Ekström, J., Sand, G., Engell, S., 2015b, A Mean Value Cross Decomposition strategy for Demand-side Management of a pulping process, Computer Aided Chemical Engineering, Elsevier, 37, pp. 1931–1936.

Harjunkoski, I., Merkert, L., Hadera, H., Gajic, D., Onofri, L., Shemeikka, A., 2015, Dicing the load: flexible production saves energy costs, ABB Review, 3, 15, pp. 18–24.

Hasan, M. M. F., Karimi, I. A., Alfadala, H. E., 2009, Optimizing Compressor Operations in an LNG Plant, In Advances in Gas Processing, Elsevier, 1, 179–184.

ISO, 2011, 50001: Energy management system.

Karagiannopoulos, S., Vrettos, E., Andersson, G. and Zima, M., 2014, Scheduling and real-time control of flexible loads and storage in electricity markets under uncertainty, 11th International Conference on the European Energy Market EEM14, 28–30 May, Krakow, Poland.

Merkert, L., Harjunkoski, I., Isaksson, A., Säynevirta, S., Saarela, A., Sand, G., 2015, Scheduling and energy – industrial challenges and opportunities, Computers and Chemical Engineering, 72, 183–198.

Pulkkinen, P. and Ritala, R., 2008, TMP production scheduling under uncertainty: Methodology and case studies, Chemical Engineering and Processing, Elsevier, 47, pp. 1492–1503.

Santos, A., Dourado, A., 1999, Global optimization of energy and production in process industries: a genetic algorithm application, Control Engineering Practice, 7, 4, 549–554.

Scheihing, P., 2014, Save energy through the superior energy performance program, Chemical Engineering Progress, September 2014, pp. 48–51.

Vosloo, J. C., 2006, Control of an underground rock winder system to reduce electricity costs on RSA gold mines, MSc thesis, Nort West University, South Africa.

Zargarzadeh, M., Karimi, I. A., Alfadala, H., 2007, Olexan, a tool for online exergy analysis, Proceedings of 17 the European Symposium on Computer Aided Process Engineering.

Renewable-Based Self-sustainable Operation of Isolated Islands

María Cristina Rodríguez-Rivero

Abstract Isolated regions require small to medium-size independent electrical grids to cover the energy needs of their populations. The development of renewable-source-based power generation systems is key to make these regions less dependent on external fossil fuels. At the same time they are the best places to assess the behaviour and reliability of renewable-based systems and their possible scale-up. This chapter focuses on the study of a hybrid wind-pumped hydro generation system designed to provide electricity to the entire population of the island of El Hierro (Canary Islands, Spain), with the initial aim of covering 100 % of its energy needs, including water desalination processes. The involvement and support of national and international governmental organizations and the welcome of the local population were key in the promotion of the project. After its approval, civil engineering works are crucial in the development of a project of such magnitude. In a more general context, this chapter also dedicates a section to describe the basics in the use of simulation and optimisation algorithms as a tool to help on the decision-making process for the design and selection of power plants and their components.

Keywords Isolated region · Self-sustainability · Wind-hydro pumping · Electric system · Energy storage

1 Background

Isolated small islands are highly dependent on imported fossil fuels for their energy needs. In addition, these regions require small individual electrical systems, in comparison with larger inland territories. Both the size of these individual grids and

M.C. Rodríguez-Rivero (✉)
Department of Engineering, University of Cambridge, 17 Charles Babbage Road, Cambridge CB3 0FS, UK
e-mail: mcr51@cam.ac.uk

M. Martín (ed.), *Alternative Energy Sources and Technologies*,
DOI 10.1007/978-3-319-28752-2_16

the dependence on external resources lead to less stability and security in the energy production and make these remote regions very vulnerable to changes in oil prices. Finally, the low level of demand makes complicated the use of economies of scale, which leads to higher investment and operating costs per unit of generated energy (transport, materials, equipment and fuels). Consequently and particularly in the case of developing regions the population in these areas usually faces insufficient and unreliable power supply. This situation can lead to a lack of basic needs not only for covering commercial and residential general purposes but also to guarantee freshwater supply. In a large amount of remote islands, seawater desalination or groundwater pumping is ubiquitous and these processes can consume large amounts of power.

In remote islands and small isolated regions in inland environments there is a clear underutilization of renewable energy sources. An increasing knowledge on the exploitation of these sources and an increasing development and efficiency of the renewable-based generation technologies have led in the last decades to diverse projects for the electrification of remote regions. These projects have a variety of goals, being the main objectives twofold: first a mean to avoid the dependence of fossil fuels in small areas (becoming the lowest cost source of energy generation with an adequate implementation) and second, to study and analyse the behaviour and reliability of a variety of renewable-sourced electricity production strategies for their implementation in developing and under-developed regions and for the scale-up to denser populated places. Of course, renewable-sourced electricity production also contributes to the fight against climate change and help in achieving EU targets of 20 % renewables in 2020 and the reduction of greenhouse gas emissions by 80 % by 2050.

Renewable energies are dependent on natural resources that most of times are intermittent or even non-existent. Moreover, accurate long-term prediction of the resource availability is very challenging (see Chap. "Emerging Optimal Control Models and Solvers for Interconnected Natural Gas and Electricity Networks": Forecasting of energy and availability of natural gas and Chap. "Solar Energy as Source for Power and Chemicals": Solar energy for a description of forecasting methods). Therefore, the use of energy storage capabilities is key to promote variable renewable energy sources integration in electric grids since they allow developing a stable and reliable energy generation.

Several energy storage systems (EES) are being examined in a variety of projects. These systems provide regulating power and balance deviation produced by the intermittence of solar photovoltaic and wind power generation. Batteries, flywheels, capacitors, pumped water, super conducting magnetic, cryogenic, and thermal and compressed air energy storage systems are among the most popular EES (Arabkoohsar et al. 2015; Bhuiyan 2012).

Every remote region presents different challenges in terms of developing a green safe independent grid. These challenges are highly related to the variety, amount and availability of resources. However, some common factors can be pointed out as Indradip Mitra does when surveying different remote regions (Mitra 2006). According to his study in small off-grid areas diesel generation is the most common

alternative as a means of rural electrification, becoming solar photovoltaic energy increasingly important in the last decades. In places such as India and Bangladesh attempts have been also done with half megawatt biomass gasifiers and wind-diesel hybrid-based medium voltage minigrids though still a large amount of population does not have access to electricity.

From the islands surveyed, wind power is by far the most utilised renewable energy resource for electricity generation with over 50 % of islands using this form of energy followed by hydropower. Nearly, 10 % of islands utilising renewable energy for electricity generation have also used biomass.

In Europe several projects aiming at the electrification of remote islands through renewable energy have been done or are currently on progress. A large renewable-based island project was initiated in Samsoe in Denmark in 1996, followed by Pellworm (Germany), Aeroe (Denmark), El Hierro (Spain), Gotland (Sweden), Hiumaa (Estonia), Lemnos (Greece), Achill (Ireland) and Ithaka (Greece).

This chapter will focus on the electrification of El Hierro, where a hybrid wind-pumped hydro capability has been installed. It aims to supply energy to a population of approximately 11,000 inhabitants. Sections covering the state of the art and the European context of renewable-based systems precede the motivation and incentives that promoted the development of the wind-hydro system in El Hierro. A general overview about how optimisation and simulation algorithms help in the prediction of the system behaviour and the decision-making process follows. Finally, the characteristics of the currently working power plant are detailed.

2 Island Description

The Atlantic archipelago of the Canary Islands (Islas Canarias) comprises seven main islands lying north-east of Africa, being one of the outermost regions of the European Union. The shortest distance between the archipelago and Spanish mainland is around 1000 km.

El Hierro is the more western of the islands and the smallest in area and population (see Fig. 1). Its 10,700 residents live mainly on tourism and agriculture, where the service sector is the most prominent and farming sets aside as a secondary family activity. The island's climate is characterised by a relatively constant warm temperature throughout the year, scarce precipitation and high insulation time. Its mild climate, famous diving sites and nature encourage a more ecologically minded tourism.

El Hierro is home to native and unique plant and animal species, among which the giant lizard (Gallotia simonyi) is particularly known. It is a mountainous and steep island with volcanic origin. Table 1 shows a brief summary of the main demographic and meteorological characteristics of the region.

Fig. 1 Location of the island of El Hierro and the archipelago of the Canary Islands from a satellite view. Imagery: © 2015 NASA, TerraMetrics, Map data © 2015 Google, Inst Geogr. Nacional

Table 1 Demographic and meteorological characteristics of the island of El Hierro

Parameter/unit	Value
Area/km^2	268.7
Altitude/m	1501
Total population (2014)[1]	10,675
Annual growth (2013–2014)[1]	−2.8 %
Average annual growth (2000–2014)[1]	1.3 %
Average minimal temperature[2a]/°C	16.4
Average annual minimal temperature[2b]/°C	18.7
Average maximal temperature[2c]/°C	26.5
Average annual maximal temperature[2d]/°C	23.3
Average annual rainfall/mm	206
Average minimal insulation hours (in December)	139
Average maximal insulation hours (in May)	243
Average annual insulation hours	2382
Wind conditions (Alíseos)	North–east trade winds

[1]Instituto Canario de Estadística (ISTAC)
http://www.gobiernodecanarias.org/istac/
[2]AEMET
http://www.aemet.es/es/serviciosclimaticos/datosclimatologicos
[a]Referred to the coldest month—February—for a period ranging from 1981 to 2010
[b]Average of the mean minimal temperatures for all 12 months for a period ranging from 1981 to 2010
[c]Referred to the hottest month—September—for a period ranging from 1981 to 2010
[d]Average of the mean maximal temperatures for all 12 months for a period ranging from 1981 to 2010

3 Electricity Sector

The generation of electricity in the Canary Islands is characterised by a high percentage of thermal generation, with a small remaining part from renewable energies (Table 2). The total electric power installed amounted 3208 MW in 2014 (Red Electrica de España S.A.U. 2014) with slight changes in the last years except for the new power plant subject of this chapter, corresponding to 12 MW in El Hierro. In addition, the annual power demand fell 0.5 % from 2013 to 2014 (Red Electrica de España S.A.U. 2014). Tables 2 and 3 and Figs. 2 and 3 summarise data about generation capabilities and installed capacities in the archipelago.

The remote location of the islands makes them singular in terms of energy generation and distribution, with the need of an independent power grid from mainland Spain. The depth among the islands is considerable and as a consequence grid connection by sea cable is not feasible among them, with the exception of the two more eastern ones—Lanzarote and Fuerteventura. These islands were connected through a 14.5-km-long submarine cable in 2005, with a transmission voltage of 66 kV. A project to instal a new backing cable (135 kV) is in progress at the moment. This connected system and the electrical grids corresponding to the two capital islands are the largest in the archipelago. The remaining three isolated systems are smaller where the system of El Hierro is 80 times smaller than that of Gran Canaria, and this in turn, 80 times smaller than the system in mainland Spain (see Fig. 3).

The isolation of the small to mid-sized electricity grids and the different growth rates of the demand make the energy policy on each island to be tailored in different manners to specifically meet the needs of every island.

In addition to the common industrial, residential and farming use of electricity, almost a 30 % of the electricity generated in the Canary Islands is used for the

Table 2 Energy production in the archipelago (Gobierno de Canarias 2014)

Energy source	Energy production/GWh	Relative percentage change comparing year 2013 to year 2012
Fuel/gas	4874	−7.5
Diesel motors	2260	3.2
Gas turbine	364	−4.3
Vapour turbine	2250	−16.6
Combined cycle	3429	9.4
In generation consumption[a]	−422	−3.5
Hydro-wind	1	–
Other hydraulic	3	14.5
Wind	401	10.6
Solar photovoltaic	287	−0.8
Renewable thermal	8	−1.5
Total generation	8579	−0.5

[a]Data related to fuel/gas and combined cycle generation

Table 3 Installed facilities in the archipelago (Gobierno de Canarias 2014)

Energy source	Installed power/MW	Relative percentage change comparing year 2013 to year 2012
Fuel/gas	1918	0.0
Diesel motors	566	0.0
Gas turbine	639	0.0
Vapour turbine	713	0.0
Combined cycle	920	0.0
Hydro-wind	12	–
Other hydraulic	1.5	0.0
Wind	154	0.0
Solar photovoltaic	166	0.6
Renewable thermal	3	0.0
Others	33	0.0
Total generation	3208	0.4

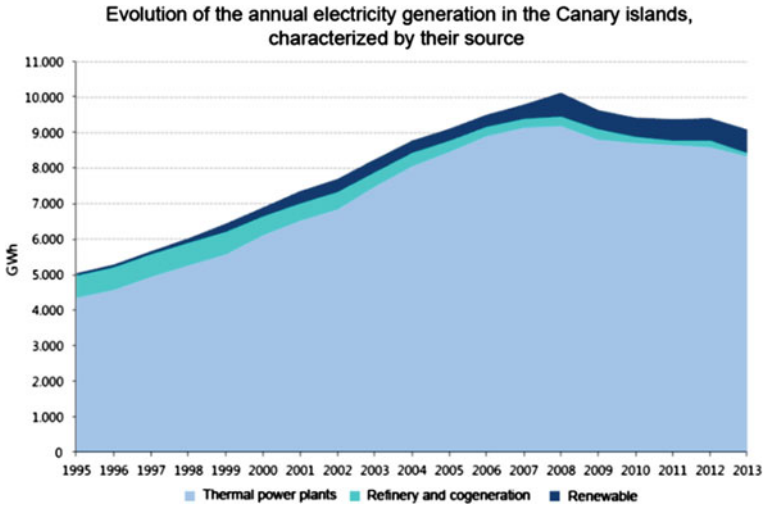


Fig. 2 Evolution of the annual electricity generation in the Canary islands in the period from 1995 to 1997, characterised by their source (Gobierno de Canarias 2014, with permission)

production of drinking water through pumping ground water and desalination (Bueno and Carta 2006). In El Hierro, consumptions from 40 to 50 % for this purpose have been reported (Gorona del Viento S.A. 2014).

From the 90s, the Government of the region encouraged an increase in the use of renewable energies and planned to exploit primarily the wind potential of the islands. However, due to system reliability concerns wind penetration is limited and the use of energy storage systems has been then contemplated for the islands. Thus,

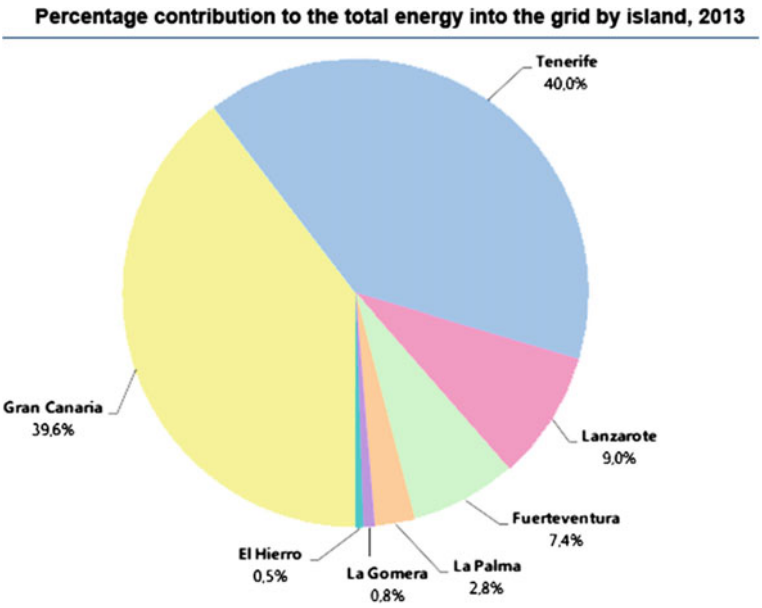


Fig. 3 Contribution to the electricity generation in the year 2013 by every island (Gobierno de Canarias 2014, with permission)

hybrid generation schemes such as wind-diesel systems for sea water desalination, hydrogen production or water pumping have been considered (Bueno and Carta 2005b).

The total consumption of fossil fuels reached 1,776,360 Tm for the archipelago in 2013. El Hierro had a consumption of 10,275 Tm, increasing a 1.1 % compared to 2012 (Gobierno de Canarias 2014).

3.1 The Electric System of El Hierro

Before the new hydro-wind plant was in operation the total generation capacity in El Hierro was 15.1 MW (December 2013). It combined 14.9 MW from thermal generation (ten diesel motors) and small wind and solar photovoltaic systems connected to the grid. The system total generation was 45.90 GWh in 2013. The 99.4 % of the demand was covered from the diesel generators and the remaining 0.6 % of the demand was generated from renewable sources (Gobierno de Canarias 2014). The renewable sources in the island were limited to two wind turbines (140 kW each) connected to the main grid and isolated panels for solar photovoltaic (8488 kW) and heating (390 m^2) (Bueno and Carta 2005b). The peak demand in the island for the period from 2005 to 2013 was 8.6 MW, reached in 2013 (Gobierno de Canarias 2014).

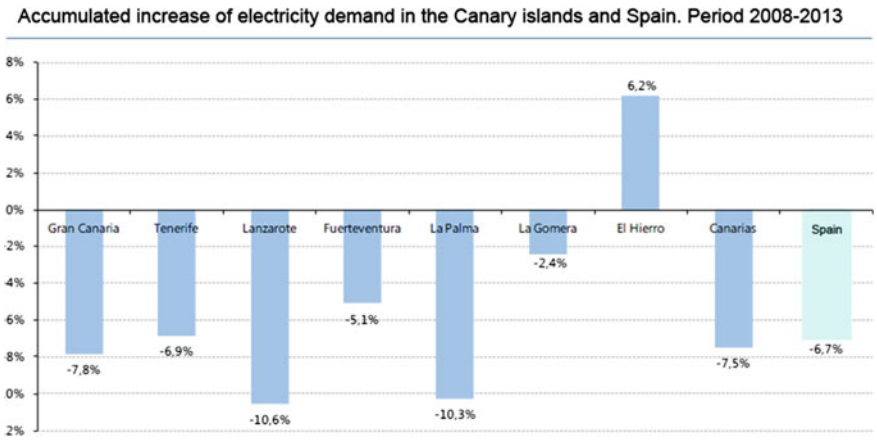


Fig. 4 Accumulated increase of electricity demand in the Canary Islands and Spain for the period between 2008 and 2013 (Gobierno de Canarias 2014, with permission)

El Hierro experienced a 3.2 % increase in the average annual growth of installed generation for the period 2008–2013 (see Fig. 4), with a total increase from 3.8 to 15.1 MW for the period between the years 1985 and 2013 (Gobierno de Canarias 2014). In a similar way, in terms of electricity generation into the grid, El Hierro has increased from 6.56 to 44.02 GWh between 1985 and 2013 with a 1.2 % of annual growth rate in the period 2008–2013 (Gobierno de Canarias 2014).

4 Historical Approach of the Self-sustainable System in El Hierro

El Hierro was the first in the Canary Islands to adopt a Sustainability Plan. This plan was sanctioned by the Cabildo of El Hierro in November 1997, including the idea of developing a generation system to make the island energetically self-sufficient. The idea was supported by the UNESCO, designing the island a Biosphere Reserve in 2000, with a 60 % of its territory protected.

This award increased the motivation for preserving the natural habitat and traditions of the island. In the socio-politic aspect the idea of creating a greener island that attracted more ecologically minded practices and tourism was then a key point for the development of the project. From the energy production point of view, a renewable generation electricity system would bring the benefits of a secure, economic and self-sufficient supply.

Such an electric system should be designed to cover current and future demand, as well as being as profitable as possible. Therefore, the equilibrium between the economy of the island and the management of the natural resources, wastes and energy were crucial to determine the type of power plant to be designed.

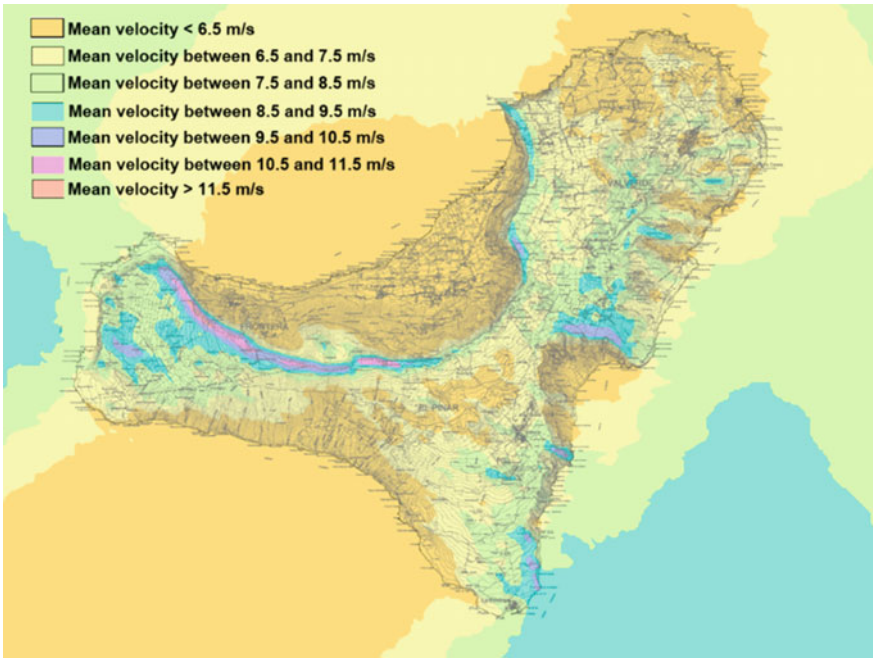


Fig. 5 Wind map of the island of El Hierro, showing average wind speeds (Cabildo of El Hierro)

In terms of available resources, El Hierro has a large number of sun hours with high radiance and good wind potential, with constant availability due to the trade winds. These winds blow throughout the year from the Tropic of Cancer to the Equator. However, the changing topography of the island and the sea depth in the coast reduce the option of exploiting large solar or wind plants (see Fig. 5 for wind data). Yet, the altitude variation potentially allows the exploitation of hydropower plants though the low rainfalls preclude the use of conventional hydropower energy. Additionally, spring water scarcity is a key factor that implies the need of desalinated water and halts crop harvesting for its use as biomass. Finally, the potential of ocean and geothermal energies have not been explored in detail in the region.

4.1 European Context

A brief introduction to the panorama of self-sufficient renewable energy-based regions in Europe should contain the examples of the island of Kythnos in Greece and the island of Samsø in Denmark.

In both cases the population is smaller than El Hierro's and the characteristics of the grids are different. Kythnos has approximately 2000 inhabitants and an installed capacity that ranges from 300 kW in winter to 2000 kW during the summer

(touristic high season). Samsø has 4,100 inhabitants, a total production capacity of 11 MW and a system that is connected to the mainland grids.

Kythnos installed the first commercial wind park in Europe in 1982 with five turbines of 20 kW each. It is one of the first references to study the wind energy penetration into electricity systems. With also a good solar potential, the installed renewable energy capacity in Kythnos increased in the following years, creating a hybrid system incorporating also a 100 kW solar photovoltaic system and diesel generators. The wind turbines were replaced around a decade ago and some photovoltaic installations became obsolete but high emphasis is being put on microgrids relying on solar energy, with the particular example of the Gaidouromadra valley community, with the intention of using the solar potential of Kythnos and neighbouring areas in similar ways (Cattaneo 2008).

More than a decade later, in 1997, Samsø in Denmark started a programme aiming to convert all energy supply into 100 % renewable energy within a 10-year period. An Energy and Environmental Office was established to counsel the citizens wanting to start their own renewable energy projects. This was the first project of its kind in Europe, being a reference to promote awareness among the local population. The objectives were manifold (Saastamoinen 2009), mainly aiming to cut in energy consumption and increase the efficiency of heat (in individual heating systems), electricity and transport systems (gradual conversion of the transport sector from petrol and oil to electrical power, and later to hydrogen). Since the start of the project, Samsø invested on various local energy production initiatives, particularly on wind power, biomass and solar energy with currently 10 offshore and 11 onshore wind turbines, where most of the investments have come from local residents. A qualified guess is that a total of 53.3 M€ has been invested into the project, of which 4 M€ has come from public subsidies. An independent evaluation in 2007 deemed the project almost completely successful, though some consider that still more efforts are needed to change the behaviour in energy saving and conservation (Saastamoinen 2009).

Both projects are examples of both the ambition to develop independent grids in remote areas and to create a social awareness of the need of more sustainable practices and energy production processes.

5 Wind-Pumped Hydro Power Plant in El Hierro

The electricity system in El Hierro has to supply energy to cover the population needs in terms of common electricity consumption (domestic and industrial use) and water consumption for drinking and irrigation, where seawater desalination plants in El Hierro can consume between the 30 and 50 % of the generated electricity in peak times. Until recently, the total electricity demand of the island was covered by diesel generators with an installed capacity able to cover peak demands of around 9 MW (approximate value of the peak demand in the last years). Fortunately, due to the even mild temperatures, the sea buffering effect on the

climate of the region and an even distribution of visitors through the year, there are few seasonal variations in the energy demand.

In the 80s, the Department of Alternative Energies in the energy company UNELCO—only energy company in the Canary Islands at that time—had already thought about the possibility of harnessing the constant winds, particularly in the west region of the island, and benefit from the steep orography through an hydraulic jump. But it was not until the decade of the 90s and early 2000 that clear signs of support towards renewable energy projects were a fact, as mentioned before. Finally, in 2004 the Energy and Transport commissary supported the plan with 2 million of Euros.

Under this favourable context, Unelco-Endesa—UNELCO was absorbed by the multinational Endesa in 1988—and the Technological Institute of the Canary Islands (ITC) performed technical studies under the same idea of combining wind and hydro power. These studies backed a project for building a hydro-wind power plant and the public–private consortium Gorona del Viento El Hierro S.A. was created in 2004 with the participation of the insular Council (60 %), Endesa (30 %) and the ITC (10 %), with the aim of making the project come true.

The Spanish government provided the consortium with financial support through the Institute for Diversification and Saving of Energy IDAE with 35 million euros for the construction and commissioning of the initially 79 million euro project. The support from the Spanish State Budget was granted in the framework of the activities carried out for energy diversification, renewable energy use and respect for the environment (Gorona del Viento S.A. 2014), including the technological innovative character of the project by itself and also the character of the project as a means to examine the hydro-wind power generation model in a medium-size context and the management of hybrid systems, particularly those that allow for a higher penetration of wind power.

The information extracted from the management of the plant will be of application to other islands and to mainland integrated energy systems, facilitating the integration of wind energy in the national and international systems.

The milestones during the construction and commissioning were (Gorona del Viento S.A. 2014):

- 80s: UNELCO-El Hierro assesses the idea of introducing renewable alternatives into the generation system.
- 90s: Incorporation of the ITC and draft of the project to integrate renewable energy sources. The support from European and national organizations is found.
- 2004: Creation and foundation of Gorona del Viento El Hierro.
- June 2007: Approval of the final design, environmental impact assessment and declaration of public utility of the 11.32 MW hydro-wind power plant. Most of the needed financial investments are guaranteed at this point.
- March 2008: Positive Urban Feasibility Report and request for bids.
- August 2009: Start of civil engineering works.
- 2013: End of civil engineering works.
- June 2014: Inauguration of the wind-hydro power plant.

6 The Technical Idea

The aim of the project was focused from the first moment on the construction of a power plant that coupled wind generation with a pumped hydro storage system with enough capacity to cover 100 % of the yearly electricity demand of the island of El Hierro. This system would be interconnected to the already working electrical system grid of the island (property of Endesa).

Different ways of integrating such a combination of power sources into the grid have been described in the literature as it will be covered in the following subsection. In this particular case the wind farm was designed to supply electricity directly into the grid and, simultaneously, to start up a pumping system when the wind energy production exceeds the energy demand. The pumping station loads the energy storage system by transferring water from a deposit at a low level (lower reservoir) into a deposit at a higher level (upper reservoir). The stored potential energy of the water (water head) can then be used by hydropower turbines for two main situations: 1. for generating electricity when wind power is not enough to cover the energy demand and 2. for ensuring the stability of the network, where at least one of the turbines is used in compensating mode for power and frequency even when there is enough wind power (Chap. "Large-Scale Stochastic Mixed-Integer Programming Algorithms for Power Generation Scheduling": Power distribution: Smart Grids).

6.1 The Decision-Making Process

As mentioned before, in small not interconnected regions where wind energy is available, the use of hybrid systems becomes paramount to control the reliability of the systems, given the intermittence of winds. Therefore, hybrid schemes with energy storage systems are essential to allow wind penetration increases in conventional power grids.

Among them, the use of wind-hydro hybrid systems has been deemed as the best solution for small not interconnected regions (Anagnostopoulos and Papantonis 2007). As storage system by itself, water pumping is considered a very good alternative for the large amount of energy that can be potentially stored and for the easy regulations and high efficiency of the power output (Bueno and Carta 2005a). Therefore, the use of wind-hydro hybrid systems is a very good alternative when there are suitable locations, as El Hierro shows, with high winds and steep valleys.

Furthermore, the pumped hydro storage system adapts very well where a large storage time constant (the ratio of rated capacity to rated power) is required. This is the case on the island of El Hierro where the storage system is designed to supply the island during two full days if wind power is not available.

In a first phase, a technical feasibility study to establish the best configuration of air turbines, water turbines, pumping equipment and capacity of the deposits to be installed in the new plant was performed (Gorona del Viento S.A. 2014). These studies were based on previous analyses as those by Bueno and Carta (2005a, b).

Fig. 6 Location of Gorona del Viento power plant. Imagery: © 2015 NASA, TerraMetrics, Map data © 2015 Google, Inst Geogr. Nacional

In this way, the engineering studies done by Endesa agreed with the best location for the power plant determined in the study by Bueno and Carta. This location lays in the coastal region at the eastern part of the islands (Playa de Tijeretas, see Fig. 6). The wind park is placed in a region with average annual wind speeds of 9 ms^{-1} and the valley has an altitude gradient of 650 m with good areas to build large water reservoirs at different levels (Bueno and Carta 2005b).

Bueno and Carta (2005b) based their study on the expected demand for the year 2015, approximately 48 GWh/year according to the Energy Plan for the Canary islands for the year 2006. However, the final project was designed to cover the expected demand in the year 2030, as modular extension for the components would be challenging to execute. The capacity of the upper reservoir was increased to meet the water needs required by the island (Gorona del Viento S.A. 2014).

The main components of the system comprises two water deposits, a wind farm, a hydroelectric plant and a pumping station, in connection with the diesel power station previously installed (see Figs. 7 and 8 for diagrams). Once the power plant completes testing operations the diesel engines are meant to be used only in exceptional cases when wind power and water storage are insufficient to produce the energy required.

The particular characteristics of the components of the plant must be determined through a detailed analysis of the system and the chosen operation mode. In the literature there are some studies on hybrid systems for the production of energy in

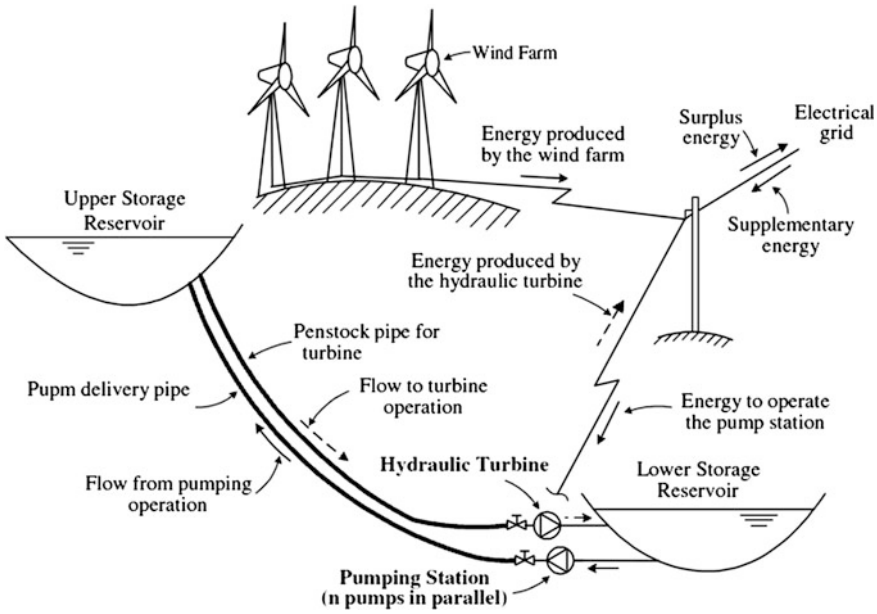


Fig. 7 Diagram of the hybrid wind-hydro plant with its main components (Anagnostopoulos and Papantonis 2007, with permission)

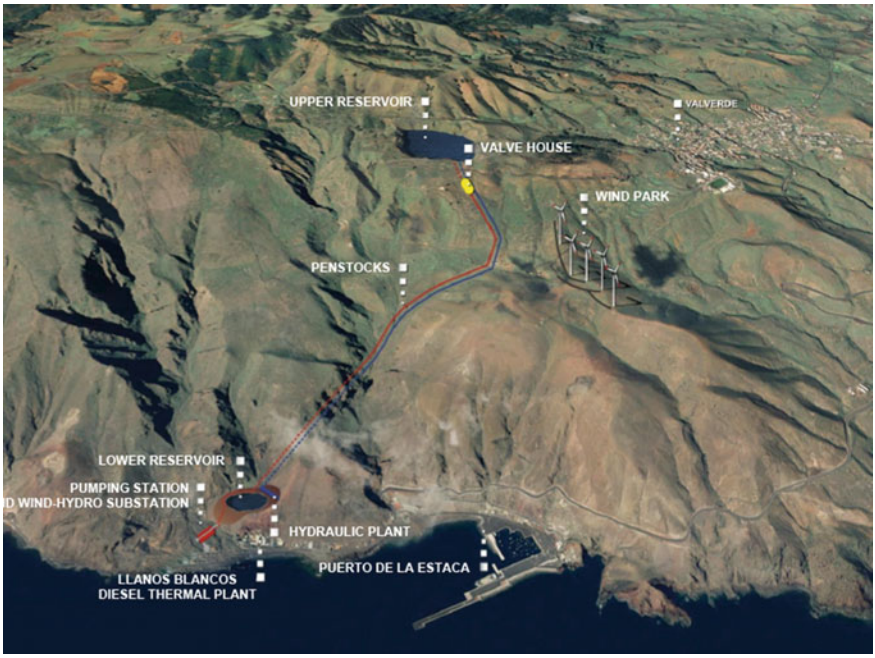


Fig. 8 Wind-pumped hydro generation plant, coupled with a previous diesel plant, in El Hierro (Gorona del Viento S.A. 2014, with permission)

islands or interconnected regions that describe different models to evaluate their technical feasibility and allow predicting the optimal solutions in terms of technical and commercial productivity (Anagnostopoulos and Papantonis 2007; Bueno and Carta 2005a, b, 2006; Benitez et al. 2008; Chap. "Large-Scale Stochastic Mixed-Integer Programming Algorithms for Power Generation Scheduling": Power distribution: Smart Grids).

In general, nonlinear constrained optimisation programmes are used to determine the best possible allocation of energy across the different generation systems and the optimum management of reservoir volumes and water flows contemplating multiple scenarios (Benitez et al. 2008).

It is worth mentioning that these studies generally evaluate two different types of operation modes for the hybrid wind-hydro plant. A first approach considers that the plant operates with two independent electrical subsystems, where the entire wind power production is intended to pump water and only the hydraulic generating subsystem is connected to the generation grid. If needed, the hydraulic generation is then combined with conventional generation systems to cover the demand (such as diesel motors). Different strategies can be defined in this context, where the control system would attempt to cover the demand with hydraulic energy as long as the level in the upper reservoir is higher than the minimum technical value for the turbines to be driven or higher than a minimum volume that allows to cover entirely the demand by itself, with or without the coupling of the conventional system (Bueno and Carta 2005a).

A second approach depicts the scenario that has been chosen in El Hierro, where both generating subsystems (wind and hydraulic) are connected to the island electrical grid. It is straightforward to think about the direct integration of the entire wind generation to the grid, using any surplus energy, not absorbable by the grid, to pump water from the lower to the upper reservoir. However, as stated above, the implications of the direct connection of wind production into the grid may cause imbalances. Therefore, the degree of wind direct contribution should be limited to a percentage of the capacity connected to the grid. The contribution percentage should be determined by dynamic studies which take into account the load of the system, the wind speed variation and other operating conditions (Bueno and Carta 2005a). Again, different strategies defining various requirements for the use of hydraulic power and conventional systems can be assessed under this approach.

The use of computer algorithms helps in the decision-making process by optimising the size and number of components (optimisation models) and by allowing to simulate daily and yearly operation of the different considered systems to evaluate the economic results (simulation models). The algorithms perform therefore an interrelated technic and economic analysis using empirical cost-estimation relations considering all possible combination of components and including design and construction costs. These algorithms calculate then the initial investment costs as well as maintenance costs and incomes, which helps in the decision-making process.

In general terms the algorithms used for an optimised design of a wind-hydro plant as that in El Hierro should consider:

- Wind data of the island (the higher resolution the better output estimation).
- Characteristics operation curves of the commercial pumps, turbines and motors considered.
- Constraints and additional components in the pump operation (variable-speed drivers).
- Fossil fuel costs.
- Financial indexes on electricity tariffs and their prediction for coming years.
- Other financial parameters such as inflation rates, loan amortization rates, etc.

The calculations are based on empirical expressions accounting for the operation of the components. These relations, coupled with different constraints of every considered strategy, result in nonlinear complex programmes. Different assumptions such as constant height difference between upper and lower reservoirs or negligible water level variation are generally used to limit the number of parameters.

Anagnostopoulos et al. presented a numerical study to find the optimal size and strategy for the pumping station unit in a hybrid wind-hydro plant, considering all wind generation to cover pumping needs exclusively (Anagnostopoulos and Papantonis 2007). They compared the typical design with identical pumps working in parallel mode (1) with two other systems: adding additional pumps of smaller size (2) and use of a variable-speed driver, (3) fixing the capacity of the water reservoirs, the wind energy and the hydraulic turbine production. They found that the configuration with variable-speed driver was the most advantageous to combine the highest energy storage rate with the lowest investment cost.

Bueno and Carta (2005a) performed a more complete study for the technical and economic sizing of the various components that constitute medium-sized wind-powered pumped hydro storage systems. They examine six different strategies, where in addition to Anagnostopoulos et al., they also considered the connection of the wind subsystem to the grid. Their model estimates the net present value, internal rate return and the payback period for each strategy studied.

They applied their model to the island of El Hierro (Bueno and Carta 2005b) where they used hourly wind speed data and the costs of the commercial components, civil works and control infrastructure, among others but considering a maximum wind penetration of 30 % (α).

Under these conditions they found that the two best strategies contemplate both systems connected to the grid. Both of them cover a minimum percentage (100–α) of the demand with the hydraulic turbines (or the diesel motors if needed) and a maximum percentage α by the wind turbine. The two strategies consider different requirements for the level of the upper reservoir (the load of the storage system). One of these two strategies, however, drastically lowers the hydraulic turbine and motor start-ups and shut-downs. Even when the obtained unit energy cost is slightly higher, this strategy facilitates control and decreases maintenance costs.

While moving forward in the designing stage, however, the system was redefined to operate with a larger wind penetration direct to the grid (baseload) and to use the surplus power, if existent, for water pumping. Thus, the hydroelectric plant

would work as load-following generation source, compensating the system to guarantee stability. A higher direct penetration of wind power avoids energy losses in the pumping and hydroelectric processes.

The size of wind turbines, water turbines and pumps are similar to those considered in the cited studies, being characteristics of small- and medium-size grids to ensure flexibility and stability of the system if one of the components fails.

The particular technical and economic studies for the particular case of El Hierro are not publically available but, from the studies of Bueno and Carta, optimisation and simulation algorithms were used to help in the prediction of the system behaviour and the decision-making process, with slight different criteria taken into account, particularly in terms of wind penetration percentage.

6.2 *The Components of the System*

The study by Bueno and Carta (2005b) for the sizing of the plant concluded that, whichever the chosen strategy, the optimal pumping configuration comprises pumps of different capacities and high power rating hydraulic turbines (3300 kW in this case compared to 750 and 1500 kW ones). Similar results are found in the power plant. The basic components and their characteristics are summarised next. Figure 8 shows a diagram of the power plant.

6.2.1 Wind Plant: 11.5 MW Power Plant (Fig. 9)

- 5 wind turbines Enercon E-70 2.3 MW each.
- Hub height: 64 m.

Fig. 9 Wind park in El Hierro (Gorona del Viento, with permission)

- Blade length: 35 m.
- Approximate weight: 250 tonnes.
- Foundation: 375 m^3 of concrete and 43.11 kg of iron per wind turbine.
- Other: total electrical wiring to the connection station approximately 49 km, optical fibre wiring for communication approximately 7.5 km. One month of transportation time between the closest dock (Puerto de la Estaca) and the final location.

6.2.2 Water Storage System/Lower Deposit

At first, the use of seawater was considered for the water storage system. However, the increase in the costs of the components that seawater would involve and the potential use of spring water also as additional drinking water storage for the island made the decision of using spring water for that purpose. Therefore, the water that fills the deposits comes from one of the desalination plants in the island, which supplies water when needed to compensate the water losses mainly by evaporation.

- Volume: 150,000 m^3 (smaller volume than the initially designed due to technical difficulties. A second lower deposit has been projected).
- Depths level and its perimeter: 43 and 165 m.
- Water maximum level: 56.00 m.
- Crest level and its perimeter: 57 and 560 m.
- Total area: 23,138 m^2.
- Height and length of outer dam: 24 and 160 m.
- Other: 22,182 m^2 of polypropylene (PP) geotextile layer (350 gm^{-2}) and 25,266 m^2 of high density polyethylene (HDPE) waterproof coating (2 mm width) (Fig. 10).

Fig. 10 Lower deposit in construction (Gorona del Viento, with permission)

Fig. 11 Pumping station (Gorona del Viento, with permission)

6.2.3 Water Storage System/Pumping Station (6 MW)

- 2 large pumps (1500 kW each).
- 6 small pumps (500 kW each).
- 1500/500 kW variators.
- Conductions: Discharge conduction: 3015 m long, 0.8 m diameter; aspiration conduction: 188 m long, 1 m diameter (Fig. 11).

6.2.4 Water Storage System/Upper Deposit

- Volume: 380,000 m^3 (enough for covering the island demand for 4–5 days).
- Depths level and its perimeter: 698 and 300 m.
- Crest level and its perimeter: 715 and 854 m.
- Total area: 50,000 m^2.
- Height and length of outer dam: 24 and 160 m.
- Other: 40,000 m^2 of polypropylene (PP) geotextile layer and 45,000 m^2 of high density polyethylene (HDPE) waterproof coating. It also serves as a drinking water reservoir (Fig. 12).

Fig. 12 View of the upper reservoir (Gorona del Viento, with permission)

6.2.5 Hydraulic Turbine Station

The selection of the turbines is dependent on many factors though two are key parameters: water head (available hydraulic jump) and flow rate. For this particular case, where the water head is large and the flow rate is low Pelton turbines are more efficient.

- 4 Pelton turbines (2.83 MW each) with flywheel (ANDRITZ AG, Austria). The flywheels help to reduce pressure surges when the flow rate changes due to changes in the grid demand and to give a more stable control speed (improved frequency stability).
- Water head: 655 m.
- Maximum flow rate: 2.0 m^3/s.
- Conduction to turbines: 2350 m long and 1 m diameter (Fig. 13).

6.2.6 Electrical Substation

Double bus bar system with connection point in Llanos Blancos diesel substation.

Fig. 13 Turbine station (Gorona del Viento, with permission)

6.3 Civil Works and Other Concerns in the Power Plant

Interesting technical and environmental situations arose from the civil works. Thus, for the construction of the lower deposit the presence of thin lapilli (small stones from volcanic eruptions, see Fig. 14) involved technical difficulties that finally made the deposit to be smaller than initially projected. The resizing of the deposit led to an unexpected limitation in the amount of water that could be released from the water storage system's upper reservoir. Thus, a second lower deposit is being designed at the moment.

Regarding the layout of the penstocks there were both environmental and technical concerns. Thus, some of the sections of penstocks passed through native vegetation (Fig. 14). This vegetation had to be temporary removed and replanted once the installations were completed. Technically, in some particular sections the stresses in some of the elbow joints of the penstocks involved the construction of large concrete blocks to compensate the excess of stress in the area.

6.4 Renewable Penetration and Generation Unit Price

The system in El Hierro is still in testing phase and few operational data are yet available. The application of the model developed by Bueno and Carta (2005b)—with a slight different approach—indicates a technical annual renewable maximum

Fig. 14 Lapilli walls (on the *right*), native vegetation (on the *left*) and pumping station (Gorona del Viento, with permission)

penetration percentage of 68.40 %. Tracking the real-time generation charts in the last months (Fig. 15), the generation from renewable sources in the island has overcome that percentage in many occasions, with a reported average of 80 % in local communication means. Short ago it reached a 100 % penetration (9th August 2015) during midday.

Gorona del Viento S.A. receives a remuneration regulated by a Governmental Order (IET/1711/2013 of 23 September). The remuneration scheme is similar to those applied in other islands' energy production plants, known as "Ordinary regime" facilities. The system sets the method for the calculation of fixed and variable costs of the installation and generation.

The fixed costs are those derived from the investment and the operation and maintenance of installations. The variable costs derive from the energy generation (also as operation and maintenance costs) and a concept known as "guarantee of available power".

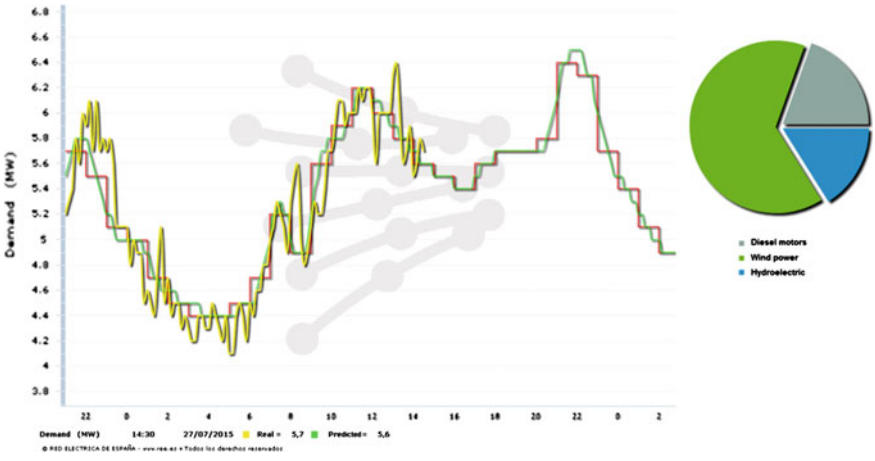


Fig. 15 Red Electrica de España website

The operation and maintenance fixed costs are mainly related to the human resources, compulsory insurances and water losses due to evaporation. This value is approximately 20,000 €/MW yearly. In addition, the remuneration on the investment derives from applying a rate of remuneration (calculated from the average yield of the Secondary market of Spanish Treasury Securities) to the calculated net annual investment.

The variable cost related to the generation corresponds to the variable costs of operation and maintenance times the generated MWh at the connection point (established as 15.57 €/MWh).

For the variable costs related to the estimation of the annually guarantee of available power (€/MW), costs incurred the previous year are considered, with values such as the energy generated from the wind and hydraulic subsystems and the proportion of wind generation to cover the direct demand and the needs of pumping are needed, as well as costs due to water evaporation and insurances.

7 Benefits and Further Initiatives

The new hybrid system in El Hierro is expected to dramatically reduce the island's oil use. If an annual renewable energy penetration of 68.40 % is achieved El Hierro would save 7364 m^3 of diesel and would decrease in CO_2 emissions of 20.91 Gg (Bueno and Carta 2005b).

Among the benefits that the new power station has brought to the island, the interest of the media has been particularly noticeable, even before its inauguration. International communication media and social networks have echoed the project. This coverage helps in the promotion of the island and involves an increasing interest in visiting El Hierro.

The international interest in renewable energy systems has allowed that Gorona del Viento participates in numerous forums (Energy Security Summit held in Prague in 2011, meeting of the Biosphere Reserve Islands in Korea, 2012 or a Summit by IRENA Renewable Islands in Malta, 2012) showing El Hierro as a place with technological innovation (Gorona del Viento S.A. 2014).

Finally, various training projects ranging from internships to development work have taken place in Gorona del Viento, with more than fifty students during the project implementation.

8 Conclusions

The wind-pumped hydro power plant described in this chapter is a good example of how the inclusion of renewable energy-based systems is achievable even in small grid contexts, where the stability of the system is a great concern. Such system brings economic and social benefits, particularly in small isolated electric grids.

In this particular case, the involvement and support of governmental organizations through investment and remuneration schemes are key for the promotion and development of the project. After the first stage of commitment from several organizations the use of simulation and optimisation algorithms as tool to help on the decision-making process are fundamental for the design and selection of power plants and their components, as it has been described in the chapter.

References

John S. Anagnostopoulos; Dimitris E. Papantonis. Pumping station design for a pumped-storage wind-hydro power plant (2007).

Arabkoohsar, A.; Machado, L.; Farzaneh-Gord, M. and Koury R.N.N. Thermo-economic analysis and sizing of a PV plant equipped with a compressed air energy storage system. Renewable Energy. Volume 83, 2015, Pages 491–509.

Benitez, L. E.; Benitez, P.C.; van Kootena, G.C. The economics of wind power with energy storage. Energy Economics Volume 30, Issue 4, July 2008, Pages 1973–1989.

Bhuiyan, F.A. Energy storage technologies for grid-connected and off-grid power system applications. 2012 IEEE Electrical Power and Energy Conference. Pages 303–310.

Bueno and Carta. Technical–economic analysis of wind-powered pumped hydrostorage systems. Part I: model development. Solar Energy 78 (2005a) 382–395

Bueno and Carta. Technical–economic analysis of wind-powered pumped hydrostorage systems. Part II: model application to the island of El Hierro. Solar Energy 78 (2005b) 396–405

Bueno and Carta. Wind powered pumped hydro storage systems, a means of increasing the penetration of renewable energy in the Canary Islands. Renewable and Sustainable Energy Reviews 10 (2006) 312–340.

Cattaneo, Agapi Fylaktou. Renewable Energy Sources in Small Communities. PLEA 2008– 25th Conference on Passive and Low Energy Architecture, Dublin, 22nd to 24th October 2008.

Gobierno de Canarias. Consejería de Empleo, Industria y Comercio. Anuario Energético de Canarias 2013, Publication year: 2014

Gorona del Viento S.A. Press dossier for the inauguration of the plant. June 2014.

Indradop, Mitra. "A renewable island life Electricity from renewables on small islands." Refocus 7:6, 2006.

Red Electrica De España S.A.U. El Sistema Eléctrico Español. Avance del Informe 2014.

Saastamoinen, Mika. Case Study 18 Samsø - renewable energy island programme Denmark. Samsoe Energy Academy. (NCRC). March, 2009.

Part V
Energy Distribution

Multi-objective Optimisation Incorporating Life Cycle Assessment. A Case Study of Biofuels Supply Chain Design

María Augusta Páez, Fernando D. Mele and Gonzalo Guillén-Gosálbez

Abstract This chapter describes how to optimise energy systems considering their economic and environmental performance simultaneously. To this end, we follow an approach that combines life cycle assessment with multi-objective optimisation. We illustrate how to apply such a framework to the strategic design and planning of biofuel supply chains. The problem is formulated in mathematical terms as a multi-objective mixed-integer linear programme. The aim of the design/planning task is to maximise the net present value while the environmental impact is minimised simultaneously. Eco-indicator 99 is the life cycle assessment methodology incorporated in the model to quantify the environmental damage. The implementation of the algorithm in a case study based on the Argentine industry reveals the conflictive trade-off between economic and environmental objectives. The proposed framework provides valuable insight into the incidence of key operational features in the optimal biofuel supply chain network.

1 Introduction

A supply chain is defined as an assembled production process and distribution system attempting to material purchase, transformation of raw materials into final products, and subsequent delivery of these to retailers. The concept of supply chain

M.A. Páez (✉) · G. Guillén-Gosálbez
School of Chemical Engineering and Analytical Science,
The University of Manchester, Manchester M13 9PL, UK
e-mail: maria.paez@postgrad.manchester.ac.uk

G. Guillén-Gosálbez
e-mail: gonzalo.guillen-gosalbez@manchester.ac.uk

F.D. Mele
Facultad de Ciencias Exactas y Tecnología, Universidad Nacional de Tucumán,
FBQF—Ayacucho 471 Rectorado—Ayacucho 491, San Miguel de Tucumán,
Tucumán 4000, Argentina
e-mail: fmele@herrera.unt.edu.ar

M. Martín (ed.), *Alternative Energy Sources and Technologies*,
DOI 10.1007/978-3-319-28752-2_17

initially appeared in the economics literature motivated by the embracement of labour policies and organisational control processes in the majority of corporations (Papageorgiou 2009; Sarkis et al. 2011).

Supply chain management deals, among other aspects, with the optimisation of two main activities: (1) production planning and inventory control, and (2) distribution and logistics. The first stage comprises the design and supervision of manufacturing processes including the storage strategies and procedures of raw materials purchase whilst the second one is related to the regulation of inventory retrieval and final product distribution. An appropriate management of these stages steers to the achievement of required performance goals; therefore, managing supply chains in an optimal manner became a key issue for practical applications in administrative and engineering literature in the early twenty-first century (Beamon 1998).

1.1 Development of Supply Chain Optimisation

Apart from management concerns, the evaluation of process supply chains was introduced in the 1970s as a response of the increasing refinement of the industrial metabolism theory and the material balance utilisation for organisational and governmental decision-making (Sarkis et al. 2011). One of the first relevant research was developed by Stern et al. (1973) where a supply chain is assessed in terms of the cumulative costs throughout various processes from the raw material inputs to the semi-finished and final products. Later, the placement of the globalisation in the early twenty-first century promoted the evolution of the supply chain management establishing three key issues: design, planning and control. The growing number of decision variables, constraints, performance measures and uncertainties in demand caused the intensification of managing complexity and highlighted the potential contribution of optimisation to face these challenges (Abo-Hamad and Arisha 2011; Papageorgiou 2009).

In this context, supply chain design optimisation emerged to lead decision-makers towards the best alternative configuration to satisfy the demand by handling the production, distribution and storage while the impact on the system performance is assessed. This task is fulfilled via the implementation of optimisation modelling techniques, which can be broadly categorised into mathematical programming and simulation-based. The prime models deal with unknown supply chain configuration using collective perception of operation dynamics, whereas simulation models perform under operational uncertainty for a fixed supply chain configuration. (Papageorgiou 2009). Furthermore, mathematical programming approaches tend to apply fairly simple capacity limitations in order to simplify the calculations during the optimisation step, while simulation approaches avoid strong simplifications to the extent possible (which results in more complex models that are hard to optimise). In essence, there is a clear trade-off between model accuracy and mathematical

complexity, and generally speaking each method lies at the extreme of the potential spectrum of approaches. A wide range of frameworks for the classification of optimisation techniques and the corresponding solving methods has been proposed in the literature. A comprehensive overview is provided in Fig. 1.

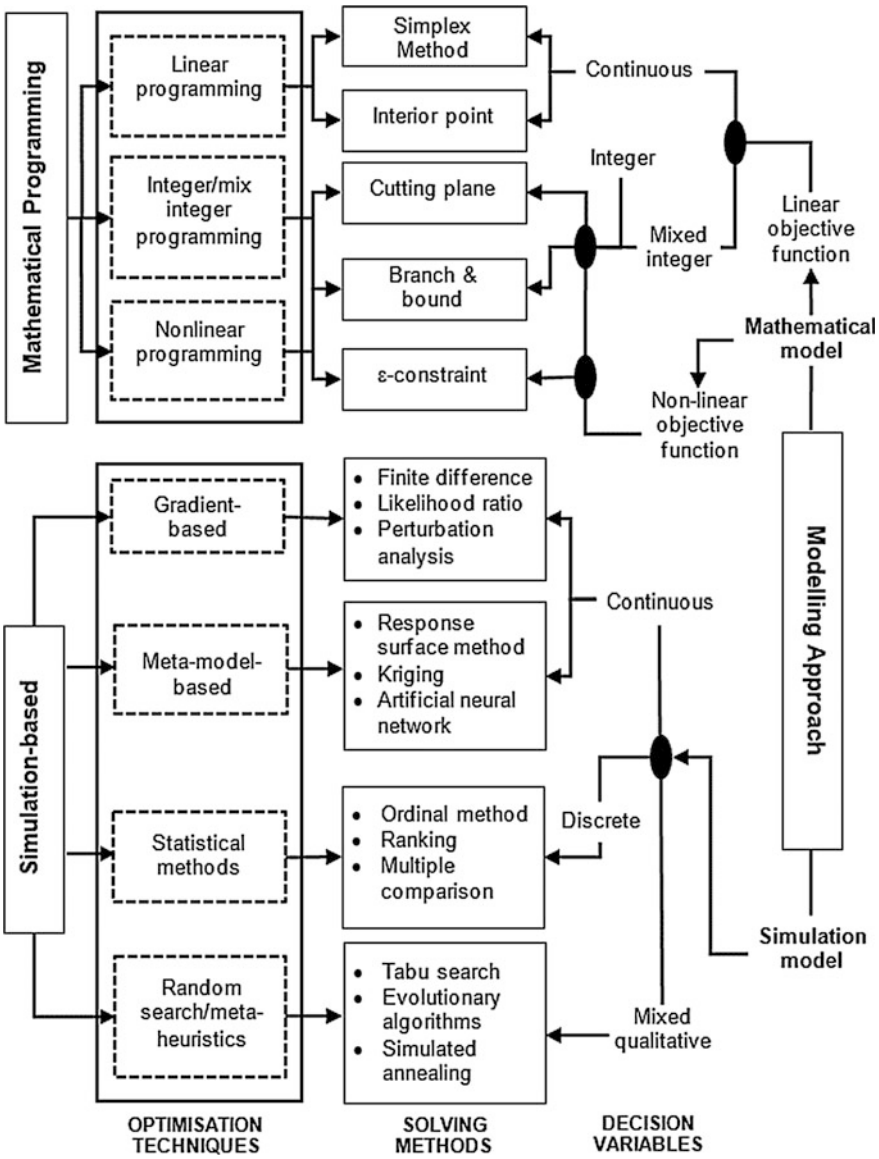


Fig. 1 Modelling techniques and solving methods for supply chain optimisation (adapted from Abo-Hamad and Arisha 2011)

The effectiveness of the optimisation is therefore determined by the modelling approach applied as well as the algorithm selected for optimising the decision variables. Another key point is the adequate selection of performance measures. Single-objective models have been the preferred approach during many years. The analysis of over 150 scientific articles evidenced that this approach has evolved throughout the time subjected to the availability of decision and information technologies. In fact, a chronological differentiation of research features can be established between 1990 and 1995 (Beamon 1998).

Prior to 1990, corporate taxes, tariffs and duties were the meaningful issues, while the production–distribution model was the preferred technology. The predominant objective was the maximisation of after tax profit. During the period 1990–1995, the variability of exchange rates was the main concern and the stochastic programming combined with valuation models was the feasible technology. The variable and overall cost minimisation became then the primary objective. From 1995 upwards, the attention on transfer price and supplier selection decisions gained prominence (Meixell and Gargeya 2005).

The raising number of metrics for performance evaluation based on economic flows and customer responsiveness led to the need to resort to multi-objective supply chain optimisations (Papageorgiou 2009). In parallel, the first MILP approaches were soon improved to model the supply chain in a more realistic manner, thereby giving rise to MINLP models, which were applied to the most dominant process industries. For instance, Neiro and Pinto (2004) formulated an MINLP model for oil and gas supply chain planning in Brazil including the crude oil provision and product distribution in four refineries, five terminals and pipeline channels.

Despite these successful improvements achieved within the framework of multi-objective optimisation, cost minimisation within a single-objective perspective still remains as the dominant approach in many sectors. Seuring (2013) studied 308 papers published in the last 15 years before 2013 and concluded that only 9 % consider more than one objective in the design and planning of supply chains.

1.2 Green Supply Chain Optimisation

The inclusion of environmental aspects in supply chain management led to the concept of Green Supply Chain Management (GrSCM). This integration was a natural reaction to the growing public pressure and the increasing environmental regulatory policies promoted by the progressive deterioration of the environment. Initially, the concept was envisaged as an extended system, in which the energy conservation and waste management were ensured at every stage of the supply chain. The performance was evaluated in terms of environmental criteria beyond traditional factors related to cost, functionality and manufacturability (Sarkis et al. 2011).

At the beginning, environmental issues were treated as constraints to be satisfied by the optimal solution of the economic optimisation problem. Few years later, the scope of the analysis was shifted to improve simultaneously the economic and environmental performance using multi-objective optimisation. In supply chain formulations, decisions to be made are modelled using both continuous and discrete variables, where the former denote mass and energy flows while the later represent facility location decisions, and selection of technologies and transportation modes. The models can be either linear or nonlinear, and may contain both continuous and discrete variables, thereby giving rise to MILP and MINLP formulations, respectively, that can be very hard to solve due to the combinatorial nature of the problem (Nikolopoulou and Ierapetritou 2012). One of the first investigations on how to include environmental concerns in supply chain management was carried out by Zhou et al. (2000) who integrated sustainability aspects in a continuous process supply chain via the implementation of goal programming applied to an MINLP model.

A key issue in the optimisation of environmental aspects concerns the selection of an appropriate environmental indicator. From the early 1970s until mid-1980s risk management perception diverted attention to waste and energy minimisation. Afterwards, up to mid-1990s, pollution prevention gained notoriety and minimisation of material usage and utility consumption were the main objectives sought. At the end of 1990s, the field of industrial ecology gained popularity and focused on maximising profit while ensuring environmental quality by adopting a life cycle perspective (Sarkis et al. 2011; Nikolopoulou and Ierapetritou 2012). It is important to highlight that despite the efforts made so far, there is at present a lack of agreement in the scientific literature on which universal indicator should be used for optimising sustainable supply chains.

1.2.1 Incorporation of Life Cycle Assessment (LCA)

In the last decade, a remarkable growth of scientific publications dealing with GrSC optimisation has been evidenced, particularly in the process systems engineering literature. The introduction of environmental aspects into optimisation models has been accomplished using systematic methodologies. Minimum environmental impact (MEI), the waste reduction algorithm (WAR), environmental fate and risk assessment toll (EFRAT), the atmospheric hazard index (AHI), and life cycle assessment (LCA) are among the most common methods. Regardless of the approach followed, what soon became clear in the optimisation of more sustainable processes was that the impact should be assessed considering all the echelons in the product supply chain. Hence, according to Eskandarpour et al. (2015), the majority of the scientific research between 1995 and 2014 integrated the principles of LCA into their supply chain optimisation models.

The combined use of LCA and multi-objective optimization was first introduced by Azapagic and Clift (1999). This framework was later on applied to supply chain management in the context of process systems engineering in the seminar works by

Mele et al. (2005) and Hugo et al. (2005). These first works in the area were followed by many others, which explored new applications and proposed some methodological advances. These include contributions to simplify the calculations by removing redundant LCA metrics from the problem (Kostin et al. 2012), the treatment of uncertainties Sabio et al. (2014), the use of eco-costs (Michelsen et al. 2006), and the inclusion of social metrics (Dehghanian and Mansour 2009), to name some.

LCA is a broad framework for assessing in a holistic and quantitative manner the environmental burdens of a product supply chain embracing extraction of raw materials, manufacturing, distribution, recycling and disposal of products. Four steps are involved in the methodology as defined by the ISO 14040 and 14044 standards. In the goal and scope step, the objectives, functional units and boundaries of the system are established whilst inlet and outlet flows, emissions and land use are quantified during the inventory analysis. Then, the environmental impacts resulted from the emissions are estimated in the impact assessment step and the outputs are analysed in the interpretation step. (Dreyer et al. 2003) The information generated in these steps is then fed into an optimisation model that optimises the values of the decision variables taking into account the effect that those variables will have on the environmental performance of the system (Pieragostini et al. 2012).

In terms of applications, we can find works in the areas of steam power generation systems (Luo et al. 2013), design of chemical supply chains (Guillén-Gosálbez et al. 2006), design of hydrogen supply chains (Almansoori and Shah 2006), strategic planning of biomass-to-energy process (Elia et al. 2011) and ethanol production (Giarola et al. 2012), among others. The main differences across these studies rely on the goal and scope definition and the impact assessment during the implementation of LCA. The inventory analysis and the interpretation stages are considered as part of the formulation and solution of the multi-objective optimisation problem. Hence, three ways for establishing the boundaries of the system have been identified: cradle-to-gave (CTG), cradle-to-gate (CTG) and gate-to-gate (GTG). In the first one the scope envisages the conventional operations in the supply chain from raw material purchase to product distribution but also embraces maintenance, disposal and recycling. Particularly, in the fuel supply chain this category is named well-to-wheel while in the biomass supply chain is called field to wheel. Cellulosic ethanol process (You et al. 2012) and sugar cane ethanol production (Kim et al. 2011) are included in this group. In the CTG case, the scope is limited to the factory gate and is commonly adopted by multiple consumers companies. Recent examples are Ruiz-Femenia et al. (2013) and Yue et al. (2014). Lastly, the exclusion of raw material extraction is evidenced in the GTG scope, which is usually applied to end-of-life products supply chains.

On the other hand, environmental assessment in LCA covers midpoint and endpoint indicators. Midpoint indicators reflect a point in the cause–effect chain (environmental mechanism) of a particular impact category, while endpoint metrics assess human health and ecosystem and resources impacts at the endpoint, which are originated by climate change, ozone depletion, as well as other categories quantified via midpoint indicators. Eskandarpour et al. (2015) pointed out that both kinds of indicators are used in the formulation of GrSC optimisation models, being

the Eco-indicator 99 (EI99), IMPACT 2002, CML92 and ReCiPe the most widely used ones.

Eco-indicator 99 is an endpoint method whose distinctive feature is to give a single score for environmental impact, instead of a bunch of per category values. It embraces the identification of concerned areas and the determination of the causes for each damage category. Eleven impacts are grouped into three damage categories (human health, ecosystem quality and resources) and finally aggregated into a single score. Impact assessment, as showed in Fig. 2, is carried out in two main stages referred to damage modelling and normalisation (involving weighting). The normalisation is based on European data related to air, water and soil emissions. The weighting procedure is performed through an expert panel approach to reflect the point of view of the society about the importance of potential impacts (Dreyer et al. 2003).

Chronologically, the implementation of the Eco-indicator 99 can be traced since the beginning of the twenty-first century. At present, this method has been replaced as main single-score method by ReCiPe. An increasing tendency to adopt a partial LCA analysis has taken place, in which only the relevant damage categories are contemplated depending on the industrial application. For instance, Hugo et al. (2005) investigated hydrogen network optimisation accounting for the trade-off between investment and greenhouse gas emissions.

1.2.2 Biofuel Green Supply Chains

The majority of studies developed in green optimisation are motivated by real-life concerns. Models are built in a specific industrial context because the formulation of

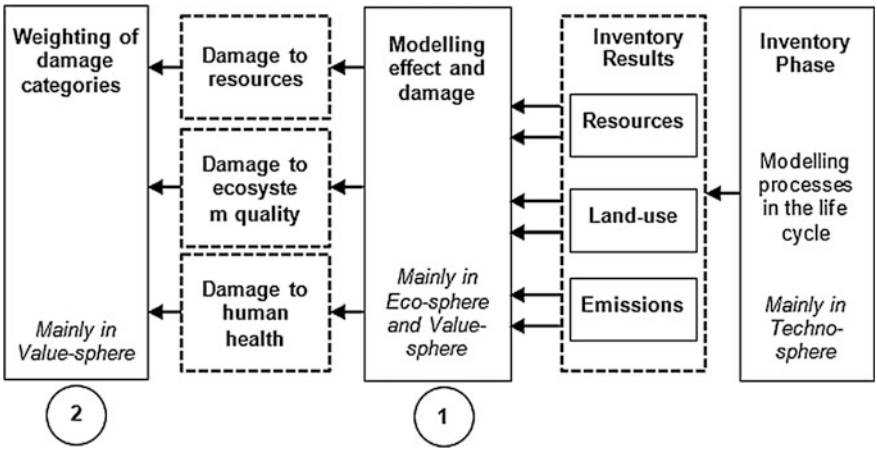


Fig. 2 Environmental impact assessment in Eco-indicator 99 methodology (adapted from Goedkoop and Spriensma 2001)

environmental aspects requires the description of a particular case. The design of supply chains in some prominent industrial sectors such as tires, steel, aluminium, aeronautics and automotive still relies on economic factors. Nevertheless, process industries like biomass-to-energy and waste management have demonstrated superior ripeness on this topic and become an emerging research field (Nikolopoulou and Ierapetritou 2012).

Particularly, in the biomass-to-energy branch, ethanol from sugar cane has deserved special attention, as it is seen as a promising green alternative to replace fossil fuels. Brandenburg et al. (2014) stated that beyond agriculture, electronics, food and beverages, biofuels contribute with 5 % of the papers in GrSC optimisation published until 2012. In a seminar paper Zamboni et al. (2009) introduced a multi-objective optimisation model to optimise a bioethanol from corn production supply chain located in Northern Italy. The simultaneous minimisation of operating costs and greenhouse gas emissions is attempted through the evaluation of biomass cultivation practice and its geographical dependence in a spatially explicit approach. Later on, Corsano et al. (2011) developed an MINLP optimisation algorithm for bioethanol from sugar cane considering residual recycle and derivative production in the environmental impact assessment. The analysis relies on disposal cost parameters related to molasses and yeast generation.

Overall, the design of supply chains in chemical engineering can be naturally formulated as a multi-objective optimisation problem where the economic and environmental performances of the network are the objectives to be optimised. Hence, most of these studies end in the production of a Pareto optimal curve to portray the trade-off between economic and environmental aspects. This set of solutions are non-dominated, thus no solution is able to increase a performance measure without decreasing at least one of the other objectives at the same time. (Pinto-Varela et al. 2011; Brandenburg et al. 2014).

2 Problem Statement

To illustrate the capabilities of multi-objective optimisation combined with LCA as applied to the design of biofuels supply chains, we consider the network depicted in Fig. 3. Three echelons are involved in the configuration: the production level including raw material suppliers and production technologies, the storage stage with the corresponding facilities and the final destination markets. The structure of such a supply chain is described next before presenting the associated mathematical formulation.

Production technologies: Bioethanol and sugar are the main final products of the process. Note that the by-products of sugar production can also be converted into bioethanol by fermentation. The model contemplates five manufacturing technologies (T), two for sugar and three for bioethanol, as schematised in Fig. 4. In the sugar case, the technologies differ in the intermediate product generated, either

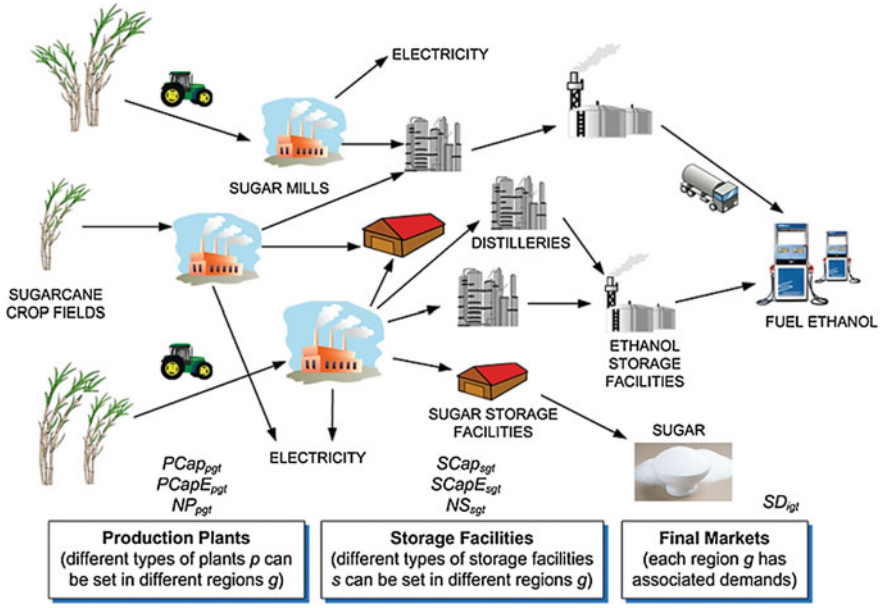


Fig. 3 Configuration of the bioethanol from sugar cane supply chain (Mele et al. 2011)

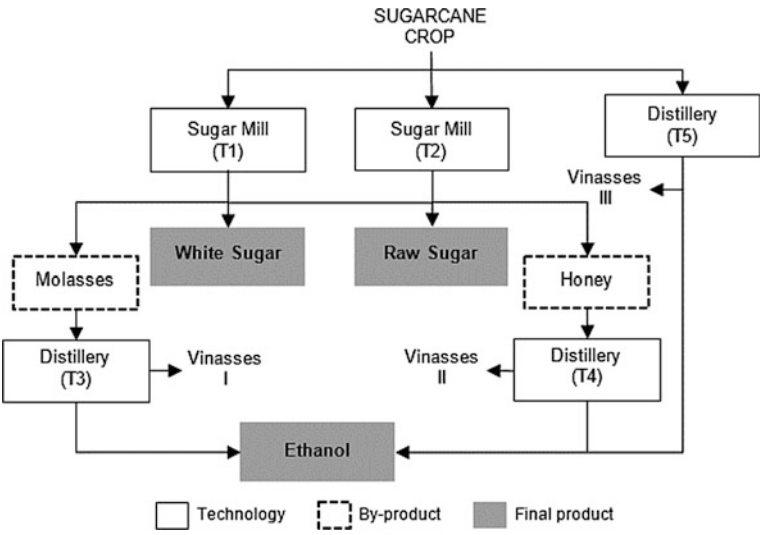


Fig. 4 Production technologies in the model of bioethanol supply chain design

molasses or honey, which will be fermented to produce ethanol. Similarly, with regards to ethanol production, there are several technologies available, each of which generates different by-products and wastes including vinasses type I, II or III.

Transportation: The final products are stored for posterior delivery to destination markets. Warehouses for liquid products (S1) and solid materials (S2) are accounted for. The storage facilities of the GrSC can be allocated in any region and are connected to each other through transport means. These include heavy trucks for sugar cane, medium-sized trucks for sugar and tank trucks for bioethanol.

The general problem we aim to solve is as follows. We are given a set of technologies to produce, store and deliver ethanol to the final customers. We are also provided with economic and environmental data associated with these technologies along with a given ethanol and sugar demand that we aim to fulfil. Hence, for a fixed evaluation time, financial parameters are considered including tax rate, salvage rate, landfill tax and interest rate. Operational and economic inputs are also given, including forecast for products demand, plant capacities and storage holding period, cost parameters for production technologies and storage facilities, fixed and upper limits for capital investment and product prices, environmental burdens of sugar and ethanol production related to manufacture technology and transport logistics (capacities, distance availability between sub-regions and costs).

The goal of the analysis is to determine the optimal network topology, including the technologies selected, their expansion over time and the associated planning decisions (e.g. transport flows, production rates, etc.) that optimise simultaneously the economic and environmental performance of the supply chain.

2.1 Mathematical Formulation

The design of a bioethanol supply chain can be formulated as a bi-criteria mixed-integer programming (MILP) model expressed in compact form as given:

$$\begin{array}{ll} \min & \left\{-\mathrm{NPV}(x,y,N);\ \mathrm{GED}_{(b)}(x,y,N)\right\} \\ \text{s.t.} & h(x,y,N)=0 \\ & g(x,y,N)\leq 0 \\ & x\in\Re, y\in\{0,1\},\ N\in Z^{+} \end{array} \tag{M1}$$

The formulation comprises three types of variables denoted as x, y and N. The first are continuous variables related to financial performance as well as production flows of feedstock, waste and final products. The second ones are binary variables used for strategic decisions concerning geographical location of production plants and warehouses, technology selection for sugar cane to be converted into biofuel and the transportation links with the corresponding modality. The last variables are integer and represent the number of manufacturing and storage facilities and transport systems.

The design envisages an overall time horizon of 4 years long, which is downsized into intervals to portray the periodical harvesting nature of the sugar cane crop. The MILP includes mass balance and capacity constraints that are modelled via equality and inequality constraints, denoted by $h(\cdot)$ and $g(\cdot)$, respectively.

The MILP contains as well two objective functions, the net present value (NPV), and the global environmental damage (GED). As will be explained later in the chapter, the latter criterion can be modelled using different indicators.

Mass balance equations: The global material balance given in Eq. (1) is the basic constraint that every potential configuration of the supply chain must satisfy. Thus, for each region (g) and material (i), the opening inventory resulted from a previous time period (ST_{isgt-1}) plus the quantity produced (PT_{igt}), the raw material purchased (PU_{igt}) and the inlet rate delivered from external facilities ($Q_{ilg'gt}$) must be equal to the final inventory (ST_{isgt}), plus the quantity distributed to retailers, the outlet flow to external facilities ($Q_{ilgg't}$) and the quantity of waste generated (W_{igt}).

$$\sum_{s\in IS(i,s)} ST_{isgt-1} + PT_{igt} + PU_{igt} + \sum_{l\in IL(i,l)} \sum_{g'\neq g} Q_{il\,g'\,gt} = \sum_{s\in IS(i,s)} ST_{isgt} + DTS_{igt} + \sum_{l\in IL(i,l)} \sum_{g\neq g'} Q_{il\,g\,g't} + W_{igt} \quad \forall i,g,t \tag{1}$$

The subscripts IS (i,s) and $IL(i,l)$ are set of ordered pairs, which associate the material (i) with the appropriate type of storage facility (s) and transportation modality (l) respectively. The overall production rate of each material is calculated as the contribution of the individual production rates (PE_{ipgt}) from the technologies (p) installed in the corresponding region:

$$PT_{igt} = \sum_{p} PE_{ipgt} \quad \forall i,g,t \tag{2}$$

The individual production rates resulted from each technology are defined through the production rate of the main product with the mass balance coefficient ρ_{pi} as expressed in Eq. (3). The subset IM (i,p) associates the main product (i) with the manufacturing technology (p).

$$PE_{ipgt} = \rho_{pi} PE_{i'pgt} \quad \forall i,p,g,t \quad \forall i' \in IM(i,p) \tag{3}$$

On the other hand, the sugar cane purchasing must be restricted by the capacity of the existing sugar cane plantations. Thus, the sugar cane delivering rate from region (g) in time interval (t) is:

$$PU_{igt} \leq CapCrop_{gt} \quad i = \text{sugarcane}, \quad \forall g,t \tag{4}$$

Similarly, the admissible inventory (ST_{isgt}) of product (i) during the time interval (t) is restricted by the warehouse capacity ($SCap_{sgt}$):

$$\sum_{i\in IS(i,s)} ST_{isgt} \leq SCap_{sgt} \quad \forall s,g,t \tag{5}$$

An average inventory (AIL_{igt}) can be established in a steady-state operation considering the amount distributed to the customers and the storage holding period (σ):

$$AIL_{igt} = \sigma DTS_{igt} \quad \forall i, g, t \tag{6}$$

The storage period denotes the number of days on average that a product will be stored in a certain facility. This concept is analogous to the turnover ratio of a warehouse, which is an indicator of the stock replacement frequency. Hence, the model manages supply and demand fluctuations by assuming that the overall storage capacity must be at least twice the average inventory level of product (i) in region (g):

$$2AIL_{igt} \leq \sum_{s \in IS(i,s)} SCap_{sgt} \quad \forall i, g, t \tag{7}$$

Lastly, the outlet flow of products delivered to the final markets must meet at least the forecasted demand (SD_{igt}):

$$DTS_{igt} \leq SD_{igt} \quad \forall i, g, t \tag{8}$$

Additionally, the binary variable $X_{lgg't}$ reflects the presence of a transportation link between regions (g) and (g'). A value of 1 determines the existence of the link and a value of 0 the converse. This variable is used in the following constraint that establishes that a region cannot purchase and export products (i) simultaneously in a given time period:

$$X_{lg\,g't} + X_{lg'\,gt} = 1 \quad \forall l, t, g, g'(g' \neq g) \tag{9}$$

Capacity constraints: The production rate of a given technology (p) in a region (g) must lie within upper ($PCap_{pgt}$) and lower capacity limits. The last are given by the minimum percentage of the available technology that is desirable for utilisation (τ):

$$\tau PCap_{pgt} \leq PE_{ipgt} \leq PCap_{pgt} \quad \forall i, p, g, t \tag{10}$$

In order to account for future capacity expansions, the existing capacity of technology (p) in any period (t) is defined as the sum of the capacity at the end of a prior period and the likely expansion executed in the present one ($PCapE_{pgt}$):

$$PCap_{pgt} = PCap_{pgt-1} + PCapE_{pgt} \quad \forall p, g, t \tag{11}$$

Capacity expansions have been also restricted by boundaries that take into account the integer variable NP_{pgt}, which represents the number of production plants allocated in region (g) at time period (t):

$$\underline{PCap_p} NP_{pgt} \leq PCapE_{pgt} \leq \overline{PCap_p} NP_{pgt} \quad \forall p, g, t \tag{12}$$

Similarly, as with production technologies, the existing capacity of a storage facility is defined in terms of the capacity at the end of a prior period and the potential expansion in the current one ($SCapE_{sgt}$), as given in Eq. (13). Besides, storage expansions are also limited by boundaries with the assistance of the integer variable NS_{sgt} as described in Eq. (14).

$$SCap_{pgt} = SCap_{pgt-1} + SCapE_{pgt} \quad \forall s, g, t \tag{13}$$

$$\underline{SCap_p} NS_{sgt} \leq SCapE_{sgt} \leq \overline{SCap_s} NS_{sgt} \quad \forall s, g, t \tag{14}$$

2.2 Estimation of Economic and Environmental Objectives

Economic objective: The net present value (NPV) quantifies the economic performance and is calculated through the discounted annual cash flows (CF_t) for a defined time interval (t):

$$NPV = \sum_t \frac{CF_t}{(1+ir)^{t-1}} \tag{15}$$

The cash flow is calculated as the difference between net earnings NE_t and the fraction of the total depreciable capital investment ($FTDC_t$):

$$CF_t = NE_t - FTDC_t \quad t = 1, \ldots, T-1 \tag{16}$$

The cash flow corresponding to the last time period ($t = T$) is perceived as an element of the fixed capital investment (FCI). Such term named salvage value (sv) will be recovered at the end of the time horizon and varies depending on the kind of industry.

$$CF_t = NE_t - FTDC_t + svFCI \quad t = T \tag{17}$$

Net earnings after tax rate (φ) are determined by the difference between revenues (Rev_t) and expenses related to facility operating (FOC_t) and transportation (TOC_t) costs:

$$\text{NE}_t = (1-\varphi)(\text{Rev}_t - \text{FOC}_t - \text{TOC}_t) + \varphi \text{DEP}_t \quad \forall t \tag{18}$$

Accordingly, revenues are estimated from the sales of the final products, which involve the delivered flow and the corresponding price (PR_{igt}):

$$\text{Rev}_t = \sum_{i \in \text{SEP}(i)} \sum_{g} \text{DTS}_{igt} \text{PR}_{igt} \quad \forall t \tag{19}$$

Operating costs are given by the product between the production rate and average inventory levels with the corresponding unit production cost (UPC_{ipgt}) and unit storage cost (USC_{isgt}), respectively:

$$\text{FOC}_t = \sum_{i} \sum_{g} \sum_{p \in IM(i,p)} \text{UPC}_{ipgt} \text{PE}_{ipgt} + \sum_{i} \sum_{g} \sum_{s \in IS(i,s)} \text{USC}_{isgt} \text{AIL}_{igt} + \text{DC}_t \quad \forall t \tag{20}$$

Disposal cost (DC_t) as part of the operating costs, is a term related to the landfill tax (LT_{igt}) to be handled as a consequence of the amount of waste generated in the process (W_{igt}):

$$\text{DC}_t = \sum_{i} \sum_{g} W_{igt} \text{LT}_{igt} \quad \forall t \tag{21}$$

Transportation cost comprises fuel, labour, maintenance and general costs, as given in Eq. (22). Fuel cost is defined as a function of the fuel price and the usage rate expressed in Eq. (23). The usage is established for a certain transportation modality (FE_l) by means of the number of trips required per period of time ($Q_{ilgg't}$/ TCap_l) and the total distance travelled ($2\text{EL}_{gg'}$). Labour cost, as appeared in Eq. (24), is referred to the driver wage (DW_{lt}) as function of the total delivery time. Maintenance cost is calculated via Eq. (25) and envisages general expenses inherent to the transport system depending on the distance driven and the cost per distance unit (ME_l). The remaining general cost includes transportation insurance, license and registration and is estimated through Eq. (26)

$$\text{TOC}_t = \text{FC}_t + \text{LC}_t + \text{MC}_t + \text{GC}_t \quad \forall t \tag{22}$$

$$\text{FC}_t = \sum_{i \in IL(i,l)} \sum_{g} \sum_{g' \neq g} \sum_{l} \text{DW}_{lt} \left[\frac{2\text{EL}_{g'g} Q_{i\,\text{lg}'\,gt}}{\text{FE}_l \text{TCap}_l} \right] \text{FP}_{lt} \quad \forall t \tag{23}$$

$$\mathrm{LC}_t = \sum_{i \in IL(i,l)} \sum_{g} \sum_{g' \neq g} \sum_{l} \mathrm{DW}_{lt} \left[\frac{Q_{i\,\mathrm{lg}'\,gt}}{\mathrm{TCap}_l} \left(\frac{2\mathrm{EL}_{g'g}}{\mathrm{SP}_l} + \mathrm{LUT}_l \right) \right] \quad \forall t \tag{24}$$

$$\mathrm{MC}_t = \sum_{i \in IL(i,l)} \sum_{g} \sum_{g' \neq g} \sum_{l} \mathrm{ME}_l \frac{2\mathrm{EL}_{g'g} Q_{i\,\mathrm{lg}'\,gt}}{\mathrm{TCap}_l} \quad \forall t \tag{25}$$

$$\mathrm{GC}_t = \sum_{l} \sum_{t' < t} \mathrm{GE}_{lt} \mathrm{NT}_{lt} \quad \forall t \tag{26}$$

Regarding capital cost, the depreciation is initially calculated implementing the straight-line method:

$$\mathrm{DEP}_t = \frac{(1 - \mathrm{sv})\mathrm{FCI}}{T} \quad \forall t \tag{27}$$

The total fixed cost investment (FCI) embodies the capacity expansions achieved in production plants and storage facilities throughout the entire time horizon. Parameters $\alpha^{\mathrm{Pr}}_{pgt}$, $\beta^{\mathrm{Pr}}_{pgt}$ and $\alpha^{\mathrm{St}}_{sgt}$, $\beta^{\mathrm{St}}_{sgt}$ are fixed and variable investment coefficients. In addition, the investment associated to the establishment of a particular transportation system (TMC_{lt}) is also accounted for in the model.

$$\begin{aligned} \mathrm{FCI} = & \sum_{p} \sum_{g} \sum_{t} \left(\alpha^{\mathrm{Pr}}_{pgt} \mathrm{NP}_{pgt} + \beta^{\mathrm{Pr}}_{pgt} \mathrm{PCapE}_{pgt} \right) \\ & + \sum_{s} \sum_{g} \sum_{t} \left(\alpha^{St}_{sgt} \mathrm{NS}_{sgt} + \beta^{St}_{sgt} \mathrm{SCapE}_{sgt} \right) + \sum_{l} \sum_{t} \mathrm{NT}_{lt} \mathrm{TMC}_{lt} \end{aligned} \tag{28}$$

The requirement of trucks is calculated as given in Eq. (29) considering material flows between regions, availability of transportation modes (avl_l) and operational issues related to distance covered, average speed and loading/unloading time.

$$\sum_{t \leq T} \mathrm{NT}_{lt} = \sum_{i \in IL(i,l)} \sum_{g} \sum_{g' \neq g} \sum_{t} \frac{Q_{i\,\mathrm{lg}\,g't}}{\mathrm{avl}_l \mathrm{TCap}_l} \left(\frac{2\mathrm{EL}_{gg'}}{\mathrm{SP}_l} + \mathrm{LUT}_l \right) \quad \forall l \tag{29}$$

The formulation splits the payment of the total capital investment in equal amounts per time period (FTDC_t):

$$\mathrm{FTDC}_t = \frac{\mathrm{FCI}}{T} \quad \forall t \tag{30}$$

Environmental objective: The environmental impact is evaluated and introduced into the model using the LCA metric Eco-indicator 99, with the hierarchical weighting perspective being used to assess the relative importance of the damages.

The calculation of this metric follows the main four LCA phases: goal and scope, life cycle inventory, damage assessment and interpretation.

1. The goal of the analysis is to determine the impact associated with a given amount of sugar and ethanol, with the boundaries of the study covering all the life cycle stages from crop harvesting to the delivery of sugar and ethanol to the final customer.
2. The life cycle inventory is carried out next. To this end, all the potential emissions and resource extractions along the process are identified and quantified. These values are calculated at every echelon of the supply chain including sugar cane purchases (PU_{igt}), production rates (PE_{ipgt}) and transportation flows ($Q_{ilgg't}$).
3. Later, the damage is estimated by matching the life cycle inventory entry with the respective damage factor. The damage factors cover 11 impacts grouped in the three main damage categories (human health, ecosystem quality and resources). Thus, for every environmental impact category (b), a set of damages are calculated using SimaPro 6.0 LCA software. All the stages in the life cycle of the sugar and bioethanol are considered, that is crop harvesting (EPU_b), production (EPE_{bp}) and transportation (EQ_{bl}), respectively.
4. The interpretation phase corresponds to the post-optimal analysis of the Pareto solutions of the multi-objective problem.

More precisely, the environmental impact of the sugar cane plantations is determined as expressed in Eq. (31). In the case of the manufacturing process, the impact is estimated through Eq. (32). In the final echelon, the environmental impact is established by Eq. (33) considering the transportation distances between regions.

$$\mathrm{IPU}_b = \mathrm{EPU}_b \sum_g \sum_t \mathrm{PU}_{igt} \quad \forall t \quad i = \text{sugarcane} \tag{31}$$

$$\mathrm{IPE}_b = \sum_{i \in MP(l)} \sum_p \sum_g \sum_t \mathrm{EPE}_{bp}\mathrm{PE}_{ipgt} \quad \forall b \tag{32}$$

$$\mathrm{IQ}_b = \sum_{i \in IL(i,l)} \sum_l \sum_g \sum_{g' \neq g} \sum_t \mathrm{EQ}_b \mathrm{EL}_{gg'} Q_{i\lg g't} \quad \forall b \tag{33}$$

The global environmental damages (GED_b) per category, which are introduced in the final model are estimated as follows:

$$\mathrm{GED}_b = \mathrm{IPU}_b + \mathrm{IPE}_b + \mathrm{IQ}_b \quad \forall b \tag{34}$$

2.3 Solution Method

The solution of the bi-criteria MILP model defined above is given by a set of Pareto options, each achieving a unique combination of objectives values, which portray

the trade-off between the conflicting objectives evaluated. Note that we can choose different damages to be optimised. In the more general case, we might be interested in optimising several damage categories simultaneously. This might lead to complex multi-objective models containing many objectives that are difficult to handle. To simplify the approach, we considered that there is a single environmental objective that must be optimised against the economic performance.

Hence, the Pareto solutions of the bi-criteria MILP are generated through the ε-constraint method. This technique divides the original problem into auxiliary single-objective optimisation frameworks. At the beginning, upper and lower boundaries for the epsilon coefficient are determined through the optimisation of two separate scalar objectives.

$$\begin{array}{ll} (\bar{x},\bar{y},\bar{N}) & \min\left\{\mathrm{GED}_{(b)}(x,y,N)\right\} \\ s.t. & h(x,y)=0 \\ & g(x,y)\le 0 \\ & x\in\Re, y\in\{0,1\},\ N\in Z^{+} \\ & \underline{\varepsilon}=\mathrm{GED}(\bar{x},\bar{y},\bar{N}) \end{array} \tag{M2a}$$

$$\begin{array}{ll} \left(\hat{x},\hat{y},\hat{N}\right) & \min\{-\mathrm{NPV}(x,y,N)\} \\ s.t. & h(x,y)=0 \\ & g(x,y)\le 0 \\ & x\in\Re, y\in\{0,1\},\ N\in Z^{+} \\ & \bar{\varepsilon}=\mathrm{GED}\left(\hat{x},\hat{y},\hat{N}\right) \end{array} \tag{M2b}$$

Finally, the economic objective is maximised and the environmental impact is transferred to an auxiliary constraint, such that the epsilon coefficient falls within the limits previously established. The single-objective problem is solved for different values of the epsilon parameters, thereby generating in each run a different point. These points are finally filtered to discard unfeasible and repeated solutions.

$$\begin{array}{ll} \min & \{-\mathrm{NPV}(x,y,N)\} \\ s.t. & h(x,y)=0 \\ & g(x,y)\le 0 \\ & \mathrm{GED}(x,y,N)\le\varepsilon \\ & \underline{\varepsilon}\le\varepsilon\le\bar{\varepsilon} \\ & x\in\Re, y\in\{0,1\},\ N\in Z^{+} \end{array} \tag{M3}$$

3 Case Study

The design of a bioethanol supply chain in Argentina is the case study through which the capabilities of the proposed formulation are demonstrated. In the last years, Argentinian policies endeavour to improve the diversification of energy provision by laying down a minimum content of biofuel in gasoline and diesel.

Currently, ethanol is primarily produced in the northwest of the country using 15 sugar cane mills (Mele et al. 2009). The minimum and maximum capacities for production technologies T1 and T2 are set to 30 and 350 million tonnes per year. Meanwhile, for the case of T3, T4 and T5, these are set to 10 and 300, respectively. In the same way, the lower and upper limits of warehouses for liquids and solids are fixed to 2 and 200 billion of tonnes, respectively. Detailed inlet information of the model is described in Tables 1, 2, 3.

Table 1 Parameters for the estimation of capital cost in different production technologies and storage facilities

	Fixed investment coefficient ($)	Variable investment coefficient ($·yr./t)
T1	5,350,000	535
T2	5,350,000	535
T3	7,710,000	771
T4	7,710,000	771
T5	9,070,000	907
S1	1,220,000	122
S2	18,940,000	1,894

Table 2 Parameters for the estimation of capital and operating costs throughout different transport systems

	Heavy truck	Medium truck	Tanker truck
Capacity (tonne per trip)	30	25	20
Availability per day (h)	18	18	18
Investment for settlement of transport system ($)	30,000	30,000	30,000
Driver wage ($/h)	10	10	10
Fuel price ($/L)	0.85	0.85	0.85
General expenses daily ($)	8.22	8.22	8.22
Maintenance expenses × 1000 ($/km)	97.6	97.6	97.6

Table 3 General financial parameters and product prices

Sugar production cost ($/t)	265
Ethanol production cost ($/t)	317
White sugar selling price ($/t)	537
Raw sugar selling price ($/t)	375
Ethanol selling price ($/t)	860
Storage cost for all materials × 1000 ($·yr./t)	365
Tax rate	0.3
Salvage rate	0.2
Interest rate	0.03
Landfill tax	0.1

For the design to reflect the bioethanol production as realistic as possible, Argentina was discretised into 24 regions corresponding to the states of the country. Table 4 provides the forecasted demand of the main products associated to each region. Tucumán, Jujuy, Salta, Santa Fe and Misiones are the regions with available sugar cane capacity for periodical harvesting.

Moreover, Table 5 displays the environmental damage factors employed in the estimation of the Eco-indicator 99 (EI99) metric. These factors correspond to the hierarchical perspective of the LCA method and are contemplated in every stage of the supply chain, including sugar cane purchase, production and transportation.

The multi-objective optimisation model was implemented in the General Algebraic Modelling System (GAMS) tool version 24.1.1 and solved using the

Table 4 Sugar and ethanol demand per region in the Argentinian supply chain

Region ID.	Name	White sugar (tonne/year)	Raw sugar (ton/year)	Ethanol (ton/year)
G01	Buenos Aires City	76614.92	38307.46	84276.41
G02	Córdoba	84126.19	42063.09	92538.81
G03	Corrientes	25438.16	12719.08	27981.97
G04	Buenos Aires Prov.	379268.90	189634.45	417195.79
G05	La Rioja	9714.57	4857.29	10686.03
G06	Mendoza	43565.35	21782.67	47921.88
G07	Neuquén	13720.58	6860.29	15092.64
G08	Entre Ríos	31547.32	15773.66	34702.05
G09	Misiones	27140.71	13570.36	29854.78
G10	Chubut	11517.28	5758.64	12669.00
G11	Chaco	26439.66	13219.83	29083.63
G12	Santa Cruz	5708.56	2854.28	6279.42
G13	Salta	30746.12	15373.06	33820.73
G14	San Juan	17526.29	8763.14	19278.92
G15	San Luis	11016.52	5508.26	12118.18
G16	Tucumán	37155.73	18557.87	40871.31
G17	Jujuy	17125.69	8562.84	18838.26
G18	Santa Fe	81121.68	40560.84	89233.85
G19	La Pampa	8412.62	4206.31	9253.88
G20	Santiago del Estero	21732.60	10866.30	23905.86
G21	Catamarca	8612.92	4306.46	9474.21
G22	Río Negro	15022.53	7511.27	16524.79
G23	Formosa	13520.28	6760.14	14872.31
G24	Tierra del Fuego	3204.81	1602.40	3525.29

The nomenclature was assigned according to the geographical location in North-South direction

Table 5 Environmental damage factors used in EI99 calculation

Supply chain stage	Unit	Human health category	Ecosystem quality category	Resources category	EI 99 metric
Sugarcane harvesting and milling	kg sugarcane	0.0563	0.0169	0.0014	0.0746
Sugar production (technology T1)	kg sugar	0.0014	0.0001	0.0004	0.0019
Sugar production (technology T2)	kg sugar	0.0017	0.0001	0.0004	0.0023
Ethanol production (any technology)	kg ethanol	0.0171	0.0001	0.0009	0.0182
Sugarcane transportation	tkm	0.0046	0.0009	0.0059	0.0113
Sugar transportation	tkm	0.0022	0.0005	0.0052	0.0079
Ethanol transportation	tkm	0.0053	0.0012	0.0108	0.0173

The values are defined as ecopoints, which are inferred as one-thousandth of the annual environmental damage load per European inhabitant

CPLEX solver. An HP ENVY laptop with an Intel Core i7 processor and 8 GB of installed memory RAM was used in the calculations.

The Pareto curve referred to EI99 against NPV was generated as depicted in Fig. 5. The capital letters are used to characterise the resulting supply chain networks. Such analysis is based on the geographical location of the facilities and the associated technologies, number of production plants and ethanol delivery to neighbour regions. The position where the transition among production technologies takes place has been also highlighted in the figure.

The results reflect the existing conflict between environmental and economic performance in dealing with bioethanol from sugar cane production. Point A and F represent the extreme solutions from minimum environmental impact and maximum revenue respectively. At the economic optimum, bioethanol is produced from honey, which is the intermediate product of sugar production, but also directly from sugar cane. Production facilities are allocated in Tucumán, Jujuy and Salta—provinces with the largest raw material capacity—where bioethanol production exceeds the demand and is delivered to four neighbour regions. Revenues come from the sale of ethanol as well as white and raw sugar as by-products, accomplishing an NPV of $\$8.553 \times 10^8$; however, the environmental damage is at the highest level in 3.034×10^9 ecopoints. The resulting configuration is not centralised since white and raw sugar demand is satisfied for remote regions located in the South of Argentina. In contrast, no facilities settlement is involved when the optimal environmental performance is targeted.

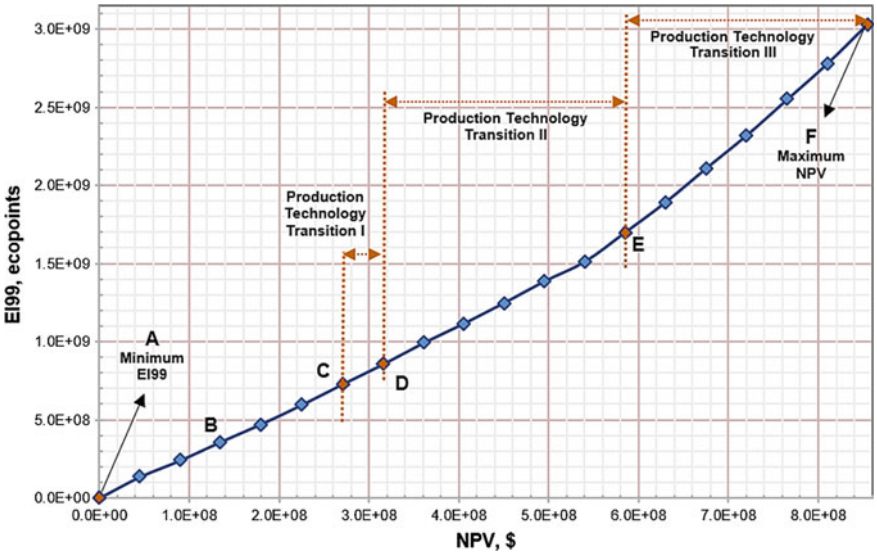


Fig. 5 Pareto multi-objective solutions for Eco-indicator 99 metric versus net present value and the corresponding sections of production technology transition

Along the Pareto curve, five well-defined intervals are identified during the characterisation of the supply chain configuration. Except for point A, sugar production dominates until point B, and honey is the only intermediate product, which subsequently generates bioethanol. Tucumán is the production centre, where bioethanol yield is for internal consumption exclusively. Moving towards point C, configuration features remain steady but Santa Fe, Salta and Jujuy contribute to raw material rate promoting bioethanol distribution to Córdoba, the second region with the highest demand on such product.

The first shift of production technology takes place with the appearance of molasses as intermediate product from sugar production. Even though Córdoba is still the only export destination, the increase of bioethanol production is enough to satisfy internal demand of Salta beyond Tucumán, which is no longer the manufacture centre. In point D, a new configuration strategy associated with the second technology transition is evidenced. Thus, molasses disappear and honey becomes the intermediate product again; up to three plants are allocated in Tucumán as the main production centre and the bioethanol yield increases. The demand of two main consumers is met, Córdoba and Santa Fe, with this last requiring a 62 % more bioethanol than Salta.

Lastly, the incorporation of the sugarcane-to-ethanol process is observed at point E. Bioethanol delivering rate rises, which is accompanied by a significant growth in revenues and deterioration of environmental performance. The demand of Buenos Aires Province, the largest consumer, is satisfied accounting for Jujuy, Tucumán and Salta as the raw material suppliers.

As stated before, the Eco-indicator 99 methodology accounts for the contribution of 11 impact categories, which are illustrated in more detail in Fig. 6 for several key Pareto points (i.e. extreme solutions and points where a technology transition takes place).

Regardless of the configuration of the supply chain and the production technology, we find that 2 out of the 11 categories are the main drivers for the environmental damage. Thus, land use contributes the most with 41 %, followed by respiratory effects caused by inorganics with 35 %. In addition, climate change category is the only category performing in the opposite direction. This fact is explained by the nature of sugar cane cultivation and growing, which are characterised by carbon dioxide trapping for photosynthesis purposes resulting in emission credits.

Even though the contribution of each category remains the same along the Pareto, the global environmental damage is compromised as the technology transition points get closer to the optimal economic performance. Thus, at point C the land-use effect decreases at 65 %, whereas at point E it is only 34 % lower. The same behaviour can be evidenced throughout all categories.

Complementary, Fig. 7 shows the contribution of different stages to the global environmental impact produced by the bioethanol supply chain. We consider only the case of net present value maximisation, since in the minimum environmental impact solution no facilities are established. The cultivation stage contributes the most representing up to 90 % of the overall Eco-indicator 99 damage. A scarce 3.1 % is the contribution of transportation, followed by the production with an almost negligible 2.5 %. It is worth mentioning that the cultivation stage includes

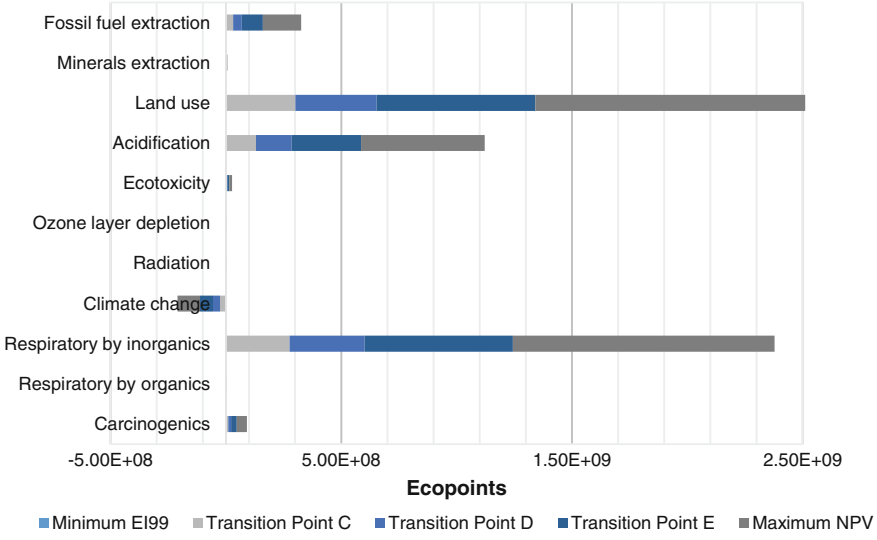


Fig. 6 Contribution of the 11 impact categories of Eco-indicator 99 in the overall environmental damage for extreme solutions and technology transition points of the Pareto

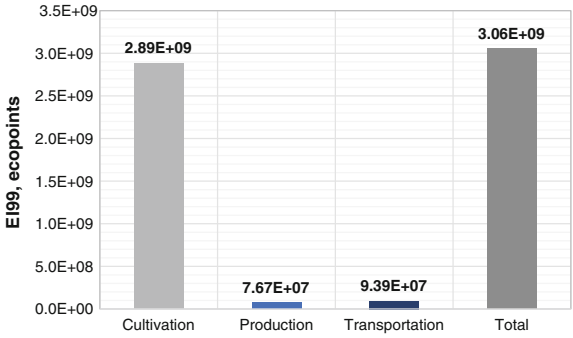


Fig. 7 Contribution of supply chain stages in Eco-indicator 99 metric for maximum NPV

sugar cane plantation, growing and harvesting as well as transportation of raw materials to mills.

On the other hand, a breakdown of total cost is provided in Fig. 8. Capital cost dominates, being 62.20 % higher than operating cost. Facility capital investment and the corresponding operating cost are the main contributors. These expenses are given by the production plants settlement (Fig. 9).

The topology of the resulting supply chain operating at the economic optimum (see Fig. 9) comprises five ethanol generation and four sugar production facilities. Two of the ethanol ones employ honey as raw material, whereas the remaining uses sugar cane directly. The facilities are located in regions with available sugar cane plantations attempting to fulfil the internal ethanol demand as much as possible. Misiones (G09) is not accounted for due to the associated lowest capacity, while Tucumán (G16) becomes the production centre. This region supplies white and raw sugar as well as ethanol to surrounding regions including Santa Fe (G18), where ethanol yield is not enough to satisfy the inner consumption.

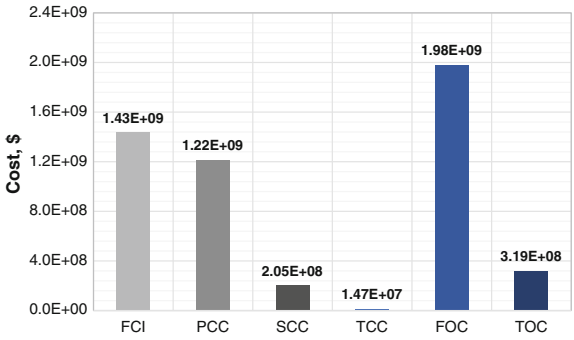


Fig. 8 Breakdown of total cost for the NPV maximisation. FCI for facility capital investment. PCC, SCC and TCC for production, storage and transportation respectively. FOC and TOC for facility and transportation operating cost respectively

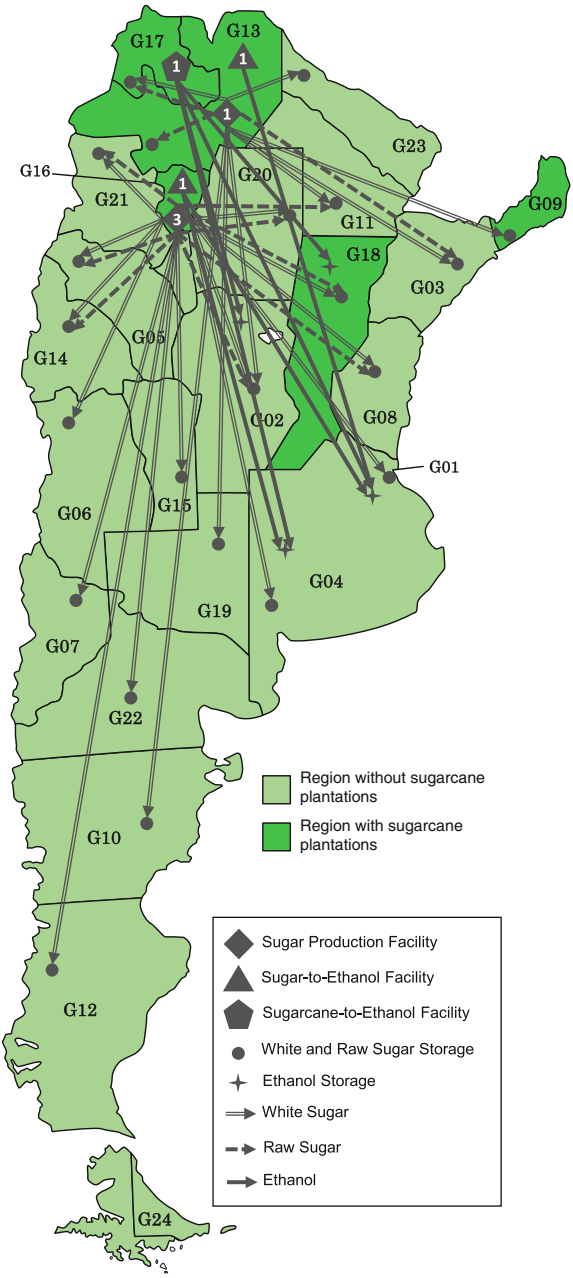


Fig. 9 Topology of the bioethanol supply chain referred to maximum NPV. The production and storage facilities are allocated and the sugar and ethanol delivery flows are illustrated

In terms of delivery flows, sugar is distributed to every region except to Tierra del Fuego (G24), which is the remotest province with the lowest demand associated. On the contrary, the amount of ethanol produced covers the demand of only the main four consumers.

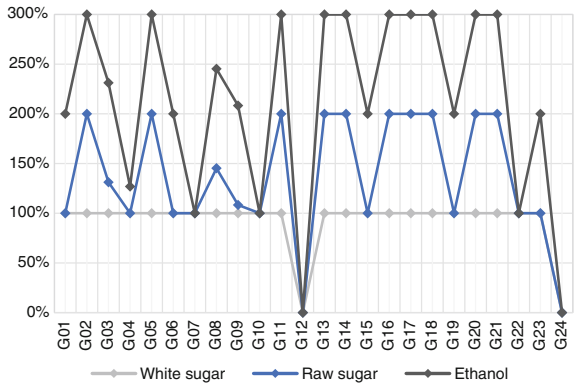


Fig. 10 Sugar and ethanol demand satisfaction per region for supply chain configuration of the maximum NPV. Nomenclature to identify the regions has been provided in Table 4 previously

Note that environmental damage reductions are accomplished in the model by modifying production rates and by switching technologies. The satisfaction levels of sugar and bioethanol demand per region are portrayed in Fig. 10. Overall, two-thirds of the ethanol demand is met in the maximum net present value solution, compared to 99 % accomplished in the case of sugar.

With regards to the sugar demand, the slight variations observed in the Pareto alternatives compared to the economic optimum are given by the provinces of Rio Negro (G22), Neuquén (G07) and Chubut (G10). The location of these regions, on the southern part of the country far from sugar cane plantations, is the main reason why their sugar demand is left unsatisfied as we move from the maximum economic performance point. In the case of ethanol demand, null satisfaction levels are observed in 10 out of the 24 regions. Furthermore, none of the resulting supply chains meet the consumption requirement of the following five regions: Neuquén, Chubut, Santa Cruz (G12), Río Negro (G22) and Tierra del Fuego (G24). This might be explained by their far location and low ethanol forecasted demand.

4 Conclusions

This chapter described in detail how to optimise more sustainable energy systems considering environmental and economic criteria simultaneously. To illustrate the benefits of combining life cycle assessment and multi-objective optimisation in this context, we addressed the design and planning of bioethanol from sugar cane supply chains. The study accounted for economic aspects and environmental concerns simultaneously as the performance measures. The environmental damage was estimated throughout all the stages involved in the supply chain by means of the LCA-based method Eco-indicator 99. The hierarchical perspective was adopted

and 11 impact categories were included. The optimisation problem was mathematically formulated as a multi-objective MILP algorithm.

The Argentinian industry was the real-life scenario to test the capabilities of the modelling approach. The set of Pareto alternatives reveal that by switching production technologies and modulating the associated production rates it is possible to reduce the environmental impact at the expense of sacrificing the economic performance. Particularly, the dominant technology to produce ethanol is the fermentation from honey as intermediate product of the upstream sugar process. The strategy to achieve environmental savings is the allocation of fewer number of ethanol plants in regions with available sugar cane capacity. This allows to satisfy internal consumption and delivery to neighbour regions with high forecasted demand. The motivation is to reduce in doing so the vinasses generation and the transportation impact, which are the main drivers of environmental deterioration.

Moreover, the analysis per impact category of the Eco-indicator 99 shows an additional conflict among environmental metrics. The main reason for this is the negative values that climate change impact can take as a result of carbon dioxide emissions credits. Future research should be focused on the treatment of uncertainties affecting the calculations and the implementation of a grid with better resolution in order to assess the spatial distribution of environmental impacts.

References

Abo-Hamad, W. & Arisha, a M.R., 2011. Simulation-Optimisation Methods in Supply Chain Applications: A Review. *Irish Journal of Management*, 30(2), pp.95–124. Available at: http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=65271281&site=ehost-live.

Almansoori, A. & Shah, N., 2006. Design and Operation of a Future Hydrogen Supply Chain. *Chemical Engineering Research and Design*, 84(6), pp.423–438. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0263876206729185.

Azapagic, a. & Clift, R., 1999. Life cycle assessment and multiobjective optimisation. *Journal of Cleaner Production*, 7(2), pp.135–143.

Beamon, B.M., 1998. Supply Chain Design and Analysis : Models and Methods. *International Journal of Production Economics*, 55(1), pp.1–22.

Brandenburg, M. et al., 2014. Quantitative models for sustainable supply chain management: Developments and directions. *European Journal of Operational Research*, 233(2), pp.299–312. Available at: http://dx.doi.org/10.1016/j.ejor.2013.09.032.

Corsano, G., Vecchietti, A.R. & Montagna, J.M., 2011. Optimal design for sustainable bioethanol supply chain considering detailed plant performance model. *Computers & Chemical Engineering*, 35(8), pp.1384–1398. Available at: http://linkinghub.elsevier.com/retrieve/pii/S009813541100010X.

Dehghanian, F. & Mansour, S., 2009. Designing sustainable recovery network of end-of-life products using genetic algorithm. *Resources, Conservation and Recycling*, 53(10), pp.559–570. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0921344909000664.

Dreyer, L.C., Niemann, A.L. & Hauschild, M.Z., 2003. Comparison of Three Different LCIA Methods: EDIP97, CML2001 and Eco-indicator 99. *The International Journal of Life Cycle Assessment*, 8(4), pp.191–200.

Elia A., J. et al., 2011. Optimal energy supply network determination and life cycle analysis for hybrid coal, biomass, and natural gas to liquid (CBGTL) plants using carbon-based hydrogen

production. *Computers and Chemical Engineering*, 35(8), pp.1399–1430. Available at: http://dx.doi.org/10.1016/j.compchemeng.2011.01.019.

Eskandarpour, M. et al., 2015. Sustainable supply chain network design: an optimization-oriented review. *Omega*, 54, pp.11–32. Available at: http://www.sciencedirect.com/science/article/pii/S0305048315000080.

Giarola, S., Zamboni, A. & Bezzo, F., 2012. Environmentally conscious capacity planning and technology selection for bioethanol supply chains. *Renewable Energy*, 43, pp. 61–72. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0960148111006811.

Goedkoop, M. & Spriensma, R., 2001. The Eco-indicator 99 - A damage oriented method for Life Cycle Impact Assessment. *Assessment*, p.144.

Guillén-Gosálbez, G. et al., 2006. Addressing the design of chemical supply chains under demand uncertainty. *Computer Aided Chemical Engineering*, 21(C), pp.1095–1100.

Hugo, A. et al., 2005. Hydrogen infrastructure strategic planning using multi-objective optimization. *International Journal of Hydrogen Energy*, 30(15), pp.1523–1534. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0360319905001163.

Kim, J. et al., 2011. Design of biomass processing network for biofuel production using an MILP model. *Biomass and bioenergy*, 35(2), pp.853–871.

Kostin, A. et al., 2012. Identifying Key Life Cycle Assessment Metrics in the Multiobjective Design of Bioethanol Supply Chains Using a Rigorous Mixed-Integer Linear Programming Approach. *Industrial & Engineering Chemistry Research*, 51(14), pp.5282–5291. Available at: http://pubs.acs.org/doi/abs/10.1021/ie2027074.

Luo, X. et al., 2013. Operational planning optimization of steam power plants considering equipment failure in petrochemical complex. *Applied Energy*, 112(1), pp.1247–1264. Available at: http://dx.doi.org/10.1016/j.apenergy.2012.12.039.

Meixell, M.J. & Gargeya, V.B., 2005. Global supply chain design: A literature review and critique. *Transportation Research Part E: Logistics and Transportation Review*, 41(6 SPEC. ISS.), pp.531–550.

Mele, F., Espuna, A. & Puigjaner, L., 2005. Environmental impact considerations into supply chain management based on life-cycle assessment. In *Innovation by life cycle management LCM 2005 international conference*.

Mele, F.D., Guillén-Gosálbez, G. & Jiménez, L., 2009. Optimal Planning of Supply Chains for Bioethanol and Sugar Production with Economic and Environmental Concerns., (October 2015), pp.997–1002. Available at: http://linkinghub.elsevier.com/retrieve/pii/S157079460970166X.

Mele, F.D. et al., 2011. Multiobjective model for more sustainable fuel supply chains. A case study of the sugar cane industry in argentina. *Industrial and Engineering Chemistry Research*, 50(9), pp.4939–4958.

Michelsen, O., Fet, A.M. & Dahlsrud, A., 2006. Eco-efficiency in extended supply chains: A case study of furniture production. *Journal of Environmental Management*, 79(3), pp.290–297.

Neiro, S.M.S. & Pinto, J.M., 2004. A general modeling framework for the operational planning of petroleum supply chains. *Computers & Chemical Engineering*, 28(6–7), pp.871–896. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0098135403002308.

Nikolopoulou, A. & Ierapetritou, M.G., 2012. Optimal design of sustainable chemical processes and supply chains: A review. *Computers and Chemical Engineering*, 44, pp.94–103. Available at: http://dx.doi.org/10.1016/j.compchemeng.2012.05.006.

Papageorgiou, L.G., 2009. Supply chain optimisation for the process industries: Advances and opportunities. *Computers and Chemical Engineering*, 33(12), pp.1931–1938.

Pieragostini, C., Mussati, M.C. & Aguirre, P., 2012. On process optimization considering LCA methodology. *Journal of Environmental Management*, 96(1), pp.43–54. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0301479711003951.

Pinto-Varela, T., Barbosa-Póvoa, A.P.F.D. & Novais, A.Q., 2011. Bi-objective optimization approach to the design and planning of supply chains: Economic versus environmental performances. *Computers & Chemical Engineering*, 35(8), pp.1454–1468. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0098135411000998.

Ruiz-Femenia, R. et al., 2013. Multi-objective optimization of environmentally conscious chemical supply chains under demand uncertainty. *Chemical Engineering Science*, 95, pp.1–11.

Sabio, N. et al., 2014. Multiobjective optimization under uncertainty of the economic and life-cycle environmental performance of industrial processes. *AIChE Journal*, 60(6), pp.2098–2121. Available at: http://onlinelibrary.wiley.com/doi/10.1002/aic.14385/abstract\nhttp://onlinelibrary.wiley.com/doi/10.1002/aic.14385/abstract;jsessionid=962ADE4225E69F85F3AAF92FA31078-CE.f04t03?systemMessage=Wiley+Online+Library+will+be+disrupted+on+the+18th+October+from+1.

Sarkis, J., Zhu, Q. & Lai, K.H., 2011. An organizational theoretic review of green supply chain management literature. *International Journal of Production Economics*, 130(1), pp.1–15. Available at: http://dx.doi.org/10.1016/j.ijpe.2010.11.010.

Seuring, S., 2013. A review of modeling approaches for sustainable supply chain management. *Decision Support Systems*, 54(4), pp.1513–1520. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0167923612001741.

Stern, M.O., Ayres, R.U. & Saxton, J.C., 1973. Tax Strategies for Industrial Pollution Abatement. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(6), pp.588–603. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4309312.

You, F. et al., 2012. Optimal design of sustainable cellulosic biofuel supply chains: multiobjective optimization coupled with life cycle assessment and input–output analysis. *AIChE Journal*, 58(4), pp.1157–1180.

Yue, D. et al., 2014. Sustainable design and operation of cellulosic bioelectricity supply chain networks with life cycle economic, environmental, and social optimization. *Industrial & Engineering Chemistry Research*, 53(10), pp.4008–4029.

Zamboni, A., Shah, N. & Bezzo, F., 2009. Spatially Explicit Static Model for the Strategic Design of Future Bioethanol Production Systems. 1. Cost Minimization. *Energy & fuels*, 23(10), pp.5121–5133. Available at: http://dx.doi.org/10.1021/ef900456w\nhttp://pubs.acs.org/doi/pdfplus/10.1021/ef900456w.

Zhou, Z., Cheng, S. & Hua, B., 2000. Supply chain optimization of continuous process industries with sustainability considerations. *Computers & Chemical Engineering*, 24(2–7), pp.1151–1158.

Large-Scale Stochastic Mixed-Integer Programming Algorithms for Power Generation Scheduling

Kibaek Kim and Victor M. Zavala

Abstract This chapter presents a stochastic unit commitment model for power systems and revisits parallel decomposition algorithms for these types of models. The model is a two-stage stochastic programming problem with first-stage binary variables and second-stage mixed-binary variables. The *here-and-now* decision is to find day-ahead schedules for slow thermal power generators. The *wait-and-see* decision consists of dispatching power and scheduling fast-start generators. We discuss advantages and limitations of different decomposition methods and provide an overview of available software packages. A large-scale numerical example is presented using a modified IEEE 118-bus system with uncertain wind power generation.

1 Stochastic Unit Commitment Model

Unit commitment (UC) is a decision-making process that schedules power generation units and production levels over a planning horizon. This process is central to ensuring efficiency and reliability of the power system operation. While fossil-fuel

K. Kim (✉)
Mathematics and Computer Science Division, Argonne National Laboratory,
9700 South Cass Avenue, Lemont, IL 60439, USA
e-mail: kimk@anl.gov

V.M. Zavala
Department of Chemical and Biological Engineering,
University of Wisconsin-Madison, 1415 Engineering Dr,
Madison, WI 53706, USA
e-mail: victor.zavala@wisc.edu

M. Martín (ed.), *Alternative Energy Sources and Technologies*,
DOI 10.1007/978-3-319-28752-2_18

power plants produce 67 % of the total electricity generation in the United States (according to 2014 data), renewable power continues to penetrate into the electricity market (U.S. Energy Information Administration 2015). This trend has motivated the development of many unit commitment models that can mitigate uncertainty of renewable supplies. In this chapter, we present a stochastic unit commitment model that determines on–off schedules of the generating units and operation levels, with the objective of cost minimization under uncertain renewable generation. We focus on a day-ahead hourly scheduling of fast and slow generators subject to real-time wind power supply for a 24-h planning horizon. In this chapter, we use the stochastic unit commitment formulation presented in (Kim and Zavala 2015). We use this model to motivate different algorithmic approaches to solve these types of problems.

We let $\mathscr{T} := \{1, \ldots, T\}$ be the set of time periods in the planning horizon. We assume that we have a set $\mathscr{W}$ of wind power generators and that the generation level W_{jwt} from each generator $w \in \mathscr{W}$ is given for each scenario $j \in \mathscr{S}$ and at time $t \in \mathscr{T}$. We also assume that the set $\mathscr{S} := \{1, \ldots, S\}$ has only a finite number of scenarios with corresponding probabilities $\{\pi_1, \ldots, \pi_S\}$. Let $\mathscr{G}$ be the set of generators considered in the model, and $\mathscr{G}_s$, $\mathscr{G}_f$ be the set of slow generators and fast generators, respectively. We assume that the slow generators are required to be scheduled a day ahead whereas the fast generators can start on demand (i.e., in real time). Each generator $g \in \mathscr{G}$ starts up, operates, and shuts down at the corresponding costs $C_g^{\mathrm{up}}, C_g^{\mathrm{fx}}$, and C_g^{dn}, respectively. We denote the binary decision variables indicating whether generator g is on or off for each scenario j at time t as x_{jgt}. We also denote the binary decision variables indicating whether generator g starts up (or shuts down) for each scenario j at time t by u_{jgt} (or v_{jgt}, respectively).

1.1 Logical Constraints for Commitment, Startup, and Shutdown Decisions

The logical relations between commitment, startup, and shutdown decisions are given, respectively, by

$$1 - x_{jg(t-1)} \geq u_{jgt}, \quad \forall j \in \mathscr{S},\ g \in \mathscr{G},\ t \in \mathscr{T}, \tag{1}$$

$$x_{jg(t-1)} \geq v_{jgt}, \quad \forall j \in \mathscr{S},\ g \in \mathscr{G},\ t \in \mathscr{T}, \tag{2}$$

$$x_{jgt} - x_{jg(t-1)} = u_{jgt} - v_{jgt}, \quad \forall j \in \mathscr{S},\ g \in \mathscr{G},\ t \in \mathscr{T}, \tag{3}$$

where the initial status x_{jg0} of generator g is given for each scenario $j \in \mathscr{S}$. Each generator g has minimum uptime and downtime denoted by UT_g and DT_g, respectively. Hence, we have the following constraints:

$$x_{jgt} \geq \sum_{\tau=\max\{1,t-\mathrm{UT}_g+1\}}^{t} u_{jg\tau}, \quad \forall j \in \mathscr{S},\ g \in \mathscr{G},\ t \in \mathscr{T}, \tag{4}$$

$$1 - x_{jgt} \geq \sum_{\tau=\max\{1,t-\mathrm{DT}_g+1\}}^{t} u_{jg\tau}, \quad \forall j \in \mathscr{S},\ g \in \mathscr{G},\ t \in \mathscr{T}, \tag{5}$$

with the initial operating status

$$x_{jgt} = 1, \quad \forall j \in \mathscr{S},\ g \in \mathscr{G}, \quad t \in \{1, \ldots, \mathrm{UT}_g^{\mathrm{init}}\}, \tag{6}$$

$$x_{jgt} = 0, \quad \forall j \in \mathscr{S},\ g \in \mathscr{G}, \quad t \in \{1, \ldots, \mathrm{DT}_g^{\mathrm{init}}\}, \tag{7}$$

where $\mathrm{UT}_g^{\mathrm{init}}$ and $\mathrm{DT}_g^{\mathrm{init}}$ are the initial uptime and downtime of generator $g \in \mathscr{G}$, respectively. The decisions for each slow generator should be the same for all the scenarios, because the slow generators cannot be adjusted in real time. This is called the *nonanticipativity constraint* and can be expressed mathematically as

$$x_{igt} = x_{jgt},\ u_{igt} = u_{jgt},\ v_{igt} = v_{jgt}, \quad \forall i, j \in \mathscr{S},\ g \in \mathscr{G}_s,\ t \in \mathscr{T}. \tag{8}$$

1.2 Generation Limits, Spinning Reserve Requirements, and Ramping Constraints

The model also determines the production levels P_{jgt} and spinning reserve amounts s_{jgt} of generator g at time t for each scenario j. The limitations of the generation level are

$$P_g^{\min} x_{jgt} \leq p_{jgt} \leq P_g^{\max} x_{jgt} - s_{jgt}, \quad \forall j \in \mathscr{S},\ g \in \mathscr{G},\ t \in \mathscr{T}, \tag{9}$$

where $P_g^{\min}$ and $P_g^{\max}$ are the minimal and maximal levels of generator $g \in \mathscr{G}$, respectively. The generation rates are physically constrained according to

$$-\mathrm{RD}_g \leq p_{jgt} - p_{jg(t-1)} \leq \mathrm{RU}_g - s_{jgt}, \quad \forall j \in \mathscr{S},\ g \in \mathscr{G},\ t \in \mathscr{T}, \tag{10}$$

$$s_{jgt} \leq \mathrm{RC}_g x_{jgt}, \quad \forall j \in \mathscr{S},\ g \in \mathscr{G},\ t \in \mathscr{T}, \tag{11}$$

where RD_g and RUg are the minimal and maximal rates of generation change for each generator $g \in \mathscr{G}$, respectively; RC_g is the ramping capacity of generator g; and the initial production level p_{jg0} of generator g is given for each scenario s. The constraints (10), (11) are called *ramping constraints*. The spinning reserve requirement SR_t is given by

$$\sum_{g\in\mathscr{G}} s_{jgt} \geq \mathrm{SR}_t, \quad \forall j \in \mathscr{S},\ t \in \mathscr{T}. \tag{12}$$

1.3 Flow Balance and Transmission Line Capacity Constraints

The system balance between the net generation level and load is given by

$$\sum_{g\in\mathscr{G}} p_{jgt} = \sum_{n\in\mathscr{N}} D_{jnt} - \sum_{w\in\mathscr{W}} W_{jwt}, \quad \forall j \in \mathscr{S},\ t \in \mathscr{T}, \tag{13}$$

where D_{jnt} is the demand load of bus n at time t for each scenario j. Let $\mathscr{L}$ be the set of transmission lines where each line $l \in \mathscr{L}$ has minimal and maximal capacities denoted by $F_l^{\min}$ and $F_l^{\max}$, respectively. Let L_{ln} be the load shift factor of transmission line l with respect to bus n for each $l \in \mathscr{L}$ and $n \in \mathscr{N}$. The load shift factor is also known as the power transfer distribution factor that represents the change of power flow on line l with respect to the change in injection at bus n. The transmission line capacity is constrained as follows:.5

$$\begin{aligned} F_l^{\min} \leq \sum_{g\in\mathscr{G}} L_{ln(g)} p_{jgt} - \sum_{n\in\mathscr{N}} L_{ln} D_{jnt} + \sum_{w\in\mathscr{W}} L_{ln(w)} W_{jwt} \leq F_l^{\max}, \\ \forall j \in \mathscr{S},\ l \in \mathscr{L},\ t \in \mathscr{T}, \end{aligned} \tag{14}$$

where $n(g)$ and $n(w)$ are, respectively, the indices of buses where generator g and wind farm w are located.

1.4 Piecewise Linear Objective Function

The objective function is the expected cost of operating generators and electricity production. While the cost functions of commitment, startup, and shutdown are linear, the production cost is a nonlinear function that is often approximated by a piecewise linear function. Let $\mathscr{K}$ be the set of linear segments to approximate the cost function. Let q_{jgkt} be the production level of generator g at cost segment k at time t for a given scenario j, and $Q_{gk}^{\max}$ be the maximum production limit of generator g at cost C_{gk}^{mar} of segment k. The piecewise linear approximation of the cost function is achieved by adding the following constraints:

$$q_{jgkt} \leq Q_{gk}^{\max} x_{jgt}, \quad \forall j \in \mathscr{S},\ g \in \mathscr{G},\ k \in \mathscr{K},\ t \in \mathscr{T}, \tag{15}$$

$$p_{jgt} = P_g^{\min} x_{jgt} + \sum_{k \in \mathscr{K}} q_{jgkt}, \quad \forall j \in \mathscr{S},\ g \in \mathscr{G},\ t \in \mathscr{T}. \tag{16}$$

1.5 Stochastic Mixed-Integer Programming Formulation

The stochastic unit commitment model is formulated as follows:

$$\min \sum_{j \in \mathscr{S}} \sum_{t \in \mathscr{T}} \sum_{g \in \mathscr{G}} \pi_j \left(C_g^{\mathrm{fx}} x_{jgt} + C_g^{\mathrm{up}} u_{jgt} + C_g^{\mathrm{dn}} v_{jgt} + \sum_{k \in \mathscr{K}} C_{gk}^{\mathrm{mar}} q_{jgkt} \right) \tag{17a}$$

$$\text{s.t. } (1) - (16) \tag{17b}$$

$$\begin{aligned} & x_{jgt} \in \{0,1\},\ 0 \le u_{jgt}, v_{jgt} \le 1,\ p_{jgt}, q_{jgkt}, s_{jgt} \ge 0, \\ & \forall j \in \mathscr{S},\ g \in \mathscr{G},\ k \in \mathscr{K},\ t \in \mathscr{T}. \end{aligned} \tag{17c}$$

This is a two-stage stochastic mixed-integer program (SMIP) where the first-stage variables represent commitment, startup, and shutdown decisions of slow generators and the second-stage variables represent all the other decisions including those of fast generators for each scenario $j \in \mathscr{S}$. The objective function (17a) is the expected cost of operating generators and producing electricity during the planning horizon $\mathscr{T}$. Note that the integrality restriction is implicitly imposed on variables u_{jgt} and v_{jgt} for any given binary value x_{jgt}. By explicitly relaxing the integrality on u_{jgt} and v_{jgt} the MIP solver can generate tight, valid cutting planes and primal solutions by heuristic procedures.

1.6 Technical Challenges of SMIP

SMIP is a challenging problem because the dimensionality increases with the number of scenarios and because the objective function is nonconvex and discontinuous in the first-stage variables. The dimensionality of SMIP makes linear algebra operations in simplex procedures expensive. Moreover, even the linear programming relaxation of the problem might not fit in memory. Consequently, off-the-shelf branch-and-cut solvers are limited to problems with few scenarios. In order to address this issue, different decomposition methods have been proposed. Benders decomposition (also known as the L-shaped method) has been the most popular method for solving SMIPs, but this algorithm does not have convergence guarantees when integer variables are present in the second stage (e.g., startup and shutdown of fast generators). Therefore, other approaches are required. These include specialized branch-and-bound techniques using tender variables (Ahmed et al. 2004) and various convexification techniques (e.g., Gomory cuts (Gade et al. 2014; Zhang and Küçükyavuz 1933),

mixed-integer rounding cuts (Kim and Mehrotra 2015), and disjunctive cuts (Sen and Higle 2005; Sherali and Fraticelli 2002)). However, they are limited to certain problem classes. Consequently, we are interested in decomposition methods that can solve SMIP to optimality or at least provide good upper and/or lower bounds. This can be achieved using a variety of methods such as dual decomposition and progressive hedging. Dual decomposition is implemented in the open-source package `DSP` (Kim and Zavala 2015), and progressive hedging is implemented in `PySP` (Watson et al. 2012). `DSP` provides a `Julia`-based modeling interface (i.e., `StochJuMP` (Huchette et al. 2014)), and `PySP` provides a `Python`-based modeling interface. The `ddsip` package (Märkert and Gollmer 2014) also implements a dual decomposition method but does not support model specification through a standard file format (i.e., SMPS (Birge et al. 1987)) and algebraic modeling languages.

2 Scenario Decomposition

In this section, we present scenario decomposition methods for the stochastic unit commitment model presented in Sect. 1. For simplicity, we write this model in the general form

$$z = \min_{x_j, y_j} \quad \sum_{j \in \mathscr{S}} \pi_j \left(c^T x_j + q_j^T y_j \right) \tag{18a}$$

$$\text{s.t.} \quad \sum_{j \in \mathscr{S}} H_j x_j = 0, \quad (\lambda) \tag{18b}$$

$$(x_j, y_j) \in G_j, \quad \forall j \in \mathscr{S}, \tag{18c}$$

where x is the first-stage decision vector and y_j is the second-stage decision vector for each scenario $j \in \mathscr{S}$. Equation (18b) represents the nonanticipativity constraints (8) with a suitable matrix $H_j \in \mathbb{R}^{S \cdot n_1 \times n_1}$, and $\lambda \in \mathbb{R}^{S \cdot n_1}$ is the corresponding dual variable. In Eq. (18c), the set G_j of feasible solutions is defined by all the constraints except the nonanticipativity constraints.

2.1 Dual Decomposition

Dual decomposition was first proposed in (CarøE and Schultz 1999). This method applies a Lagrangian relaxation of the nonanticipativity constraints (8) to obtain the Lagrangian dual function

$$D(\lambda) := \min_{x_j, y_j} \left\{ \sum_{j \in \mathscr{S}} L_j(x_j, y_j, \lambda) : (x_j, y_j) \in G_j, \quad \forall j \in \mathscr{S} \right\}, \tag{19}$$

where

$$L_j(x_j, y_j, \lambda) := \pi_j\left(c^T x_j + q_j^T y_j\right) + \lambda^T (H_j x_j). \tag{20}$$

For fixed λ, the Lagrangian dual function can be decomposed as

$$D(\lambda) = \sum_{j \in \mathscr{S}} D_j(\lambda), \tag{21}$$

where

$$D_j(\lambda) := \min_{x_j, y_j} \left\{ L_j(x_j, y_j, \lambda) : (x_j, y_j) \in G_j \right\}. \tag{22}$$

We thus seek to obtain the best lower bound for (18a–c) by solving the Lagrangian dual problem

$$z_{\mathrm{LD}} := \max_{\lambda} \sum_{j \in \mathscr{S}} D_j(\lambda). \tag{23}$$

Proposition 1 is an important property of the Lagrangian relaxation, which shows the tightness of the lower bound z_{LD} (Geoffrion 1974).

Proposition 1 *The optimal value z_{LD} of the Lagrangian dual problem (23) is equal to the optimal value of the following linear program:*

$$\min_{x_j, y_j} \left\{ \sum_{j \in \mathscr{S}} \pi_j \left(c^T x_j + q_j^T y_j\right) : \sum_{j \in \mathscr{S}} H_j x_j = 0,\ (x_j, y_j) \in \mathrm{conv}(G_j), \quad \forall j \in \mathscr{S} \right\}, \tag{24}$$

where $\mathrm{conv}(G_j)$ denotes the convex hull of G_j. Moreover, $z_{\mathrm{LD}} \geq z_{\mathrm{LP}}$ holds, where z_{LP} is the optimal value of the linear programming relaxation of (18a–c).

2.1.1 Subgradient Method

Subgradient methods have been widely used in nonsmooth optimization. Let λ^k be the dual variable at iteration $k \geq 0$, and x_s^k be an optimal solution of (22) for given λ^k. The dual variable is updated as

$$\lambda^{k+1} = \lambda^k - \alpha_k \sum_{j \in \mathscr{S}} H_j x_j^k, \tag{25}$$

where $\alpha_k \in (0, 1]$ is the step size. This method updates the duals using a subgradient of $D(\lambda)$ at λ^k, denoted by $\sum_{j \in \mathscr{S}} H_j x_j^k$. Different step-size rules have been studied for

subgradient methods (Bertsekas 2015). For example, in (Fisher 1985) the step size α_k is given by

$$\alpha_k := \beta_k \frac{z_{\text{UB}} - D(\lambda^k)}{\left\| \sum_{j \in \mathscr{S}} H_j x_j^k \right\|_2^2}, \tag{26}$$

where z_{UB} is the objective value of the best-known feasible solution to (18a–c) up to iteration k and β_k is a user-defined positive scalar. The subgradient method is summarized in Algorithm 1.

Algorithm 1 Dual Decomposition Based on Subgradient Method (DD-Sub)

1: Set $k \leftarrow 0, z_{\text{LB}} \leftarrow -\infty, z_{\text{UB}} \leftarrow \infty$ and $\gamma \leftarrow 0$.
2: **loop**
3: SOLVE (22) to obtain $D_j(\lambda^k)$ and (x_j^k, y_j^k) for given λ^k and for all $j \in \mathscr{S}$
4: **if** $D(\lambda^k) > z_{\text{LB}}$ **then**
5: $z_{\text{LB}} \leftarrow D(\lambda^k)$
6: **else**
7: $\gamma \leftarrow \gamma + 1$
8: **if** $\gamma = \gamma^{\max}$ **then**
9: $\beta_k \leftarrow 0.5\beta_k$ and $\gamma \leftarrow 0$
10: **end if**
11: **end if**
12: UPDATE z_{UB} for given x_s^k
13: $k \leftarrow k + 1$
14: **end loop**

Algorithm 1 is initialized with user-defined parameters $\lambda^0, \gamma^{\max}$, and β_0 and reduces β_k by half when the best lower bound z_{LB} is not improved for the last $\gamma^{\max}$ iterations (lines 8–10). The best upper bound z_{UB} may be obtained by solving (18a–c) for fixed x_s^k (line 12). An important limitation of the subgradient method is that finite termination cannot be proved (Fisher 1985).

2.1.2 Cutting-Plane Method

The cutting-plane method is an outer approximation scheme that solves the Lagrangian dual problem by iteratively adding linear inequalities. The outer approximation of (23) at iteration k is given by the Lagrangian master problem

$$m_k := \max_{\theta_j, \lambda} \sum_{j \in \mathscr{S}} \theta_j \tag{27a}$$

$$\text{s.t. } \theta_j \leq D_j(\lambda^l) + \left(H_j x_j^l\right)^T (\lambda - \lambda^l), \quad \forall j \in \mathscr{S}, \quad l = 0, 1, \ldots, k. \tag{27b}$$

The dual variable λ^{k+1} is obtained by solving the approximation (27a–b) at iteration k. We define the primal–dual solution of the Lagrangian master problem as the triplet (θ, λ, π). Here, $\theta := (\theta_1, \ldots, \theta_S)$ and $\pi := (\pi_1^0, \ldots, \pi_1^k, \ldots, \pi_S^0, \ldots, \pi_S^k)$, where π_s^l are the dual variables of (27b). The master problem (27a–b) exhibits a dual block-angular structure that allows for parallelism in solving the master using an interior-point solver (Lubin et al. 2013). The method is summarized in Algorithm 2.

Algorithm 2 Dual Decomposition Based on Cutting-Plane Method

1: $k \leftarrow 0$ and $\lambda^0 \leftarrow 0$
2: SOLVE (22) to obtain $D_j(\lambda^k)$ and (x_j^k, y_j^k) for given λ^k and for each $j \in \mathscr{S}$
3: $z_{\text{LB}} \leftarrow D(\lambda^k)$.
4: **repeat**
5: ADD (27b) for given $D(\lambda^k)$ and x_j^k
6: SOLVE (27) to obtain m_k and $(\theta^{k+1}, \lambda^{k+1})$
7: SOLVE (22) to obtain $D_j(\lambda^{k+1})$ and (x_j^{k+1}, y_j^{k+1}) for given λ^{k+1} and for all $j \in \mathscr{S}$
8: UPDATE $z_{\text{LB}} \leftarrow \max\{z_{\text{LB}}, D(\lambda^{k+1})\}$.
9: $k \leftarrow k+1$
10: **until** $m_{k-1} \leq D(\lambda^k)$

The function $D_j(\lambda)$ is piecewise linear concave in λ supported by the linear inequalities (27b). Assuming that the master problem (27a–b) and the subproblem (22) can be solved to optimality, Algorithm 2 terminates with an *optimal* solution of (23) after a finite number of steps because the number of linear inequalities required to approximate $D(\lambda)$ is finite. This gives the cutting-plane method a natural termination criterion (i.e., $m_{k-1} \leq D(\lambda^k)$). In other words, this criterion indicates that m_{k-1} matches the Lagrangian dual function $D(\lambda^k)$ and thus the maximum of the Lagrangian master problem matches the maximum of the Lagrangian dual problem.

2.1.3 Variants of the Cutting-Plane Method

The cutting-plane method is inherently unstable, and hence the solutions of the master problem (27a–b) oscillate significantly when the Lagrangian dual (23) is not well approximated at the beginning of the iterations. Moreover, the solution also suffers from degeneracy. Several variants of the cutting-plane method have been proposed to overcome these issues. Here we present two variants.

Interior-Point Cutting-Plane Method

The interior-point method (IPM) with early termination criteria was proposed in (Kim and Zavala 2015). The IPM solves the master problem (27a–b) suboptimally to find stronger cuts from interior feasible solutions and to avoid degeneracy.

The IPM checks the termination criteria in the following order:

1. $\sum_{j\in\mathscr{S}} \theta_j^k \geq z_{\text{UB}}$
2. $g_k(\theta^k, \lambda^k, \mu^k) < \varepsilon_{\text{IPM}}^k$

Here, $g_k(\theta^k, \lambda^k, \mu^k)$ is the relative duality gap of the primal–dual feasible solution $(\theta^k, \lambda^k, \mu^k)$ of the master (27a–b) at iteration k. The tolerance $\varepsilon_{\text{IPM}}^k$ can be relaxed when the duality gap of the dual decomposition method is large. It is updated as follows:

Algorithm 3 Dual Decomposition Based on Interior-Point Cutting-Plane Method (IPCPM)

$k \leftarrow 0, \lambda^0 \leftarrow 0$ and $z_{\text{UB}} \leftarrow \infty$
SOLVE (22) to obtain $D_j(\lambda^k)$ and (x_j^k, y_j^k) for given λ^k and for each $j \in \mathscr{S}$
ADD cutting-planes (27b) to the master (27) for given $D(\lambda^k)$ and x_j^k
$z_{\text{LB}} \leftarrow D(\lambda^k)$.
loop
  SOLVE the master (27) by the IPM to obtain $(\theta^{k+1}, \lambda^{k+1})$
  SOLVE (22) to obtain $D_j(\lambda^{k+1})$ and (x_j^{k+1}, y_j^{k+1}) for given λ^{k+1} and for each $j \in \mathscr{S}$
  if $(\theta^{k+1}, \lambda^{k+1})$ is obtained from the first termination criterion **then**
    if $\theta_j^{k+1} \leq D_j(\lambda^{k+1})$ for all $j \in \mathscr{S}$ **then**
      STOP
    else
      ADD cutting-planes (27b) to the master (27) for given $D(\lambda^{k+1})$ and x_s^{k+1}
    end if
  else if $(\theta^{k+1}, \lambda^{k+1})$ is obtained from the second termination criterion **then**
    if $\theta_j^{k+1} \leq D_j(\lambda^{k+1})$ for all $j \in \mathscr{S}$ **then**
      if $\varepsilon_{\text{IPM}}^k > \varepsilon_{\text{Opt}}$ **then**
        UPDATE $\varepsilon_{\text{IPM}}^{k+1}$ from (28)
      else
        STOP
      end if
    else
      ADD cutting-planes (27b) to the master (27) for given $D(\lambda^{k+1})$ and x_j^{k+1}
    end if
  end if
  $z_{\text{LB}} \leftarrow \max\{z_{\text{LB}}, D(\lambda^k)\}$.
  $k \leftarrow k+1$
end loop

$$\varepsilon_{\text{IPM}}^k := \min\left\{\varepsilon_{\text{IPM}}^{\max}, \frac{g_{k-1}(\tilde{\theta}^{k-1}, \tilde{\pi}^{k-1})}{\delta} + \frac{\tilde{m}_{k-1} - \sum_{j\in\mathscr{S}} D_j(\tilde{\lambda}^{k-1})}{1+|\tilde{m}_{k-1}|}\right\}, \tag{28}$$

where $\tilde{m}_{k-1} := \sum_{j\in\mathscr{S}} \tilde{\theta}_j^{k-1}$ and $\delta > 1$ is the *degree of optimality* (Gondzio et al. 2013). The interior-point cutting-plane method is summarized in Algorithm 3. Algorithm 3 terminates after a finite number of iterations with an optimal solution

of the Lagrangian dual problem (23) (see Theorem 2 in (Kim and Zavala 2015) for a proof).

Bundle Method

The bundle method is a stabilized cutting-plane method where a quadratic stabilizing term is added to the objective function of the master (27a–b). As a result, the method solves the master problem

$$\max_{\theta_j,\lambda} \sum_{j\in\mathscr{S}} \theta_j + \frac{1}{2\tau}\|\lambda - \lambda^+\|^2 \tag{29a}$$

$$\text{s.t. } \theta_j \le D_j(\lambda^l) + \left(H_j x_j^l\right)^T(\lambda - \lambda^l), \quad \forall j \in \mathscr{S}, \quad l = 0, 1, \ldots, k, \tag{29b}$$

where λ^+ is a *stability center* and $\tau > 0$ is a parameter that defines the *stabilization effect* (Lemaréchal et al. 2001). Note that the bundle method becomes equivalent to the cutting-plane method for a large τ and the subgradient method for a small τ. A bundle method for dual decomposition is summarized in Algorithm 4 as presented in (Lubin et al. 2013), where the dual of the master problem (27a–b) is solved. Algorithm 4 is convergent in the limit (Lubin et al. 2013).

Algorithm 4 Dual Decomposition Based on Bundle Method

Choose initial $\kappa \in (0,1)$ and $t > 0$, $k \leftarrow 0, \lambda^+ \leftarrow 0$ and $\lambda^0 \leftarrow 0$
SOLVE (22) to obtain $D_j(\lambda^k)$ and (x_j^k, y_j^k) for given λ^k and for each $j \in \mathscr{S}$
$z_{\text{LB}} \leftarrow D(\lambda^k)$
loop
  ADD cuts (29b) to the master (29) for given $D(\lambda^k)$ and x_j^k
  SOLVE the master (29) to obtain $(\theta^{k+1}, \lambda^{k+1})$
  $v \leftarrow \sum_{j\in\mathscr{S}} \theta_j^{k+1} - z_{\text{LB}}$
  if $v < \varepsilon(1 + |z_{\text{LB}}|)$ **then**
    STOP
  end if
  $k \leftarrow k + 1$
  SOLVE (22) to obtain $D_j(\lambda^k)$ and (x_j^k, y_j^k) for given λ^k and for each $j \in \mathscr{S}$.
  $u \leftarrow 2t\left[1 - \left(D(\lambda^k) - z_{\text{LB}}\right)/v\right]$
  UPDATE $t \leftarrow \min\left\{\max\left\{u, t/10, 10^{-4}\right\}, 10t\right\}$
  if $D(\lambda^k) - D(\lambda^+) > \kappa v$ **then**
    UPDATE $z_{\text{LB}} \leftarrow D(\lambda^k)$
    UPDATE $\lambda^+ \leftarrow \lambda^k$
  end if
end loop

We also note that stabilization of the master solution can be achieved using a trust-region constraint (as opposed to the quadratic term used in the bundle method), as is done in (Kim and Zavala 2015). The trust region avoids the need of tuning the parameter τ.

2.2 Progressive Hedging

Progressive hedging is a scenario decomposition framework that is partly motivated by the augmented Lagrangian dual of (18a–c)

$$\begin{aligned} L_r(\hat{x},\lambda) &= \sum_{j\in\mathscr{S}} \pi_j \left(c^T x_j + q_j^T y_j \right) + \sum_{j\in\mathscr{S}} \lambda^T \left(H_j x_j \right) + 0.5\rho \| \sum_{j\in\mathscr{S}} H_j x_j \|^2 \\ &= \sum_{j\in\mathscr{S}} \pi_j \left(c^T x_j + q_j^T y_j \right) + \sum_{j\in\mathscr{S}} \omega_j^T x_j + 0.5\rho \| \sum_{j\in\mathscr{S}} \pi_j x_j - \hat{x} \|^2, \end{aligned} \tag{30}$$

where $\omega_j := \lambda^T H_j$, and the second equality holds because the nonanticipativity constraints $\sum_{j\in\mathscr{S}} H_j x_j = 0$ are equivalent to $\sum_{j\in\mathscr{S}} \pi_j x_j - \mathbb{E}[x] = 0$ (Rockafellar and Wets 1991). An important observation is that the term $\| \sum_{j\in\mathscr{S}} \pi_j x_j - \hat{x} \|^2$ of (30) cannot be decomposed in scenarios. To achieve decomposition, the progressive hedging algorithm solves subproblems of the form

$$P_j(\hat{x},\omega,\rho) := \min \left\{ \pi_j \left(c^T x_j + q_j^T y_j \right) + \omega^T x_j + 0.5 \| \rho^T (x_j - \hat{x}) \|^2 : (x_j, y_j) \in G_j \right\}, \tag{31}$$

where the vector $\hat{x}$ is the expected value of x from the previous iteration, $\omega \in \mathbb{R}^{n_1}$ is a price vector, and $\rho > 0$ is a perturbation vector. For computational efficiency, the quadratic proximal term of (31) is often approximated using piecewise linear functions (Watson et al. 2012). We note that the progressive hedging subproblem (31) has structural connections with the augmented Lagrangian (30). Unlike the (augmented) Lagrangian relaxation, however, the progressive hedging method seeks to find a feasible solution (i.e., an upper bound). The method is summarized in Algorithm 5.

The choice of the perturbation vector ρ can significantly affect algorithmic performance. A small value of ρ may require many iterations to achieve changes in the first-stage variables, whereas a large value may lead to a suboptimal solution. Different strategies for computing ρ are discussed in (Watson and Woodruff 2011). We also note that Algorithm 5 cannot be guaranteed to terminate in a finite number of steps, but it often provides good-quality solutions.

Algorithm 5 Progressive Hedging (PH)

1: $k \leftarrow 0$ and $\omega_j^0 \leftarrow 0$ for each $j \in \mathscr{S}$
2: SOLVE $P_j(0,0,0)$ to obtain (x_j^k, y_j^k) for all $j \in \mathscr{S}$
3: $z_{\text{LB}} \leftarrow \sum_{j\in\mathscr{S}} P_j(0,0,0)$.
4: $k \leftarrow k+1$
5: **repeat**
6: UPDATE $\hat{x}^{k-1} \leftarrow \sum_{j\in\mathscr{S}} \pi_j x_j^{k-1}$
7: UPDATE $\omega_j^k \leftarrow \omega_j^k + \rho\left(x_j^{k-1} - \hat{x}_j^{k-1}\right)$ for each $j \in \mathscr{S}$.
8: SOLVE $P_j(\hat{x}_j^{k-1}, \omega_j^k, \rho)$ to obtain (x^k, x^k) for each $j \in \mathscr{S}$.
9: $k \leftarrow k+1$
10: **until** x_j^k are equal for all $j \in \mathscr{S}$

2.3 *Incorporating Benders-Type Cutting-Plane Procedure*

One can combine dual decomposition and Benders techniques. For instance, a Benders-type cutting-plane procedure has been proposed for the Lagrangian subproblems (22) in (Kim and Zavala 2015). The aim of the cutting-plane procedure is to eliminate infeasible first-stage solutions and to tighten the Lagrangian subproblems. The computational experiments in (Kim and Zavala 2015) show significant improvement in the quality of the bounds, number of iterations, and solution time. For simplicity, we use the general definition of set $G_j := \{(x_j, y_j) : Ax_j = b,\ T_j x_j + W_j y_j = h_j\}$ for each $j \in \mathscr{S}$.

2.3.1 Feasibility Cuts

Without loss of generality, let $(\hat{x}_1, \hat{y}_1) \in G_1$ be the feasible subproblem solution that is infeasible to the original problem (18a–c). This situation can occur because the stochastic UC problem might not have relatively complete recourse. For $j \in \mathscr{S}$, we solve the linear program

$$\max_{\mu} \left\{ \mu^T \left(h_j - T_j \hat{x}_1\right) : \mu^T W_j \leq 0,\ |\mu| \leq 1 \right\}, \tag{32}$$

where the absolute value $|\cdot|$ is taken componentwise, until $\mu_j^T \left(h_j - T_j x\right) > 0$ for some j. For some j such that $\mu_j^T \left(h_j - T_j x\right) > 0$, we add

$$\mu_j^T \left(h_j - T_j x\right) \leq 0 \tag{33}$$

to the constraint set of G_j for all $j \in \mathscr{S}$.

This cut eliminates a candidate first-stage solution $\hat{x}_1$ that does not have a feasible recourse for the subproblem. We highlight, however, that infeasibility can

be guaranteed to be eliminated only if it is detected at the root node of the branch-and-bound tree of the subproblem (22).

2.3.2 Optimality Cuts

We assume that $(\hat{x}_1, \hat{y}_1) \in G_1$ is a feasible subproblem solution that is also feasible with respect to the original problem (18a–c). For $j \in \mathscr{S}$, we solve the linear program

$$\max_{\pi} \left\{ \pi^T \left(h_j - T_j \hat{x}_1 \right) : \pi^T W_j \leq q_j \right\}. \tag{34}$$

The optimality cut is generated by

$$c^T x + \sum_{j \in \mathscr{S}} \pi_j^T \left(h_j - T_j x \right) \leq z_{\text{UB}} \tag{35}$$

for a given best upper bound z_{UB} and added to the constraint set of G_j for all $j \in \mathscr{S}$.

We note that the optimality cut (35) is parameterized by the best-known upper bound z_{UB} and thus can be tightened as better upper bounds are obtained. In other words, the optimality cut seeks to eliminate first-stage solutions that go above a known upper bound.

Procedure 1 summarizes the Benders-type cutting-plane procedure for solving the Lagrangian subproblems (22) by adding the valid inequalities (33) and (35). The procedure can also be applied to other scenario decomposition methods. Moreover, the procedure terminates in a finite number of steps (see Theorem 1 in (Kim and Zavala 2015)).

Procedure 1 Cutting-Plane Procedure for Lagrangian Subproblems (CPSub)

Require: λ^k
for all $s \in \mathscr{S}$ **do**
  repeat
    SOLVE subproblem (22) to obtain $D_j(\lambda^k)$ and (x_j^k, y_j^k) for λ^k
    $isFeasible \leftarrow true$
    for all $j' \in \mathscr{S} \backslash \{j\}$ **do**
      SOLVE feasibility cut generator (32) to obtain $\mu_{j'}$ for x_j^k
      if $\mu_{j'}^T (h_{j'} - T_{j'} x_j^k) > 0$ **then**
        ADD feasibility cut (33) to all the subproblems (22)
        $isFeasible \leftarrow false$
      end if
    end for
  until $isFeasible = true$
  UPDATE z_{UB} by solving (18) for fixed x_j^k
  GENERATE optimality cut (35) by solving (34) for x_j^k and for all $j \in \mathscr{S}$
  ADD optimality cut (35) to all the subproblems (22)
end for

3 Numerical Example

We present a numerical example of the stochastic unit commitment model. We also illustrate how to solve the model using open-source software packages. In this model, thermal power generators are scheduled over a day. The schedules are subject to uncertainty in wind power generation. We use a modified IEEE 118-bus system with 54 generators, 118 buses, and 186 transmission lines provided in (Kim and Zavala 2015). We assume that 17 of the 54 generators are allowed to start on demand (second stage) whereas the other generators should be scheduled in advance (first stage). We also consider three identical wind farms, each consisting of 120 wind turbines. The demand load is 3,095 MW on average, with a peak of 3,733 MW. The wind power generation level is 494 MW on average, with a peak of 916 MW for the 64 scenarios generated. Figure 1 shows the 64 scenarios (gray lines) of wind power generation and the mean levels (red lines). We used real wind speed data predicted from the observations of 31 weather stations in the state of Illinois.

Table 1 presents the size of the stochastic unit commitment instances with 4, 8, 16, 32, and 64 scenarios. The first stage has 10,727 constraints and 2,592 variables including 864 integer variables, and the second stage has 27,322 constraints and 9,072 variables including 432 integer variables.

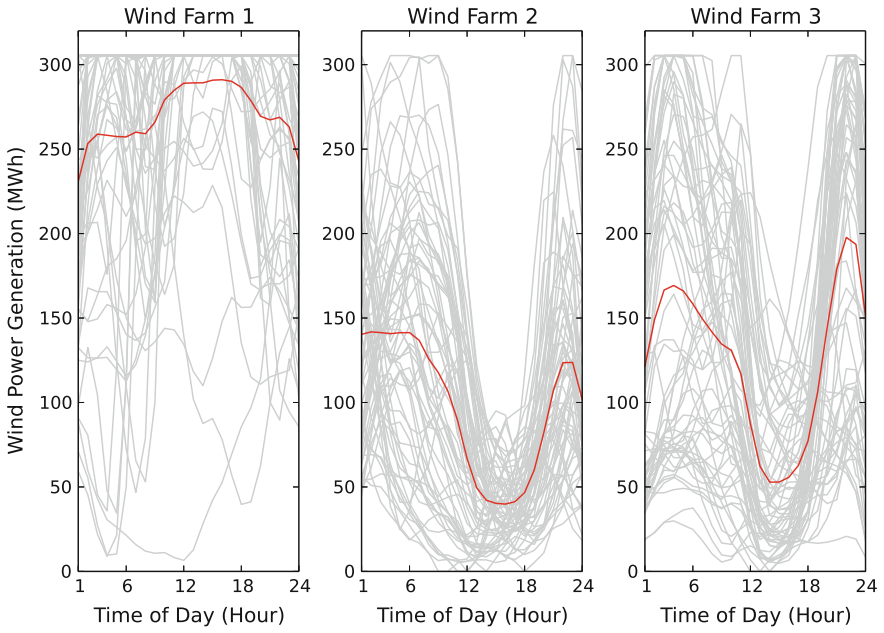


Fig. 1 Wind power generation scenarios for wind farms considered in the stochastic unit commitment model

Table 1 Characteristics of stochastic unit commitment instances

# Scenarios	# Rows	# Columns	# Integers
4	120,015	38,880	2,592
8	229,303	75,168	4,320
16	447,879	147,744	7,776
32	885,031	292,896	14,688
64	1,759,335	583,200	28,512

Table 2 Numerical results for the extensive form of the stochastic unit commitment problems

# Scenarios	Branch-and-cut nodes	Upper bound	Lower bound	Gap (%)	Time (sec.)
4	88,831	907035.3	906089.9	0.01	6632
8	58,235	904068.1	903567.8	0.05	>21,600
16	3505	900806.1	900200.3	0.07	>21,600
32	9	907536.0	901759.8	0.64	>21,600
64	1	∞	33605.4	∞	>21,600

We also solved the extensive form of the problems using `SCIP` (Achterberg 2009) (version 3.1.1) with a 6 h time limit and a 0.01 % optimality gap tolerance. Table 2 presents the computational results for the extensive form. None of the instances except the 4-scenario instance can reach the 0.01 % optimality gap within the 6 h limit. Moreover, the 64-scenario instance finds a poor lower bound without an upper bound.

To avoid scalability issues, we solved the stochastic UC problems using the scenario decomposition methods presented in Sect. 2. Specifically, we used the `DSP` package (Kim and Zavala 2015) (version 0.2.0) and `PySP` (Watson et al. 2012) in the `Pyomo` package (version 4.0.9682) for the dual decomposition and progressive hedging methods, respectively. `DSP` and `PySP` solve the subproblems with 0.01 % of optimality gap using `SCIP` (version 3.1.1) and `CBC` (https://projects.coin-or.org/Cbc) (version 2.8), respectively. We could not use `SCIP` for `PySP` because the interface of `SCIP` for `PySP` does not allow us to set the optimality gap tolerance. All computations were run on *Blues*, a 310-node computing cluster at Argonne National Laboratory. Each node on the Blues cluster has two 8-core 2.6 GHz Xeon processors and 64 GB of RAM.

3.1 Lower and Upper Bounds from Dual Decomposition

We first solved the problems using the dual decomposition method implemented in the `DSP` solver. We used the interior-point cutting-plane method (IPCPM) for solving the master problem (27a–b). We also used the Benders-type cutting-plane

Table 3 Numerical results for stochastic unit commitment problem from DSP

# Scenarios	Iterations	Upper bound	Lower bound	Gap (%)	Time/Iter. (sec.)	Total time (sec.)
4	1	907046.1	906979.1	<0.01	551	551
8	1	904006.6	903953.5	<0.01	667	667
16	1	900706.3	900650.7	<0.01	764	764
32	19	903227.7	903149.9	<0.01	390	7424
64	16	895118.0	895044.6	<0.01	895	14,320

procedure (CPSub). DSP is capable of solving a problem in parallel with up to S^2 number of computing cores (i.e., up to 4,096 cores for the 64-scenario instance). However, in our numerical example, we used only up to 2,048 cores on the Blues cluster.

Table 3 shows that upper and lower bounds are obtained with <0.01 % optimality gap for all the problem instances. Moreover, the DSP solutions terminate after the first iteration for the 4-, 8-, and 16-scenario instances because the CPSub procedure excluded infeasible first-stage solutions and tightened the scenario subproblems. We note that the solution time per iteration does not increase with the number of scenarios because of parallelization of DSP.

We present results for the different variants of dual decomposition discussed in Sect. 2. We compare: the subgradient method (DDSub), IPCPM *without* CPSub (IPCPM), and IPCPM *with* CPSub (DSP). Figure 2 shows the best upper bound and

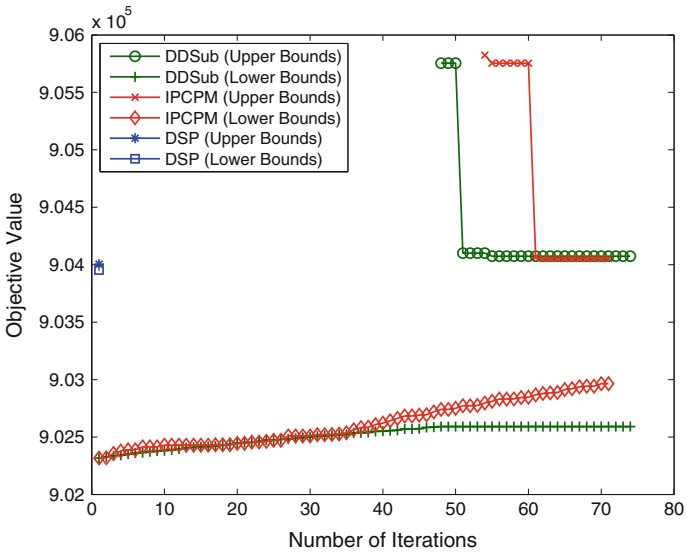


Fig. 2 Upper bounds and lower bounds obtained with DSP with and without Procedure 1 and with the subgradient method

the best lower bound obtained at each iteration. As can be seen, DSP obtains upper and lower bounds with <0.01 % duality gap at the first iteration and achieved termination, whereas IPCPM and DDSub are not able to find upper bounds for the first 53 iterations and the first 47 iterations, respectively. The results clearly indicate that *the problem does not have relatively complete recourse.* Moreover, DSP finds tighter lower bounds than do DDSub and IPCPM because of the ability to tighten the subproblems by Procedure 1. The figure also shows that IPCPM finds better lower and upper bounds than does DDSub.

3.2 Upper Bounds from Progressive Hedging

We now present results for the progressive hedging method implemented in the `PySP` package. The subproblem solutions are parallelized using the `Pyro` package (Pyro: Python remote objects 2015) that provides capabilities for distributed computing. In our numerical example, we set the initial perturbation vector $\rho := 1.0$ and later adjust it in proportion to the objective function coefficient (Watson et al. 2012). We also set the `–enable-ww-extension` option, which enables additional heuristics for finding feasible solutions (Watson and Woodruff 2011). We note that one could tune the parameters and devise other heuristics.

Table 4 presents the results from solving the stochastic UC problems by the progressive hedging method in `PySP`. `PySP` finds tight upper bounds for the 4-, 8-, 16-, and 32-scenario problems. For the 64-scenario problem, however, `PySP` cannot find an upper bound after 53 iterations and 6 h of solution time. We also note that parallelization keeps the time per iteration constant as we increase the number of scenarios. The number of iterations, however, also tends to increase as we increase the number of scenarios, because the number of nonanticipativity constraints increases.

Table 4 Numerical results for stochastic unit commitment problem from `PySP`

# Scenarios	Iterations	Upper bound	Time/Iter. (sec.)	Total time (sec.)
4	5	907042.5	245	1224
8	13	904041.2	203	2645
16	26	900712.1	240	6228
32	20	903355.7	603	12,069
64	53	∞	407	>21,600

4 Summary

We have presented a stochastic unit commitment model, which can be formulated as a two-stage stochastic mixed-integer programming problem. The first-stage schedules slow generation units a day ahead under uncertain wind power production. The second stage dispatches electricity production in order to meet demand load, while scheduling fast generation units. The problem is challenging because it is a large-scale mixed-integer programming problem. Moreover, standard Benders decomposition cannot be applied because the recourse function is nonconvex and discontinuous on the first-stage variables. As a result, we have presented scenario decomposition methods, which split the original problem into the scenario number of subproblems and solve the smaller subproblems. This decomposability naturally leads to parallel algorithms capable of running on high-performance computing systems. We have presented dual decomposition and progressive hedging with various enhancement techniques that incorporate Benders-type cuts in the scenario decomposition framework. In the numerical example, we used an IEEE 118-bus system with 64 scenarios of wind power productions. We solved the stochastic unit commitment problem using open-source software packages, `DSP` and `PySP`, that implement parallel dual decomposition and progressive hedging, respectively. The numerical results show that state-of-the-art solution methods are able to solve a large-scale stochastic unit commitment problem. However, more research is needed to scale up these methods and address a large number of scenarios.

Acknowledgments This material is based upon work supported by the U.S. Department of Energy, Office of Science, under contract number DE-AC02-06CH11357. We gratefully acknowledge the computing resources provided on Blues, a high-performance computing cluster operated by the Laboratory Computing Resource Center at Argonne National Laboratory. We thank Julie Bessac for providing wind speed prediction data.

References

Achterberg, T.: Scip: solving constraint integer programs. Mathematical Programming Computation **1**(1), 1–41 (2009)

Ahmed, S., Tawarmalani, M., Sahinidis, N.V.: A finite branch-and-bound algorithm for two-stage stochastic integer programs. Mathematical Programming **100**(2), 355–377 (2004)

Bertsekas, D.P., Scientific, A.: Convex Optimization Algorithms. Athena Scientific (2015)

Birge, J.R., Dempster, M.A., Gassmann, H.I., Gunn, E.A., King, A.J., Wallace, S.W.: A standard input format for multiperiod stochastic linear programs. IIASA Laxenburg Austria (1987)

CarøE, C.C., Schultz, R.: Dual decomposition in stochastic integer programming. Operations Research Letters **24**(1), 37–45 (1999)

Fisher, M.L.: An applications oriented guide to Lagrangian relaxation. Interfaces **15**(2), 10–21 (1985)

Forrest, J.: Cbc. URL https://projects.coin-or.org/Cbc

Gade, D., Küçükyavuz, S., Sen, S.: Decomposition algorithms with parametric gomory cuts for two-stage stochastic integer programs. Mathematical Programming **144**(1–2), 39–64 (2014)

Geoffrion, A.M.: Lagrangean relaxation for integer programming. Springer (1974)
Gondzio, J., Gonzalez-Brevis, P., Munari, P.: New developments in the primal–dual column generation technique. European Journal of Operational Research **224**(1), 41–51 (2013)
Huchette, J., Lubin, M., Petra, C.: Parallel algebraic modeling for stochastic optimization. In: Proceedings of the 1st First Workshop for High Performance Technical Computing in Dynamic Languages, pp. 29–35. IEEE Press (2014)
Kim, K., Mehrotra, S.: A two-stage stochastic integer programming approach to integrated staffing and scheduling with application to nurse management. Operations Research **63**(6), 1431–1451 (2015)
Kim, K., Zavala, V.M.: Algorithmic innovations and software for the dual decomposition method applied to stochastic mixed-integer programs. Optimization Online (2015)
Lemaréchal, C.: Lagrangian relaxation. In: Computational combinatorial optimization, pp. 112–156. Springer (2001)
Lubin, M., Martin, K., Petra, C.G., Sandkç, B.: On parallelizing dual decomposition in stochastic integer programming. Operations Research Letters **41**(3), 252–258 (2013)
Märkert, A., Gollmer, R.: Users Guide to ddsip–A C package for the dual decomposition of two-stage stochastic programs with mixed-integer recourse (2014)
Pyro: Python remote objects (2015). URL http://pythonhosted.org/Pyro4/
Rockafellar, R.T., Wets, R.J.B.: Scenarios and policy aggregation in optimization under uncertainty. Mathematics of operations research **16**(1), 119–147 (1991)
Sen, S., Higle, J.L.: The C^3 theorem and a D^2 algorithm for large scale stochastic mixed-integer programming: set convexification. Mathematical Programming **104**(1), 1–20 (2005)
Sherali, H.D., Fraticelli, B.M.: A modification of benders' decomposition algorithm for discrete subproblems: An approach for stochastic programs with integer recourse. Journal of Global Optimization **22**(1–4), 319–342 (2002)
U.S. Energy Information Administration: Monthly energy review. Tech. rep., U.S. Energy Information Administration (2015)
Watson, J.P., Woodruff, D.L.: Progressive hedging innovations for a class of stochastic mixed-integer recsource allocation problems. Computational Management Science **8**(4), 355–370 (2011)
Watson, J.P., Woodruff, D.L., Hart, W.E.: PySP: modeling and solving stochastic programs in Python. Mathematical Programming Computation **4**(2), 109–149 (2012)
Zhang, M., Küçükyavuz, S.: Finitely convergent decomposition algorithms for two-stage stochastic pure integer programs. SIAM Journal on Optimization **24**(4), 1933–1951 (2014)