

Enterprise-wide optimization for industrial demand side management: fundamentals, advances, and perspectives

Qi Zhang^{a,*}, Ignacio E. Grossmann^a

^aCenter for Advanced Process Decision-making, Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Abstract

The active 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, enterprise-wide optimization (EWO) for industrial DSM has emerged as a major area of interest for both researchers and practitioners. In this work, we introduce the reader to the fundamentals of power system economics, provide a definition of DSM that reflects more strongly the consumer's perspective, and present a comprehensive review of existing works on EWO for industrial DSM. The review is organized into four parts, which correspond to the four main challenges that we identify as: (1) accurate modeling of operational flexibility, (2) integration of production and energy management, (3) decision-making across multiple time and space scales, and (4) optimization under uncertainty. Finally, we highlight research gaps and future opportunities in this area.

Keywords: enterprise-wide optimization, demand side management, electricity markets

1. Introduction

In the light of increasing volatility in electricity price and availability, demand side management (DSM), which refers to the active management of electricity demand, has been recognized as an effective approach to improving power grid performance and consumer benefits. For electricity consumers, DSM constitutes the opportunity to benefit from financial incentives by adjusting their electricity consumption. Especially for the process industry, which is a major electricity consumer, DSM is becoming increasingly critical for maintaining profitability.

Industrial DSM integrates production and energy management, which requires detailed understanding of the production process as well as knowledge about power system economics. Since decision-making in industrial DSM is highly complex, a systems approach has to be taken in order to find optimal solutions while considering all relevant options and critical constraints. We can establish such a systematic framework for DSM by leveraging the methodologies in mathematical modeling and optimization that have been developed in the area of process systems engineering (PSE). In particular, enterprise-wide optimization (EWO) (Grossmann, 2005, 2012) has proven to be very effective in industrial DSM. Typically, electricity

*Corresponding author

Email address: qz@andrew.cmu.edu (Qi Zhang)

prices vary on an hourly basis, and price fluctuation patterns are observed over the course of a day or a week, resulting in the need of scheduling models with adequate time resolutions. Moreover, many DSM opportunities have to be evaluated on a long-term basis, often involving investment decisions and under considerable uncertainty. EWO is capable of capturing all these considerations as it considers operational, tactical, and strategic decision-making at the plant and supply chain levels.

This review is an extension of our recently published book chapter (Zhang and Grossmann, 2016) that provides an overview of existing works on planning and scheduling for industrial DSM. In this work, besides extending the review to EWO and placing a stronger focus on the modeling aspects, we also introduce the reader to the mechanisms of deregulated electricity markets, which are in several ways distinctly different from typical commodity markets in the process industries. We emphasize that understanding how electricity markets operate is crucial for exploring DSM opportunities.

The remainder of this paper is organized as follows. In Section 2, we present the fundamentals of electricity markets, followed by a definition of DSM and a classification of DSM activities in Section 3. Section 4 describes the special characteristics of DSM in power-intensive process industries, and identifies four major challenges in EWO for industrial DSM: (1) accurate modeling of operational flexibility, (2) integration of production and energy management, (3) decision-making across multiple time and space scales, and (4) optimization under uncertainty. A state-of-the-art review of existing works addressing these four topics is presented in Sections 5–8. Finally, in Section 9, we close with concluding remarks on future opportunities and challenges in the area of industrial DSM.

This review reflects the expanding frontier of PSE, which is increasingly addressing interdisciplinary problems that reach beyond traditional chemical engineering. We dedicate this paper to Professor Roger Sargent, the founder and intellectual leader of PSE, whose pioneering work laid the technical foundation for the advances in this area and continues to inspire us to tackle major engineering challenges of our time.

2. Fundamentals of electricity markets

In a power grid, electrical energy is generated by power plants, sent through high-voltage transmission lines to substations that transform it to lower voltage, and then distributed to consumers (see Figure 1 for a schematic of the North American power system). In a regulated environment, generation, transmission, and distribution are operated by one entity, which is referred to as a vertically integrated utility; this utility holds the monopoly for the supply of electricity. In contrast, in a deregulated environment, different parts are owned and operated by different participants, and electricity markets provide a platform that fosters competition among those participants. The deregulation of electricity markets started in the 1990s (mainly in the UK and the U.S.). Since then, experience has shown that allowing more competition in electricity markets can significantly increase grid efficiency.

In this section, we focus on aspects of electricity markets that are important for the understanding of demand side management, and omit topics of less interest for the end-consumer, such as the capacity market and financial transmission rights. For more details, we refer to the excellent book on power system economics by Kirschen and Strbac (2004). Note that due to different historical developments, the organization of electricity markets as well as the applied terminology varies among different geographical regions; in this paper, we mostly make reference to the U.S. electricity markets.

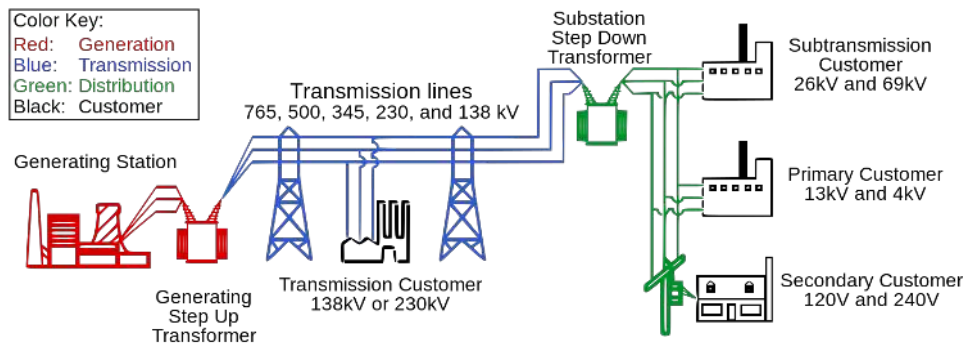


Figure 1: Schematic of an electric power system in North America (FERC, 2004).

2.1. Participants in deregulated electricity markets

As it is common in commodity markets, electricity markets are divided into a wholesale market and a retail market, which differ in the amount of power that can be traded. Figure 2 depicts the flow of electrical energy in a power grid and shows in which markets the involved entities participate. In the wholesale market, generating companies sell their produced electricity to retailers or large consumers that are allowed to participate in this market. The transmission is operated by transmission companies, which own transmission assets such as cables and transformers. In the retail market, retailers sell the electricity that they purchase in the wholesale market to smaller end-consumers; here, the distribution is handled by distribution companies. Depending on the market structure and the level of deregulation, some of the entities could be subsidiaries of the same company; for instance, a utility from which small consumers purchase electricity could be a retailer and at the same time own the distribution network.

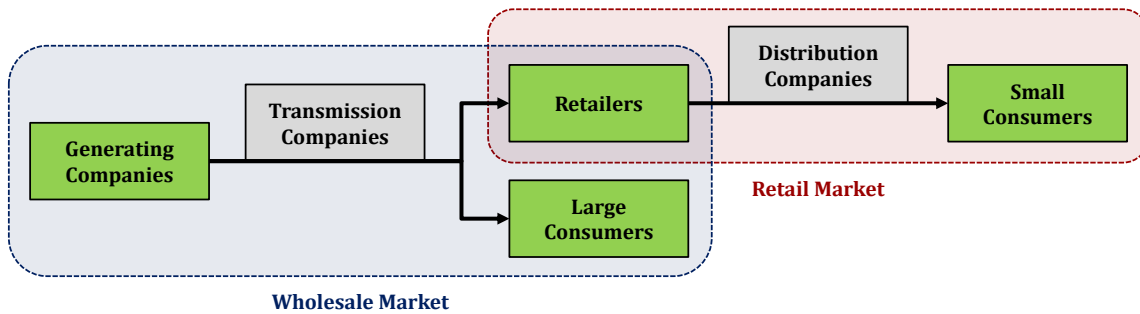


Figure 2: Participants in the electricity wholesale and retail markets.

Typically, a grid operator, commonly referred to as the independent system operator (ISO), controls the transmission network and operates the wholesale market. The ISO's primary responsibility is to maintain grid reliability, i.e. make sure that the required amount of electrical energy can be supplied at all times.

2.2. Electricity as a commodity and as a service

Electricity markets are established based on the premise that electrical energy can be treated as a commodity; however, it is distinctly different from most other commodities mainly because it is linked with a physical system that functions much faster than any market. In a power grid, electricity supply and

demand have to match at all times. Since it is difficult to store electrical energy, it has to be produced at the same time as it is consumed. Delays are unacceptable; hence, the grid has to react very quickly to any changes and disturbances.

The mechanisms of an electrical energy market, in which electricity is traded as a commodity, are not fast enough to ensure the balance of electricity supply and demand in real time. Instead, we rely on backup capacities that can be dispatched when supply-demand mismatch occurs. These backup capacities do not represent actual power generation but a guarantee to dispatch when required; hence, they are traded as services in the so-called ancillary services market. Ancillary services are categorized according to the required response time for dispatch. For example, regulation services, which are used for frequency and voltage control, have to be dispatched within seconds; longer response times (10–30 minutes) are acceptable in the case of operating reserves, which are needed in larger quantities during contingency events.

Ancillary services are typically provided by flexible generating facilities such as gas-fired power plants. Providing ancillary services limits the operation of the generation process. Consider the following example: We have a generator with a capacity of 90 MW. We trade 50 MW in the electrical energy market and the remaining 40 MW as operating reserve in the ancillary services market. The actual amount of power that we generate will be between 50 and 90 MW; however, we cannot schedule to produce more than 50 MW because then we would not be able to dispatch 40 MW of operating reserve when requested. As a result, a generating facility that provides operating reserves cannot operate at full capacity, which often leads to reduced generation efficiency.

2.3. Bilateral trading and trading in electricity pools

Electricity is traded either through bilateral contracts or through auctions in competitive electricity pools. In bilateral trading, a buyer and a seller enter into a contract without the involvement of a third party. The most common kinds of bilateral contracts are customized long-term contracts, which are negotiated privately between both parties; hence, the terms of such contracts are flexible.

Trading in pools is a very unusual form of commodity trading, but is well-established in electricity markets. In the operation of power systems, power produced by all generators is pooled on its way to the consumers, which creates desirable economies of scale. Pooling is possible because units of electrical energy produced by different generators are indistinguishable. Following the same idea, competitive electricity pools have been created in which electricity is traded through auctions.

The auction-based trading is managed by the ISO and works as follows: Generating companies submit bids, which state at which prices they are willing to supply which quantities of power for the period under consideration. These bids are ranked in order of increasing price; the resulting inverse supply curve showing the bid price as a function of cumulative bid quantity is shown in Figure 3. Similarly, consumers submit offers to purchase certain amounts of power at certain prices. Offers are ranked in order of decreasing price, resulting in the inverse demand curve shown in Figure 3. The market is cleared at the price where the supply and demand curves intersect, denoted by P^* . The quantity of electrical energy produced and consumed in the grid is then Q^* . All generators with bid prices below or at the clearing price are instructed to produce the amount of power corresponding to their accepted bids, and all consumers with offer prices above or at the clearing price are allowed to draw the corresponding amount of energy from the system. Different markets are cleared at different frequencies. For instance, a day-ahead market is cleared once every day while a real-time market is typically cleared every five minutes.

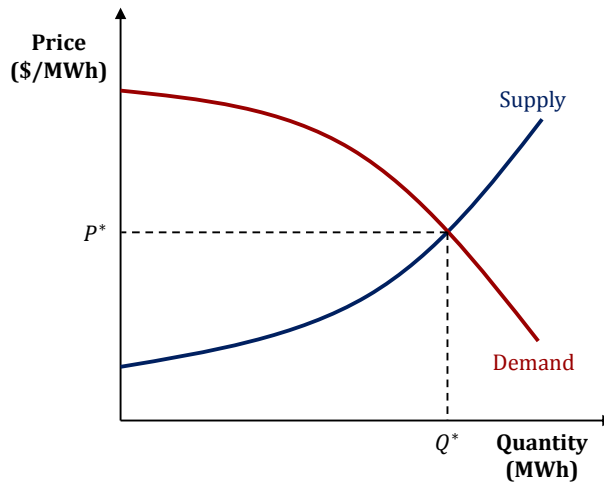


Figure 3: Inverse electricity supply and demand curves. The clearing price P^* and quantity Q^* mark the intersection of the two curves.

The market clearing price represents the price of producing one additional unit of electrical energy and is therefore called the system marginal price (SMP). Irrespective of their bids, generators are paid the SMP for every unit of energy that they produce; similarly, all consumers pay the SMP for every unit of energy that they consume. Noncontrollable generation resources, such as solar and wind power plants, usually bid with a price of zero in order to guarantee the clearing of their bids while knowing that they will sell at a most likely nonzero SMP.

2.4. Locational marginal pricing

Having one single SMP that applies to all generators and consumers is only appropriate if electricity transmission is unconstrained, which in practice is seldom true due to the physical limitations of the transmission network. Because of transmission losses and congestion, an additional unit of electrical energy in a specific geographical region may have to be provided solely by local generators. As a result, the marginal price depends on the location at which the energy is produced or consumed and is thus called locational marginal price (LMP).

The system operator clears the market by solving an optimization problem that considers transmission constraints and maximizes the economic welfare generated by the system while satisfying all security considerations. In this way, economic efficiency is achieved, and LMPs are obtained as a byproduct of the optimization. LMPs can vary significantly; they can even take negative values, which occur especially frequently in the real-time market.

3. Demand side management

Traditionally, in power systems engineering, the focus has been on improving the power supply for given demands (loads) in the grid. Only in recent years, it has been recognized that there is also considerable flexibility on the demand side that can be leveraged to effectively change the load curve according

to the power grid’s needs. The management of flexible loads is generally referred to as demand side management. 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 and public policy.

3.1. Two perspectives on DSM

The opportunities in DSM can be viewed from two distinct perspectives: the system operator’s perspective and the electricity consumer’s perspective. On the one hand, the system 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 to provide the flexibility to quickly react to supply-demand mismatch 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 markets, 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 4 illustrates the activities on the system 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 system operator’s perspective. The system operator’s main task 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 the consumers’ load adjustment capabilities (Albadi and El-Saadany, 2008; Mohsenian-Rad et al., 2010); however, such simplified models are often not sufficiently accurate for capturing the real DSM potential that each consumer has as well as the costs occurring on the consumer side when implementing different DSM measures.

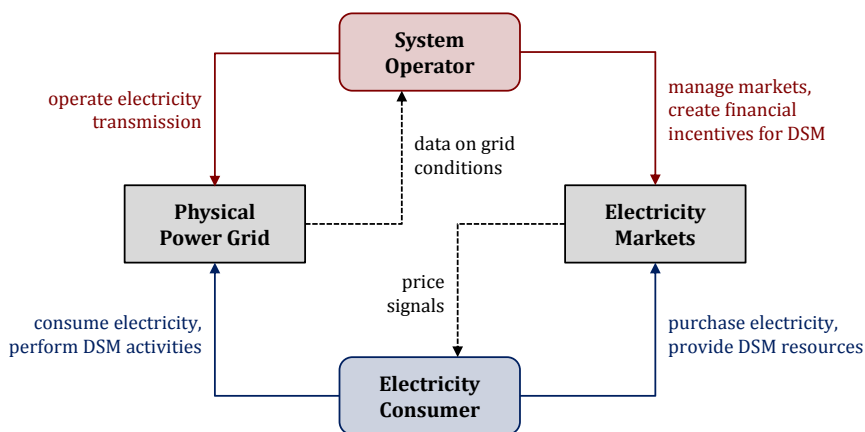


Figure 4: Two perspectives on DSM, which requires both the system operator and the consumer to interact through physical grid operations and the electricity markets.

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 in which production levels are controlled (Samad and Kiliccote, 2012).

3.2. Definition and classification of DSM activities

When it comes to defining DSM, most descriptions represent the system operator's perspective. For example, in a report released by The World Bank (Charles River Associates, 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 system 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 system 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.

Figure 5 shows a general classification of DSM activities, which consist of various DSM programs introduced by the system 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 Associates, 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 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 system operator such that these capacities can be dispatched to maintain grid stability or in times of emergency. The system 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. With dispatchable DR,

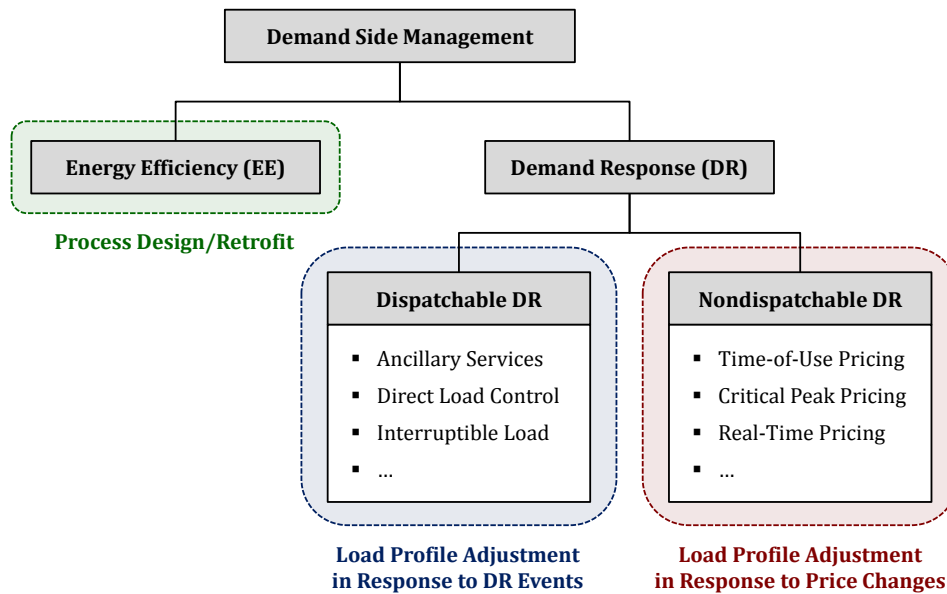


Figure 5: Classification of DSM activities. Rectangular boxes depict DSM programs introduced by the system operator, rounded rectangles indicate measures to be taken by the consumer.

consumers participate in the ancillary services market. Nondispatchable DR resources are not controlled by the system operator; instead, consumers choose to adjust their power consumption profiles based on price signals from the electrical energy 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.

EE is 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, a high degree of flexibility in the consumer’s process is required in order to react to time-sensitive prices and unforeseeable DR events.

4. Enterprise-wide optimization for 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.

4.1. *Special considerations in industrial DSM*

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 other 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 bilateral power contracts that 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 \(2012\)](#) report five real-world case studies in which industrial DSM has been successfully implemented and has generated considerable cost savings.

4.2. *Major challenges in 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 PSE community in the late 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 Harjunkski 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 advances in EWO made over the last few decades; however, there are unique challenges that require special attention. The four main challenges that we see in EWO 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 the ability of a process 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 schedule is determined, then 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. Therefore, 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. *Multiscale decision-making*

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. Furthermore, DSM should be considered at the supply chain level in order to account for location-dependent electricity market conditions and constraints on product distribution.

4. *Optimization under uncertainty*

Due to the high level of uncertainty in the power grid, electricity prices are extremely volatile. Also, all dispatchable DR activities are intrinsically uncertain because the consumer does not know in advance when the system operator will request the dispatch of those DR services. Additional uncertainties exist on the production side, e.g. regarding product demand and processing time. 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. Sorted in thematic groups, the table shows various features of the different models; we will refer back to this overview in the next sections.

5. 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 five years.

5.1. 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 production environments (continuous and batch).

Table 1 further 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 the 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.

5.2. 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.

Table 1: Overview of reviewed works, listed in chronological order, with model features.

	Continuous Production Batch Production	Air Separation Aluminum	Cement	Chlor-Alkali Flour /Pulp Machining Steel	Power Contracts Onsite Generation Dispatchable DR	Planning & Scheduling Scheduling & Control Supply Chain	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 Kebblis (2007)	✓	✓					✓				✓	
Babu and Ashok (2008)	✓		✓				✓					✓
Castro et al. (2009)	✓		✓				✓	✓			✓	
Yusta et al. (2010)	✓			✓			✓			✓		
Nolde and Morari (2010)	✓			✓			✓				✓	
Hait 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. (2014a)	✓	✓				✓	✓	✓			✓	
Mitra et al. (2014b)	✓	✓				✓	✓	✓			✓	
Zhang and Hug (2014)	✓		✓		✓		✓	✓			✓	
Shrouf et al. (2014)	✓			✓			✓				✓	
Marchetti et al. (2014)	✓	✓				✓	✓	✓			✓	
Zhang et al. (2015a)	✓	✓			✓	✓	✓	✓			✓	
Zhang and Hug (2015)	✓		✓		✓		✓	✓			✓	
Hadera et al. (2015a)	✓			✓	✓	✓		✓			✓	
Hadera et al. (2015b)	✓			✓	✓	✓		✓			✓	
Tan et al. (2015)	✓			✓			✓				✓	
Zhang et al. (2015b)	✓	✓			✓		✓				✓	
Zhang et al. (2016e)	✓	✓			✓		✓				✓	
Zhang et al. (2016b)	✓	✓			✓		✓	✓			✓	
Zhang et al. (2016d)	✓	✓			✓		✓	✓			✓	
Hadera et al. (2016)	✓			✓	✓	✓		✓			✓	
Pattison et al. (2016b)	✓	✓				✓	✓				✓	
Zhang et al. (2016f)	✓	✓				✓	✓				✓	
Zhang et al. (2016a)	✓	✓			✓	✓	✓	✓			✓	
Pattison et al. (2016a)	✓	✓				✓	✓				✓	
Zhou et al. (2017)	✓	✓					✓				✓	

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; Shah 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 Castro 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 Kebliis (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 Kebliis (2007) to improve the tightness of the formulation, and develop additional constraints to impose restrictions on the transitions between different modes. As illustrated in Figure 6, the dynamics of a plant in terms of its mode transition behavior can be visualized with a mode transition graph. The arcs in the graph show the directions of the allowed transitions with the corresponding operational constraints.

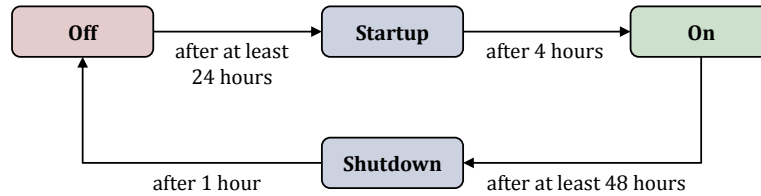


Figure 6: Example of a mode transition graph, which shows the different operating modes and all possible transitions with the corresponding operational constraints.

To model mode transitions, we first have to introduce a binary variable y_{mt} , which equals 1 if mode m is selected in time period t , and enforce that one and only one mode is selected in each time period as follows:

$$\sum_m y_{mt} = 1 \quad \forall t. \quad (1)$$

Next, the binary mode transition variable $z_{mm't}$ is introduced; $z_{mm't}$ has to equal 1 if and only if the plant switches from mode m to mode m' at time point t , which corresponds to the following condition:

$$y_{m,t-1} = 1 \wedge y_{m't} = 1 \Leftrightarrow z_{mm',t-1} = 1 \quad \forall (m, m') \in TR, t \quad (2)$$

where TR is the set of possible mode transitions. This condition can be modeled as follows:

$$y_{m,t-1} + y_{m't} - 1 \leq z_{mm',t-1} \quad \forall (m, m') \in TR, t \quad (3a)$$

$$z_{mm',t-1} \leq y_{m,t-1} \quad \forall (m, m') \in TR, t \quad (3b)$$

$$z_{mm',t-1} \leq y_{m't} \quad \forall (m, m') \in TR, t \quad (3c)$$

which can be replaced with the following set of constraints (Sahinidis and Grossmann, 1991; Erdirik-Dogan and Grossmann, 2008), derived by using a different interpretation of condition (2):

$$y_{mt} = \sum_{m' \in TR_m^f \cup \{m\}} z_{m'm,t-1} \quad \forall m, t \quad (4a)$$

$$y_{m,t-1} = \sum_{m' \in TR_m^t \cup \{m\}} z_{mm',t-1} \quad \forall m, t \quad (4b)$$

where TR_m^f denotes the set of modes from which mode m can be directly reached; similarly, TR_m^t denotes the set of modes which can be directly reached from mode m . The formulation given by Eqs. (4) is tighter and has fewer constraints than Eqs. (3). Mitra et al. (2013) show that Eqs. (4) can be further improved by subtracting Eq. (4b) from Eq. (4a), resulting in the following formulation

$$\sum_{m' \in TR_m^f} z_{m'm,t-1} - \sum_{m' \in TR_m^t} z_{mm',t-1} = y_{mt} - y_{m,t-1} \quad \forall m, t \quad (5)$$

which has fewer constraints and also fewer variables since the binary variables for self-transitions have been eliminated. Interestingly, if the process has only two states, say "on" and "off", because $y_{on,t} + y_{off,t} = 1$, Eq. (5) reduced to:

$$z_{off,on,t-1} - z_{on,off,t-1} = y_{on,t} - y_{on,t-1} \quad \forall t \quad (6)$$

which is widely known from the standard unit commitment formulation (Arroyo and Conejo, 2000).

Using the binary variables, mode transition constraints can now be derived, such as:

$$y_{m't} \geq \sum_{k=1}^{\theta_{mm'}} z_{mm',t-k} \quad \forall (m, m') \in TR, t \quad (7a)$$

$$z_{mm',t-\bar{\theta}_{mm'm''}} = z_{m'm''t} \quad \forall (m, m', m'') \in SQ, t \quad (7b)$$

where Eq. (7a) 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. (7b) is 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, 2016a,b,d,e,f; Zhou et al., 2017). In the most recent development, Zhang et al. (2016e) have further generalized the model such that it can be used to represent continuous process networks. It should be mentioned that in practice, the constraints on transition times are obtained from a careful analysis of the underlying control system. Depending on the process dynamics, varying operating conditions can have a significant impact on the response of the system as, for example, Cao et al. (2015) have demonstrated for an industrial air separation plant.

5.3. 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; Méndez 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 Figure 7, 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:

$$\sum_e \sum_{p \in P_e} y_{tpe} = 1 \quad \forall t \neq |T| \quad (8a)$$

$$\hat{\tau}_t \geq \sum_e \sum_{p \in P_e} L_{ep} y_{tpe} \quad \forall t \neq |T| \quad (8b)$$

$$\hat{\tau}_t + \sum_{i \in \bar{I}_e} \frac{\bar{\mu}_{ri} \xi_{it}}{\rho_i^{\max}} \leq \sum_{p \in P_e} U_{ep} y_{tpe} + M \left(1 - \sum_{i \in \bar{I}_e} \bar{\mu}_{ri} z_{it} \right) \quad \forall t \neq |T|, e, r \in R^{\text{TC}} \quad (8c)$$

where the binary variable y_{tpe} equals 1 if some tasks are processed within price period $p \in P_e$ during time interval t . Eq. (8a) states that there is one active price period during each time interval. Eqs. (8b) and (8c) 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 , ρ_i^{\max} is the maximum processing rate of task i , M is a big-M parameter, z_{it} is a binary variable that equals 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.

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 Figure 8, which shows six different task-interval overlap

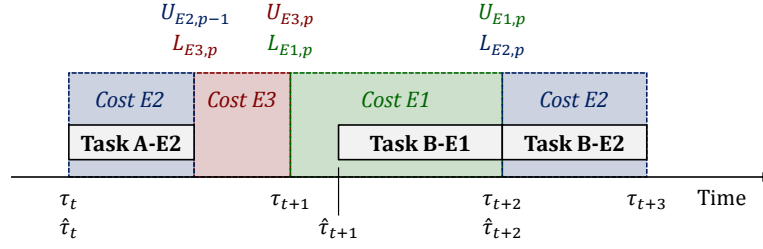


Figure 7: Illustration of continuous time representation with timing constraints (Castro et al., 2009).

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.

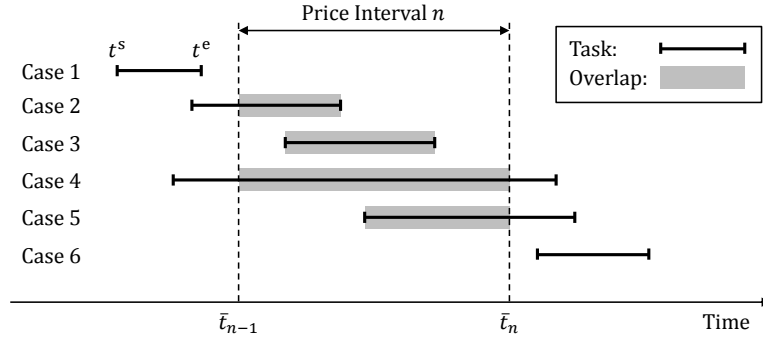


Figure 8: Six possible task-interval overlap cases (Nolde and Morari, 2010).

Haït and Artigues (2011) propose a reformulation of the task-interval overlap constraints, which reduces the number of binary variables and constraints. In the reformulation, t_l^s and t_l^e denote the starting and ending times of task l , respectively. Price interval n starts at time \bar{t}_{n-1} and ends at time \bar{t}_n . Haït and Artigues (2011) introduce the binary variable y_{ln}^s that equals 1 if task l starts before or during price interval n , which is enforced by the following constraints:

$$t_l^s \geq \bar{t}_n (1 - y_{ln}^s) \quad \forall l \in L, n = 1, \dots, |N| - 1 \quad (9a)$$

$$t_l^s \leq \bar{t}_n + |T|(1 - y_{ln}^s) \quad \forall l \in L, n = 1, \dots, |N| - 1 \quad (9b)$$

$$y_{ln}^s \leq y_{l,n+1}^s \quad \forall l \in L, n = 1, \dots, |N| - 1 \quad (9c)$$

$$y_{l,|N|}^s = 1 \quad \forall l \in L \quad (9d)$$

where L , N , and T denote the sets of tasks, price intervals, and time periods, respectively. Similarly, binary variables associated with the ending times, y_{ln}^e , are introduced, and the corresponding constraints are also included in the formulation. The following constraints are then added to model the six different task-interval overlap cases:

$$0 \leq o_{ln} \leq (\bar{t}_n - \bar{t}_{n-1})(y_{ln}^s - y_{l,n-1}^e) \quad \forall l \in L, n \in N \quad (10a)$$

$$\sum_n o_{ln} = t_l^e - t_l^s \quad \forall l \in L \quad (10b)$$

$$o_{ln} \geq t_l^e - \bar{t}_n + (\bar{t}_n - \bar{t}_{n-1})y_{l,n-1}^s - |T|(1 - y_{ln}^e) \quad \forall l \in L, n \in N \quad (10c)$$

$$o_{ln} \geq (\bar{t}_n - \bar{t}_{n-1})(y_{l,n-1}^s - y_{ln}^e) \quad \forall l \in L, n \in N \quad (10d)$$

$$o_{ln} \geq \bar{t}_n(1 - y_{l,n-1}^s) - t_l^s - (\bar{t}_n - \bar{t}_{n-1})y_{ln}^e \quad \forall l \in L, n \in N \quad (10e)$$

where o_{ln} denotes the amount of time in which task l overlaps with price interval n . Note that o_{ln} can only be greater than zero if $y_{ln}^s = 1$ and $y_{l,n-1}^e = 0$; hence, Eqs. (10a) and (10b) are sufficient for modeling Cases 1, 3, and 6. In order to also cover the remaining cases, Eqs. (10c)–(10e) are required, which respectively incorporate Cases 2, 4, and 5.

A continuous-time model incorporating the same task-interval overlap 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 Trespalcios, 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 have 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.

5.4. 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. Therefore, [Ierapetritou et al. \(2002\)](#) develop 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 computation 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. [Pattison et al. \(2016b\)](#) solve an optimization problem involving a system of differential algebraic equations (DAEs) by using a sequential approach in which an NLP is solved at each iteration and discontinuities are dealt with in the integration step. However, it should be noted that if a simultaneous dynamic optimization approach ([Biegler, 2007](#)) were applied, the resulting model would be an MINLP.

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. \(2015a, 2016e\)](#). 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., 2016c](#)). 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.

5.5. Main insights from case studies

A number of case studies have been presented in the reviewed references, some using real-world data from industry. 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](#)).

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 relationship is not that obvious when dispatchable DR is involved. While it is still true that flexibility decreases with increasing plant utilization, higher utilization also implies higher target production levels, which allow more dispatchable DR to be provided. For instance, [Zhang et al. \(2016d\)](#) find in their case studies that the largest cost savings are achieved at a high, yet not maximum level of plant utilization (around 95%) if the plant provides interruptible load.

6. 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, on-site 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. (2015a, 2016) 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. (2016b) 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. (2016e) propose a more general block contract formulation which can be expressed as the following disjunction:

$$\bigvee_{b \in B_c} \left[\begin{array}{l} X_{cbt} \\ H_{cb't} = H_{cb'}^{\max} \quad \forall b' \in B_c, b' < b \\ H_{cb} \leq H_{cb}^{\max} \\ H_{cb't} = 0 \quad \forall b' \in B_c, b' > b \end{array} \right] \quad \forall c, t \quad (11a)$$

$$\bigvee_{b \in B_c} X_{cbt} \quad \forall c, t \quad (11b)$$

$$X_{cbt} \in \{\text{true}, \text{false}\} \quad \forall c, b \in B_c, t \quad (11c)$$

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. The Boolean variable X_{cbt} is true if block b is the highest block reached for contract c in time period t . Disjunction (11a) 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 (11b), one and only one X_{cbt} has to be

true. By applying the hull reformulation (Balas, 1985), Eqs. (11) can be transformed into a set of mixed-integer linear constraints and incorporated into the MILP scheduling model. Figure 9 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 Figure 9b, penalties incur for both over- and underconsumption.

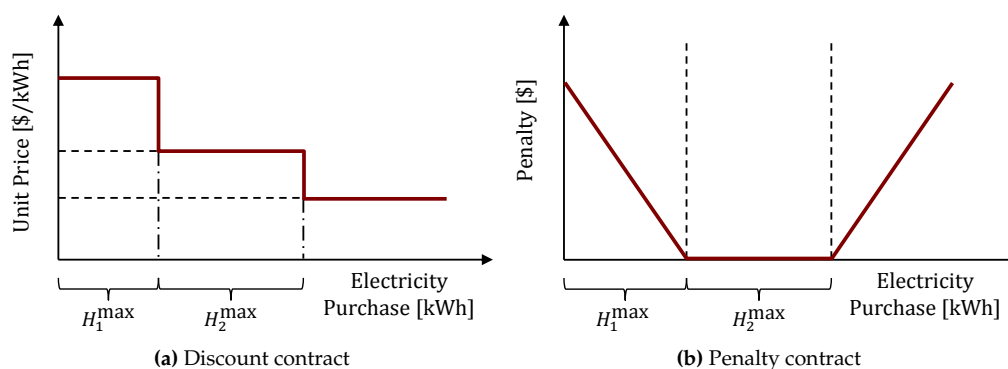


Figure 9: Discount and penalty contracts modeled as block contracts (Zhang et al., 2016e).

Vujanic et al. (2012), Zhang and Hug (2014, 2015), and Zhang et al. (2015a,b, 2016d) develop systematic approaches for the optimization of dispatchable DR activities, e.g. the provision of ancillary 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 Section 8.

7. Multiscale decision-making

Electricity prices and DR opportunities are time- as well as location-sensitive, where the latter is due to locational marginal pricing and plant-specific power contracts. All the works reviewed in the previous two sections consider the short-term scheduling of a single plant. If we go beyond that problem and consider long-term decisions and interactions at the supply chain level, multiscale optimization is required.

7.1. 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.

Mitra et al. (2014a) 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 than the 8760 time periods that would be required to represent hourly discretization over one year.

Zhang et al. (2016a) apply a similar approach, but represent each season of the year by two 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. (2014a), a consistent inventory profile throughout the year is achieved since inventory can be carried over from one season to the next, which allows the consideration of seasonal changes in product demand. Using this modeling idea, which is illustrated in Figure 10, Zhang et al. (2016a) solve a long-term electricity procurement problem over a one-year planning horizon.

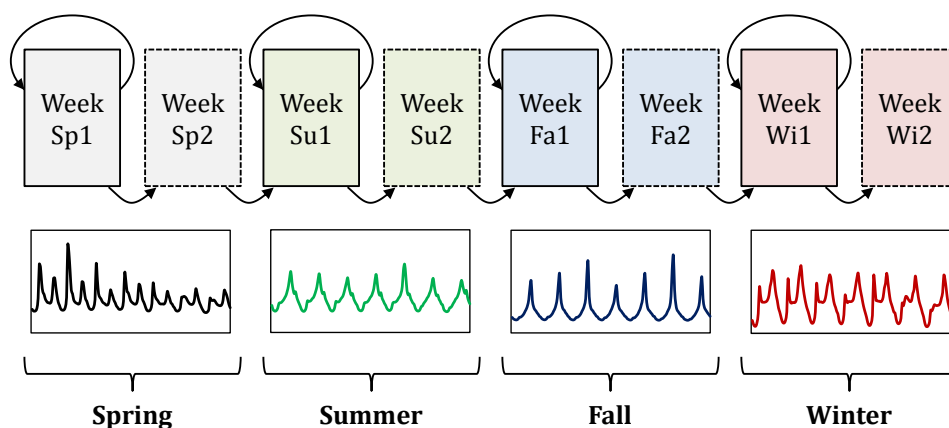


Figure 10: Multiscale model with each of the four seasons of the year represented by two weeks, where the first has a cyclic schedule and the second is noncyclic (Zhang et al., 2016a).

Pattison et al. (2016b) integrate process control, i.e. operations at an even smaller time scale, with scheduling when optimizing a cryogenic air separation process subject to time-sensitive electricity pricing. The authors propose a scheduling approach based on scheduling-oriented low-order dynamic models identified from historical process data. Compared to applying a high-dimensional detailed dynamic model, the computational effort can be significantly reduced while preserving essential information required for the optimal schedule to be feasible from a dynamic point of view. In a subsequent work, Pattison et al. (2016a) apply the proposed model in a moving horizon closed-loop scheduling framework in order to account for real-time changes in market conditions and disturbances.

7.2. Across multiple space scales

The vast majority of the reviewed works consider the operation of a single production plant. In order to capture potential synergies between multiple process units, Zhang et al. (2016e) propose a scheduling

model for optimizing power-intensive process networks. The proposed model has been adopted by Zhang et al. (2016a) to simultaneously optimize the long-term electricity procurement strategy for multiple air separation plants. A similar approach is taken by Zhou et al. (2017) who also consider the interactions between multiple air separation units. Although Zhou et al. (2017) consider demand rather than electricity price changes, the proposed model can be easily extended to be applied in a classical DSM setting.

At the supply chain level, the interactions between different production sites are more complex since also the distribution of raw materials and products has to be considered. Moreover, the impact of different site-specific electricity prices has to be taken into account. Marchetti et al. (2014) consider the simultaneous optimization of production and distribution operations in power-intensive industrial gas supply chains. The proposed integrated model includes prespecified routes that can be assigned to available trucks for the transportation of liquid products. Results from case studies show that coordinating production and distribution can significantly improve economic performance. However, the proposed model only considers a time resolution of 12 h and is therefore incapable of capturing electricity prices that change with higher frequency.

Acknowledging the restrictions of the model proposed by Marchetti et al. (2014), Zhang et al. (2016f) develop a multiscale production routing framework involving two time grids, a fine one for production scheduling and a coarse one for distribution planning. In order to solve large instances of the problem, an iterative MILP-based heuristic approach is proposed, which solves the integrated MILP model with a restricted set of candidate routes at each iteration and dynamically updates the set of candidate routes for the next iteration. The results of an extensive computational study show that the proposed algorithm finds high-quality solutions in reasonable computation times, and in large instances, it significantly outperforms a standard two-phase heuristic approach and a solution strategy involving a one-time heuristic pre-generation of candidate routes similar to the one proposed by Marchetti et al. (2014). By applying the multiscale framework, the authors solve an industrial test case that considers a real-world industrial gas supply chain with 2 plants, approximately 240 customers, 20 vehicles, and a planning horizon of 4 weeks.

8. 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.

8.1. Modeling electricity price uncertainty

Ierapetritou et al. (2002) present a two-stage stochastic programming framework in which the electricity prices for the first three 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 three days. While the first-stage decisions are related to the first three 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:

$$\min_x \max_{c,z} \sum_t \alpha_t E_t \quad (12a)$$

$$\text{s.t. } \alpha_t = \sum_{t'} \zeta_{tt'} c_{t'} \quad \forall t \quad (12b)$$

$$c_{tk}^{\min} z_{tk} \leq \bar{c}_{tk} \leq c_{tk}^{\max} z_{tk}, z_{tk} \in \{0, 1\} \quad \forall t, k \quad (12c)$$

$$c_t = \sum_k \bar{c}_{tk}, \sum_k z_{tk} = 1 \quad \forall t \quad (12d)$$

$$\sum_t z_{tk} \leq \Gamma_k \quad \forall k \quad (12e)$$

$$\text{s.t. } x \in X \quad (12f)$$

where x denotes the operational decision variables including E_t , which is the amount of electricity purchased in time period t . Given the scheduling constraints that are generally represented by Eq. (12f), this bilevel formulation minimizes the cost of purchasing electricity for the worst case within the uncertainty set defined by Eqs. (12b)–(12e). As shown in Eq. (12b), 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 (Düzgün and Thiele, 2015). Eqs. (12c)–(12d) ensure that only one price range k is chosen in each time period t , and the parameter Γ_t in Eq. (12e) 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.

8.2. 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. \(2014a\)](#) 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. \(2014a\)](#) 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. Because of the large size of the stochastic program, [Mitra et al. \(2014b\)](#) further propose a tailored hybrid bilevel decomposition algorithm to solve industrial-scale instances of the problem.

[Zhang et al. \(2016b\)](#) 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; (4) multicut Benders decomposition ([Birge and Louveaux, 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 has led to the following remarkable insight: 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.

8.3. 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. Thus, 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. \(2015a\)](#) 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:

$$U(R) = \left\{ w : \left(D_t = R_t w_t, 0 \leq w_k \leq 1 \forall k = 1, \dots, t, \sum_{k=1}^t w_k \leq \Gamma_t \right) \forall t \right\} \quad (13)$$

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 static robust optimization is its 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 decision rules. [Zhang et al. \(2016d\)](#) 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. (13), the following affine decision rule is proposed:

$$Q_{it} = \bar{Q}_{it} + \sum_{k=t-\zeta}^t q_{itk} w_k \quad (14)$$

where Q_{it} is the actual amount of product i produced in time period t , which consists of the target production rate \bar{Q}_{it} (for the case where reserve demand is zero) and a linear recourse term. The recourse term allows Q_{it} to depend on uncertain parameters realized in previous time periods, which makes this a multistage formulation. The parameter ζ can be any integer between zero and $t - 1$; the greater ζ is, the more past information is considered in the recourse. Note that the coefficients q_{itk} are 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 studies show that significant financial benefits can be achieved by providing

interruptible load. The results further demonstrate the value of recourse as cost savings increase with the increasing ζ ; however, this higher flexibility in the recourse comes at the cost of increased model size (more variables q_{itk}) and computation time.

9. Conclusions and perspectives

As indicated by the review presented in Sections 4–8, EWO for industrial DSM is a relatively new research area, and there is considerable need for further development. In the following, we highlight some remaining challenges and opportunities in this area:

- Almost all existing works on planning and scheduling for DSM use linear process models. While many of these MILP models are adequate for capturing major plant operations, it would certainly be desirable to incorporate more accurate nonlinear process models.
- Power contracts can involve complex price structures that cannot be considered in the current models. Also, existing DSM programs are constantly revised and new ones are created. Therefore, 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 Section 7. Appropriate models for such problems have the potential to significantly improve economic performance.
- Although many high-impact strategic and tactical decisions, such as decisions regarding facility location, capacity expansion, and fleet sizing, are made at the supply chain level, little work has been done on optimization of power-intensive supply chains. However, this topic will become increasingly important with the trend toward more tightly integrated supply chains and more volatile electricity markets.
- It will be worth exploring synergies from joint efforts of 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 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.

Matching electricity supply and demand in the power grid has become increasingly challenging due to high fluctuations in electricity demand and increasing penetration of intermittent renewable energy into the electricity supply mix. This development calls for more effective DSM measures in order to maintain grid efficiency and reliability. DSM problems often require interdisciplinary approaches, and the PSE community is particularly well-equipped to make significant contributions in this area.

Acknowledgments

The authors gratefully acknowledge the financial support from the National Science Foundation under Grant No. 1159443.

References

- Albadi, M. H., El-Saadany, E. F., 2008. A summary of demand response in electricity markets. *Electric Power Systems Research* 78 (11), 1989–1996.
- Arroyo, J. M., Conejo, A. J., 2000. Optimal Response of a Thermal Unit to an Electricity Spot Market. *IEEE Transactions on Power Systems* 15 (3), 1098–1104.
- Artigues, C., Lopez, P., Haït, 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, C. A., 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., El Ghaoui, L., Nemirovski, A., 2009. *Robust Optimization*. Princeton University Press, New Jersey.
- Ben-Tal, A., Goryashko, A., Guslitzer, E., Nemirovski, A., 2004. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming* 99 (2), 351–376.
- Biegler, L. T., 2007. An overview of simultaneous strategies for dynamic optimization. *Chemical Engineering and Processing: Process Intensification* 46 (11), 1043–1053.
- Birge, J. R., Louveaux, F., 2011. *Introduction to Stochastic Programming*, 2nd Edition. Springer Science+Business Media.
- Birge, J. R., Louveaux, F. V., 1988. A multicut algorithm for two-stage stochastic linear programs. *European Journal of Operational Research* 34 (3), 384–392.
- Cao, Y., Swartz, C. L. E., Baldea, M., Blouin, S., 2015. Optimization-based assessment of design limitations to air separation plant agility in demand response scenarios. *Journal of Process Control* 33, 37–48.
- 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 Journal* 60 (6), 2083–2097.
- Castro, P. M., Harjunkski, I., Grossmann, I. E., 2009. New Continuous-Time Scheduling Formulation for Continuous Plants under Variable Electricity Cost. *Industrial & Engineering Chemistry Research* 48 (14), 6701–6714.
- Castro, P. M., Harjunkski, I., Grossmann, I. E., 2011. Optimal scheduling of continuous plants with energy constraints. *Computers & Chemical Engineering* 35 (2), 372–387.
- Castro, P. M., Sun, L., Harjunkski, I., 2013. Resource-Task Network Formulations for Industrial Demand Side Management of a Steel Plant. *Industrial & Engineering Chemistry Research* 52, 13046–13058.
- Charles River Associates, 2005. *Primer on Demand-Side Management*. Tech. Rep. February, The World Bank.
- Conejo, A. J., Nogales, F. J., Arroyo, J. M., 2002. Price-Taker Bidding Strategy Under Price Uncertainty. *IEEE Transactions on Power Systems* 17 (4), 1081–1088.
- Daryanian, B., Bohn, R. E., Tabors, R. D., 1989. Optimal Demand-Side Response to Electricity Spot Prices for Storage-Type Customers. *IEEE Transactions on Power Systems* 4 (3), 897–903.

- Ding, Y. M., Hong, S. H., Li, X. H., 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.
- 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, M. A., Grossmann, I. E., 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.
- Erdirik-Dogan, M., Grossmann, I. E., 2008. Simultaneous planning and scheduling of single-stage multi-product continuous plants with parallel lines. *Computers & Chemical Engineering* 32 (11), 2664–2683.
- 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, J. W., 2011. A new approach to scheduling in manufacturing for power consumption and carbon footprint reduction. *Journal of Manufacturing Systems* 30 (4), 234–240.
- FERC, 2004. Final Report on the August 14, 2003 Blackout in the United States and Canada: Causes and Recommendations. Tech. rep., Federal Energy Regulatory Commission.
- FERC, 2010. National Action Plan on Demand Response. Tech. rep., Federal Energy Regulatory Commission.
- Floudas, C. A., Lin, X., 2004. Continuous-time versus discrete-time approaches for scheduling of chemical processes: a review. *Computers & Chemical Engineering* 28 (11), 2109–2129.
- Gellings, C. W., 1985. The concept of demand-side management for electric utilities. *Proceedings of the IEEE* 73 (10), 1468–1470.
- Gellings, C. W., 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, S. C., 1981. A Review of Production Scheduling. *Operations Research* 29 (4), 646–675.
- Grossmann, I. E., 2005. Enterprise-wide optimization: A new frontier in process systems engineering. *AIChE Journal* 51 (7), 1846–1857.
- Grossmann, I. E., 2012. Advances in mathematical programming models for enterprise-wide optimization. *Computers & Chemical Engineering* 47, 2–18.
- Grossmann, I. E., Trespalacios, F., 2013. Systematic Modeling of Discrete-Continuous Optimization Models through Generalized Disjunctive Programming. *AIChE Journal* 59 (9), 3276–3295.
- 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, 117–136.
- Hadera, H., Labrik, R., Sand, G., Engell, S., Harjunkoski, I., 2016. An Improved Energy-Awareness Formulation for General Precedence Continuous-Time Scheduling Models. *Industrial and Engineering Chemistry Research* 55 (5), 1336–1346.
- Hadera, H., Wide, P., Harjunkoski, I., Sand, G., Engell, S., 2015b. 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 & Chemical Engineering* 35 (12), 3044–3047.
- Harjunkoski, I., Maravelias, C. T., Bongers, P., Castro, P. M., Engell, S., Grossmann, I. E., Hooker, J., Méndez, C., Sand, G., Wasick, J., 2014. Scope for industrial applications of production scheduling models and solution methods. *Computers & Chemical Engineering* 62, 161–193.
- Ierapetritou, M. G., 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, M. H., Kebblis, M. F., 2007. Operations planning with real time pricing of a primary input. *Computers & Operations Research* 34 (3), 848–867.
- Kirschen, D. S., 2003. Demand-Side View of Electricity Markets. *IEEE Transactions on Power Systems* 18 (2), 520–527.
- Kirschen, D. S., Strbac, G., 2004. *Fundamentals of Power System Economics*. John Wiley & Sons.
- Kondili, E., Pantelides, C. C., Sargent, R. W. H., 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. G., 2008. Process scheduling under uncertainty: Review and challenges. *Computers & Chemical Engineering* 32 (4-5), 715–727.
- Maravelias, C. T., 2012. General Framework and Modeling Approach Classification for Chemical Production Scheduling. *AIChE Journal* 58 (6), 1812–1828.
- Maravelias, C. T., Sung, C., 2009. Integration of production planning and scheduling: Overview, challenges and opportunities. *Computers & Chemical Engineering* 33 (12), 1919–1930.
- Marchetti, P. A., Gupta, V., Grossmann, I. E., Cook, L., Valton, P.-M., Singh, T., Li, T., André, J., 2014. Simultaneous production and distribution of industrial gas supply-chains. *Computers & Chemical Engineering* 69, 39–58.
- Méndez, C. A., Cerdá, J., Grossmann, I. E., Harjunkski, 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.
- Mitra, S., Grossmann, I. E., Pinto, J. M., 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, I. E., Pinto, J. M., 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., Pinto, J. M., Grossmann, I. E., 2014a. 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.
- Mitra, S., Pinto, J. M., Grossmann, I. E., 2014b. Optimal multi-scale capacity planning for power-intensive continuous processes under time-sensitive electricity prices and demand uncertainty. Part II: Enhanced hybrid bi-level decomposition. *Computers and Chemical Engineering* 65, 102–111.
- Mitra, S., Sun, L., Grossmann, I. E., 2013. Optimal scheduling of industrial combined heat and power plants under time-sensitive electricity prices. *Energy* 54, 194–211.
- Mohsenian-Rad, A. H., Wong, V. W. S., 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, M. A., Watson, D. S., Kilicote, 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 & Chemical Engineering* 34 (11), 1899–1903.
- Pantelides, C. C., 1994. Unified Frameworks for Optimal Process Planning and Scheduling. In: *Foundations of computer-aided process operations*. New York, pp. 253–274.
- Pattison, R. C., Touretzky, C. R., Harjunkski, I., Baldea, M., 2016a. Moving Horizon Closed-Loop Production Scheduling Using Dynamic Process Models Richard. *AIChE Journal*.
- Pattison, R. C., Touretzky, C. R., Johansson, T., Harjunkski, I., Baldea, M., 2016b. Optimal Process Operations in Fast-Changing Electricity Markets: Framework for Scheduling with Low-Order Dynamic Models and an Air Separation Application. *Industrial & Engineering Chemistry Research* 55 (16), 4562–4584.
- Reklaitis, G. V., 1982. Review of scheduling of process operations. *AIChE Symposium Series* 78 (214), 119–133.
- Rockafellar, R. T., Uryasev, S., 2000. Optimization of Conditional Value-at-Risk. *Journal of risk* 2, 21–42.
- Sahinidis, N. V., Grossmann, I. E., 1991. Reformulation of Multiperiod MILP Models for Planning and Scheduling of Chemical Processes. *Computers & Chemical Engineering* 15 (4), 255–272.
- Samad, T., Kilicote, S., 2012. Smart grid technologies and applications for the industrial sector. *Computers & Chemical Engineering* 47, 76–84.
- Shah, N., Pantelides, C. C., Sargent, R. W. H., 1993. A general algorithm for short-term scheduling of batch operations - II. Computational issues. *Computers & Chemical Engineering* 17 (2), 229–244.
- Shrouf, F., Ordieres-Meré, J., García-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, C. T., 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, Y. Y., Huang, Y. L., Liu, S. X., 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.
- Verderame, P. M., Elia, J. A., Li, J., Floudas, C. A., 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, K. R., 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, N. H., 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, J. M., Torres, F., Khodr, H. M., 2010. Optimal methodology for a machining process scheduling in spot electricity markets. *Energy Conversion and Management* 51 (12), 2647–2654.
- Zareipour, H., Cañizares, C. A., 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, A. M., Grossmann, I. E., Sundaramoorthy, A., Pinto, J. M., 2016a. Long-term electricity procurement for large industrial consumers under uncertainty. Submitted for publication.
- Zhang, Q., Cremer, J. L., Grossmann, I. E., Sundaramoorthy, A., Pinto, J. M., 2016b. Risk-based integrated production scheduling and electricity procurement for continuous power-intensive processes. *Computers & Chemical Engineering* 86, 90–105.
- Zhang, Q., Grossmann, I. E., 2016. Planning and Scheduling for Industrial Demand Side Management: Advances and Challenges. In: Martin, M. (Ed.), *Alternative Energy Sources and Technologies: Process Design and Operation*. Springer, pp. 383–414.
- Zhang, Q., Grossmann, I. E., Heuberger, C. F., Sundaramoorthy, A., Pinto, J. M., 2015a. Air Separation with Cryogenic Energy Storage: Optimal Scheduling Considering Electric Energy and Reserve Markets. *AIChE Journal* 61 (5), 1547–1558.
- Zhang, Q., Grossmann, I. E., Sundaramoorthy, A., Pinto, J. M., 2016c. Data-driven construction of Convex Region Surrogate models. *Optimization and Engineering* 17 (2), 289–332.
- Zhang, Q., Morari, M. F., Grossmann, I. E., Sundaramoorthy, A., Pinto, J. M., 2016d. An adjustable robust optimization approach to provision of interruptible load by continuous industrial processes. *Computers & Chemical Engineering* 86, 106–119.
- Zhang, Q., Sundaramoorthy, A., Grossmann, I. E., Pinto, J. M., 2016e. A discrete-time scheduling model for continuous power-intensive process networks with various power contracts. *Computers & Chemical Engineering* 84, 382–393.
- Zhang, Q., Sundaramoorthy, A., Grossmann, I. E., Pinto, J. M., 2016f. Multiscale production routing in multicommodity supply chains with complex production facilities. Submitted for publication.
- 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.
- Zhang, X., Hug, G., Kolter, Z., Harjunkoski, I., 2015b. Industrial Demand Response by Steel Plants with Spinning Reserve Provision. In: *Proceedings of the IEEE North American Power Symposium*.
- Zhou, D., Zhou, K., Zhu, L., Zhao, J., Xu, Z., Shao, Z., Chen, X., 2017. Optimal scheduling of multiple sets of air separation units with frequent load-change operation. *Separation and Purification Technology* 172, 178–191.