

Risk-based integrated production scheduling and electricity procurement for continuous power-intensive processes

Qi Zhang^a, Jochen L. Cremer^b, Ignacio E. Grossmann^{a,*}, Arul
Sundaramoorthy^c, Jose M. Pinto^d

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

^b*Faculty of Mechanical Engineering, RWTH Aachen University, 52056 Aachen, Germany*

^c*Prazair, Inc., Business and Supply Chain Optimization R&D, Tonawanda, NY 14150,
USA*

^d*Prazair, Inc., Business and Supply Chain Optimization R&D, Danbury, CT 06810, USA*

Abstract

For optimal operation of power-intensive plants, production scheduling and electricity procurement have to be considered simultaneously. In addition, uncertainty needs to be taken into account. For this purpose, an integrated stochastic mixed-integer linear programming model is developed that considers the two most critical sources of uncertainty: spot electricity price, and product demand. Conditional value-at-risk is incorporated into the model as a measure of risk. Furthermore, scenario reduction and multicut Benders decomposition are implemented to solve large-scale real-world problems. The proposed model is applied to an illustrative example as well as an industrial air separation case. The results show the benefit from stochastic optimization and the effect of taking a risk-averse rather than a risk-neutral approach. An interesting insight from the analysis is that in risk-neutral optimization, accounting for electricity price uncertainty does not yield significant added value; however, in risk-averse optimization, modeling price uncertainty is crucial for obtaining good solutions.

Keywords: Production scheduling, electricity procurement, demand response, stochastic programming, conditional value-at-risk

*Corresponding author

Email address: grossmann@cmu.edu (Ignacio E. Grossmann)

1. Introduction

Due to high fluctuations in electricity demand and increasing penetration of intermittent renewable energy into the electricity supply mix, it is becoming increasingly difficult to match electricity demand and supply in the power grid (Hand et al., 2012). As a result, electricity prices have become extremely volatile and difficult to predict, which poses immense challenges to power-intensive industries, such as air separation, aluminum, and chlor-alkali manufacturing.

For large industrial electricity consumers, there are two ways of dealing with uncertainty in electricity price: (1) dynamically adjust the production schedule to changes in the spot price, i.e. shift the electricity load to lower-price periods, which is also referred to as demand response (Charles River Associates, 2005); (2) remove price uncertainty by signing power contracts with agreed fixed prices. Both strategies can be very effective in reducing the electricity cost, but they also have their limitations and drawbacks. A plant's capability for demand response is limited by the flexibility of the production process, which has to be carefully evaluated in order to avoid detrimental disruptions caused by sudden changes in the plant operation. Power contracts provide fixed electricity prices; however, this reduction in risk usually comes at the cost of higher expected average prices. Moreover, power contracts require the consumers to commit themselves in advance to the amount that they are going to purchase for a certain period of time. This commitment reduces the consumers' demand response opportunities since there is less room for adjustments in response to real-time price changes. Hence, there is a trade-off between purchasing power from contracts and from the spot market.

It is clear that often only a combination of the two aforementioned strategies will lead to the best result. Here, the major challenge in the decision-making is uncertainty. This uncertainty does not only occur in the electricity price; another source of uncertainty that has a possibly even greater impact on the production schedule is product demand. Major operational decisions and decisions regarding the commitment to power contracts have to be made before the actual spot electricity price and product demand are known for the time horizon of interest. There is only limited room for reactive actions as soon as these decisions are made. Therefore, it is crucial to account for these uncertainties in the decision-making process.

The high potential impact of large industrial electricity consumers participating in demand response is widely acknowledged (Paulus and Borggrefe, 2011; Samad and Kiliccote, 2012; Merkert et al., 2014) and has been the focus of increased research efforts in recent years. Deterministic scheduling models considering demand response have been proposed for various industrial power-intensive

processes such as steelmaking (Ashok, 2006; Castro et al., 2013), cement production (Castro et al., 2009, 2011), electrolysis (Babu and Ashok, 2008), and air separation (Ierapetritou et al., 2002; Karwan and Kebli, 2007; Mitra et al., 2012; Zhang et al., 2014). However, these models do not consider power contracts but rather assume that power can only be purchased from the spot market.

Although many works have addressed problems involving power contracts from an electricity producer’s or retailer’s point of view (Conejo et al., 2008; Lima et al., 2015; Carrión et al., 2007a; Hatami et al., 2009), the literature is scarce in papers considering the consumers’ perspective. Conejo et al. (2005) solve a medium-term electricity procurement problem that considers a set of bilateral contracts, hourly changing spot prices, and the possibility of producing electricity with an onsite generating facility. The self-generated power can be used for own consumption or sold to the spot market. A subsequent work (Conejo and Carrión, 2006) addresses a similar problem for a shorter time horizon, while considering cost volatility by using an estimate of the covariance of the spot price. While the models proposed in these two papers are deterministic, Carrión et al. (2007b) apply stochastic programming to explicitly model uncertainty in electricity prices; furthermore, the conditional value-at-risk (CVaR) is included in the model as a measure of risk, which is used to show the clear trade-off between expected cost and risk. A similar trade-off is shown by Zare et al. (2010) who apply the concept of information gap decision theory to evaluate the robustness of a solution against high spot prices or high procurement costs. Beraldi et al. (2011) consider the short-term electricity procurement problem involving bilateral contracts and the day-ahead market; here, a stochastic programming model is solved in a rolling-horizon fashion.

In all the works reviewed in the previous paragraph, the consumer’s electricity demand profile is assumed to be known and therefore fixed. This implies that a separate production scheduling problem has to be solved first in order to determine the electricity demand, which then can be used as input in the electricity procurement problem. However, this sequential approach is likely to be suboptimal since the production scheduling problem does not take the full electricity price information into account. Recently, Zhang et al. (2015) have introduced a general deterministic model that simultaneously optimizes the production schedule for a given power-intensive process and the electricity procurement strategy involving various power contracts.

In this work, in addition to taking an integrated approach, we apply stochastic programming to model uncertainty in both spot electricity price and product demand, which to the best of our knowledge has not been considered in this context before. Also, risk is taken into account by incorporating the CVaR into the model. Scenario reduction and multicut Benders decomposition are applied

to the resulting mixed-integer linear programming (MILP) model in order to solve large-scale industrial problems. We apply the proposed framework to an illustrative example and an industrial air separation case. In the analysis of the results, we emphasize the difference between risk-neutral and risk-averse optimization. Moreover, a comprehensive assessment of the added value gained from stochastic optimization compared to the deterministic approach is conducted by computing the value of stochastic solution for a large number of instances.

The remainder of this paper is organized as follows. In Section 2, the problem statement is presented before the MILP model is developed in Section 3. Section 4 describes the applied scenario generation and reduction techniques, while Section 5 outlines the multicut Benders decomposition algorithm. In Sections 6 and 7, the proposed model is applied to an illustrative example and an industrial case study, respectively. The main insights from the results are discussed in Section 8. Finally, in Section 9, we close with a summary and concluding remarks.

2. Problem statement

We consider a power-intensive continuously operated plant that can produce a given set of products. Inventory capacities exist for storable products, and additional products can be purchased at given costs. It is assumed that for fixed product demand, all production costs, besides the cost of electricity, are constant. In this way, for optimization purposes, the total operating cost only consists of the electricity cost and the cost of purchasing additional products.

Electricity can be purchased from the spot market, or from power contracts that have fixed pre-agreed electricity price and availability conditions. While purchases from the spot market can be made a day in advance (day-ahead) or on the spot (real-time), one has to commit to the electricity purchase from power contracts for a longer period of time, e.g. for one week.

The goal is to optimize the production and electricity procurement schedules in terms of expected cost and/or risk over a given time horizon. For this purpose, uncertainty in spot electricity price and product demand is considered. The decisions can be divided into two sets: one containing *here-and-now* decisions that have to be made at the beginning and cannot be changed over the course of the scheduling horizon, the other containing *wait-and-see* decisions that can be adjusted after realization of the uncertainty. In this problem, the here-and-now decisions are the following in each time period of the scheduling horizon:

- the mode of operation for the production process,
- and the amount of electricity purchased from each power contract,

where the mode of operation refers to the state in which the plant operates, e.g. “off”, “on”, or “startup”.

The wait-and-see decisions are:

- the actual production rates,
- the amounts of products stored,
- the amounts of products purchased,
- and the amount of electricity purchased from the spot market.

3. Model formulation

The stochastic scheduling problem is formulated as a mixed-integer linear program (MILP). The underlying production scheduling model is based on a formulation developed in previous works (Mitra et al., 2012, 2013; Zhang et al., 2014). Hence, we only provide brief descriptions of the production scheduling constraints and focus more on the electricity procurement and uncertainty modeling. Note that unless specified otherwise, all continuous variables in this model are constrained to be nonnegative. A list of indices, sets, parameters, and variables used in the model formulation is given in the Nomenclature section.

3.1. Uncertainty modeling strategy

We adopt a stochastic programming (Birge and Louveaux, 2011) approach to model the uncertainty in electricity price and product demand. In stochastic programming, uncertainty is represented by discrete scenarios, and decisions are made at different stages, which are defined such that realization of uncertainty is observed between two stages, and at each stage, actions depending on previous observations are taken.

Depending on the type of spot market, the real electricity price can be observed minutes (real-time), hours, or one day (day-ahead) in advance. Also, production rates can be adjusted in every time period, which makes this problem a multi-stage problem. However, the resulting multi-stage stochastic programming problem is extremely large and computationally intractable. Therefore, we approximate the multi-stage problem with a two-stage stochastic programming problem where we assume that all uncertainty for the entire scheduling horizon is realized right after the here-and-now decisions are made.

We define the set of product demand scenarios, S^D , with probabilities $\varphi_{\bar{s}}^D$ of each demand scenario \bar{s} ; similarly, S^P denotes the set of electricity price scenarios with $\varphi_{\hat{s}}^P$ being the probability of price scenario \hat{s} . Each pair of demand scenario \bar{s} and price scenario \hat{s} corresponds to a general scenario $s \in S$ with the

probability $\varphi_s = \varphi_s^D \varphi_s^P$. For the sake of brevity, we refer to the general scenario set S in most part of the model formulation.

3.2. Plant model

In this framework, we assume that the plant can operate in different operating modes. For each mode, the feasible operating region is defined by a polytope in the product space, and a linear electricity consumption function with respect to the production rates is given for each mode. At any point in time, the plant can only run in one operating mode. For a given operating mode, the operating point has to lie within the corresponding polytope. These relationships can be expressed by the following constraints:

$$PD_{its} = \sum_m \overline{PD}_{mits} \quad \forall i, t \in \overline{T}, s \quad (1a)$$

$$\overline{PD}_{mits} = \sum_{j \in J_m} \lambda_{mjts} v_{mji} \quad \forall m, i, t \in \overline{T}, s \quad (1b)$$

$$\sum_{j \in J_m} \lambda_{mjts} = y_{mt} \quad \forall m, t \in \overline{T}, s \quad (1c)$$

$$EU_{ts} = \sum_m \left(\delta_m y_{mt} + \sum_i \gamma_{mi} \overline{PD}_{mits} \right) \quad \forall t \in \overline{T}, s \quad (1d)$$

$$\sum_m y_{mt} = 1 \quad \forall t \in \overline{T} \quad (1e)$$

where J_m is the set of vertices of the polytope associated with mode m . The binary variable y_{mt} is 1 if mode m is selected in time period t . The amount of product i produced in time period t of scenario s is denoted by PD_{its} . Associated with PD_{its} is the disaggregated variable \overline{PD}_{mits} for mode m , which is expressed as a convex combination of the corresponding vertices, v_{mji} . The amount of electricity consumed, EU_{ts} , is a linear function of PD_{its} with a constant δ_m and coefficients γ_{mi} specific to the selected mode. Note that while y_{mt} is a first-stage variable, PD_{its} and EU_{ts} are second-stage variables that depend on scenario s .

3.3. Transition constraints

A transition occurs when the system changes from one operating point to another. In particular, constraints have to be imposed on transitions between different operating modes, which is achieved by Eqs. (2)–(4). The binary variable $z_{mm't}$ takes the value 1 if and only if the plant switches from mode m to mode m' at time t , which is enforced by the following constraint:

$$\sum_{m' \in TR_m^f} z_{m'm,t-1} - \sum_{m' \in TR_m^t} z_{mm',t-1} = y_{mt} - y_{m,t-1} \quad \forall m, t \in \overline{T} \quad (2)$$

where $TR_m^f = \{m' : (m', m) \in TR\}$ and $TR_m^t = \{m' : (m, m') \in TR\}$ with TR being the set of all possible mode-to-mode transitions.

The restriction that the plant has to remain in a certain mode for a minimum amount of time after a transition is expressed in the following constraint:

$$y_{m't} \geq \sum_{k=1}^{\theta_{mm'}} z_{mm',t-k} \quad \forall (m, m') \in TR, t \in \bar{T} \quad (3)$$

with $\theta_{mm'}$ being the minimum stay time in mode m' after switching to it from mode m .

For predefined sequences, each defined as a fixed chain of transitions from mode m to mode m' to mode m'' , we can specify a fixed stay time in mode m' by imposing the following constraint:

$$z_{mm',t-\bar{\theta}_{mm'm''}} = z_{m'm''t} \quad \forall (m, m', m'') \in SQ, t \in \bar{T} \quad (4)$$

where SQ is the set of predefined sequences and $\bar{\theta}_{mm'm''}$ is the fixed stay time in mode m' in the corresponding sequence.

3.4. Mass balance constraints

The plant produces a certain set of products, of which some may be storable. As stated in Eq. (5a), the inventory level at time t , IV_{its} , is the inventory level at time $t-1$ plus the amount produced minus the amount sold, SL_{its} , and minus the amount wasted, PW_{its} , in time period t . PW_{its} takes a nonzero value if the demand is satisfied and the inventory has reached its maximum capacity. Eq. (5b) sets bounds on the inventory levels, and Eq. (5c) states that also products purchased from other sources, denoted by PC_{its} , can be used to satisfy demand. Note that all variables involved in the mass balance constraints are second-stage variables.

$$IV_{its} = IV_{i,t-1,s} + PD_{its} - SL_{its} - PW_{its} \quad \forall i, t \in \bar{T}, s \quad (5a)$$

$$IV_{it}^{\min} \leq IV_{its} \leq IV_{it}^{\max} \quad \forall i, t \in \bar{T}, s \quad (5b)$$

$$SL_{its} + PC_{its} = D_{its} \quad \forall i, t \in \bar{T}, s \quad (5c)$$

3.5. Energy balance constraints

As stated in Eq. (6a), the plant can be powered by electricity purchased from contracts and from the spot market, denoted by EC_{ct} and ES_{ts} , respectively. EW_{ts} is the amount of electricity “wasted” in time period t of scenario s , which takes a nonzero value if the committed electricity purchase from contracts exceeds the electricity consumption in that particular scenario s . In practice,

this amount of “wasted” electricity simply does not get delivered, yet you still have to pay for it. Here, EC_{ct} is a first-stage variable, whereas ES_{ts} and EW_{ts} are second-stage variables. Eq. (6b) restricts the amount of electricity that can be purchased from the spot market in each time period.

$$EU_{ts} = \sum_c EC_{ct} + ES_{ts} - EW_{ts} \quad \forall t \in \bar{T}, s \quad (6a)$$

$$ES_{ts} \leq ES_t^{\max} \quad \forall t \in \bar{T}, s \quad (6b)$$

3.6. Power contract model

Power contracts can be very complex in their price structures. Here, we apply a relatively simple model that incorporates the main features of any common power contract. We assume that for each contract c , the price consists of two components: a time-dependent, and an amount-dependent component. The time-dependent price component roughly follows the expected spot electricity price profile, and is typically given for so-called time-of-use (TOU) periods. Figure 1a shows an example with four TOU periods over the course of 24 hours. The time-dependent price level depends on the type of TOU period: off-peak, mid-peak, or on-peak. The amount-dependent component sets a base price that depends on the total amount of electricity purchased during the entire scheduling horizon. As illustrated in Figure 1b, electricity is offered at discounted rates when certain purchase amounts have been reached.

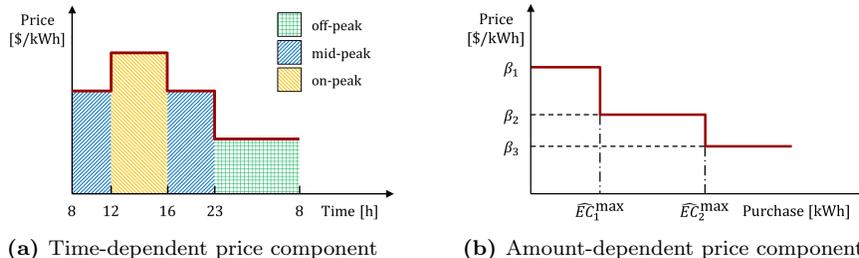


Figure 1: Contract prices typically consist of a time-dependent and an amount-dependent component.

The time-dependent price component is simply expressed through the cost coefficient α_{ct}^{EC} associated with the variable EC_{ct} , which is the amount of electricity purchased from contract c in time period t . To accommodate the amount-dependent price component, we apply a block contract model which can be

formulated as follows:

$$\widehat{EC}_c = \sum_{t \in \overline{T}} EC_{ct} \quad \forall c \quad (7a)$$

$$\bigvee_{b \in B_c} \left[\begin{array}{c} X_{cb} \\ BC_c = \beta_{cb} \widehat{EC}_c \\ \widehat{EC}_{c,b-1}^{\max} \leq \widehat{EC}_c \leq \widehat{EC}_{cb}^{\max} \end{array} \right] \quad \forall c \quad (7b)$$

$$\bigvee_{b \in B_c} X_{cb} \quad \forall c \quad (7c)$$

$$X_{cb} \in \{true, false\} \quad \forall c, b \in B_c \quad (7d)$$

where \widehat{EC}_c is the amount of electricity purchased from contract c over the entire scheduling horizon. B_c is the set of blocks of contract c , and X_{cb} is a boolean variable that is true if block b of contract c is chosen. \widehat{EC}_{cb}^{\max} is the amount of electricity that one has to purchase in block $b \in B_c$ before reaching the next block. Hence, block b is chosen if \widehat{EC}_c takes a value between $\widehat{EC}_{c,b-1}^{\max}$ and \widehat{EC}_{cb}^{\max} . In that case, the base cost, BC_c , is \widehat{EC}_c times the corresponding unit cost coefficient β_{cb} . Eq. (7c) states that only one block can be chosen for each contract.

By applying the hull reformulation (Balas, 1985), the disjunction and the logic constraints given by Eqs. (7b)–(7d) can be transformed into the following set of mixed-integer linear constraints:

$$\widehat{EC}_c = \sum_{b \in B_c} \widehat{EC}_{cb} \quad \forall c \quad (8a)$$

$$BC_c = \sum_{b \in B_c} \beta_{cb} \widehat{EC}_{cb} \quad \forall c \quad (8b)$$

$$\widehat{EC}_{c,b-1}^{\max} x_{cb} \leq \widehat{EC}_{cb} \quad \forall c, b \in B_c, b > 1 \quad (8c)$$

$$\widehat{EC}_{cb} \leq \widehat{EC}_{cb}^{\max} x_{cb} \quad \forall c, b \in B_c \quad (8d)$$

$$\sum_{b \in B_c} x_{cb} = 1 \quad \forall c \quad (8e)$$

where x_{cb} is a binary variable, and \widehat{EC}_{cb} is the disaggregated variable associated with \widehat{EC}_c .

As stated before, we assume that contract terms require that the electricity purchase decisions for the scheduling horizon are made before the beginning of that time horizon. These decisions cannot be changed later, i.e. consumers have to purchase the amount of electricity to which they has committed themselves regardless of their actual need. Hence, all contract-related decisions are first-stage decisions. Moreover, in each TOU period, which generally consists of

multiple time periods, the power purchase from a contract has to remain constant for the entire duration of that TOU period, which is expressed through the following constraint:

$$EC_{ct} = \overline{EC}_{cp} \quad \forall c, p \in P_c, t \in \widehat{T}_{cp} \quad (9a)$$

$$\overline{EC}_{cp} \leq \overline{EC}_{cp}^{\max} \quad \forall c, p \in P_c \quad (9b)$$

where \overline{EC}_{cp} is the electricity purchased from contract c in each time period within the TOU period p . P_c denotes the set of TOU periods for contract c , while \widehat{T}_{cp} is the set of time periods in TOU period p of contract c . $\overline{EC}_{cp}^{\max}$ bounds the amount that can be purchased from contract c in TOU period p .

3.7. Boundary conditions

The scheduling problem is formulated for a given time horizon. For the problem to be well-defined, initial conditions and terminal constraints are required, which are given in the following:

$$IV_{i,0,s} = IV_i^{\text{ini}} \quad \forall i, s \quad (10a)$$

$$y_{m,0} = y_m^{\text{ini}} \quad \forall m \quad (10b)$$

$$z_{mm't} = z_{mm't}^{\text{ini}} \quad \forall (m, m') \in TR, -\theta^{\max} + 1 \leq t \leq -1 \quad (10c)$$

$$IV_{i,t^{\text{fin}},s} \geq IV_i^{\text{fin}} \quad \forall i, s \quad (10d)$$

with $\theta^{\max} = \max\left(\max_{m,m'}\{\theta_{mm'}\}, \max_{m,m',m''}\{\bar{\theta}_{mm'm''}\}\right)$, which defines for how far back in the past the mode switching information has to be provided.

3.8. Total expected operating cost

The total expected operating cost, TC , consists of the cost of purchasing electricity from contracts, the expected cost of purchasing electricity from the spot market, and the expected cost of purchasing products, as stated in the following:

$$TC = \sum_c BC_c + \sum_{t \in \overline{T}} \sum_c \alpha_{ct}^{\text{EC}} EC_{ct} + \sum_s \varphi_s \sum_{t \in \overline{T}} \left(\alpha_{ts}^{\text{ES}} ES_{ts} + \sum_i \alpha_{it}^{\text{PC}} PC_{its} \right) \quad (11)$$

where φ_s denotes the probability of scenario s .

3.9. Conditional value-at-risk

There are many risk measures that are widely used in practice, such as the variance of the loss distribution, shortfall probability, downside risk, and value-at-risk (VaR) (Rockafellar, 2007). The risk measure of choice in this work is

the conditional value-at-risk (CVaR). For a given $\alpha \in (0, 1)$, the α -CVaR is defined as the expected loss greater than the α -VaR, which is the α -quantile of the loss distribution. Since the CVaR was introduced by [Rockafellar and Uryasev \(2000\)](#), it has become very popular because of its ability to consider the probability density in the tail of the loss distribution, its mathematical properties from being a coherent risk measure, and the ease of incorporating it into stochastic optimization models. For more information on the use of CVaR, we refer to [Rockafellar and Uryasev \(2000\)](#) and [Sarykalin et al. \(2008\)](#).

Applied to a scenario-based formulation and defined in terms of cost, the α -CVaR corresponds to the mean cost computed over the scenarios that have costs greater than the α -quantile. Because the model here considers price as well as demand uncertainty, we have to be cautious when defining the risk measure. If we simply apply the CVaR to the scenario set S , we bias toward high-demand scenarios in the sense that scenarios contributing to the CVaR will most likely be associated with high product demand as illustrated in [Figure 2a](#). The obvious explanation is that higher demand automatically leads to higher cost; thus, high demand is deemed to be risky and therefore contributes to the CVaR. However, higher demand also results in higher revenue, which is not considered in the CVaR defined in terms of cost; hence, high-demand scenarios are falsely regarded as unfavorable.

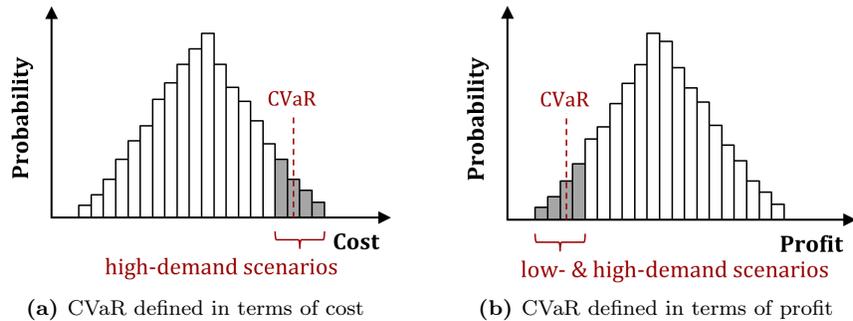


Figure 2: CVaR can be defined in terms of cost or in terms of profit, which can have different implications on the solution.

The simple solution to the problem is to define the CVaR in terms of profit by incorporating a constant revenue term, R_s , for each scenario. Now the α -CVaR is defined as the expected profit computed over the scenarios that have profit values smaller than the α -VaR, which is now the $(1 - \alpha)$ -quantile of the profit distribution. As illustrated in [Figure 2b](#), low-demand scenarios are likely to contribute to the CVaR because of the reduced revenues and therefore smaller

profits; however, high demand may also result in low profit, e.g. if the plant does not have sufficient production capacity to satisfy the demand such that additional products have to be purchased at higher cost.

The CVaR, denoted by CV , is incorporated into the model by adding the following constraints to the formulation:

$$CV = \kappa - \frac{1}{1-\alpha} \sum_s \varphi_s \omega_s \quad (12a)$$

$$\kappa - \left[R_s - \sum_c BC_c - \sum_{t \in \bar{T}} \sum_c \alpha_{ct}^{\text{EC}} EC_{ct} - \sum_{t \in \bar{T}} \left(\alpha_{ts}^{\text{ES}} ES_{ts} + \sum_i \alpha_{it}^{\text{PC}} PC_{its} \right) \right] \leq \omega_s \quad \forall s \quad (12b)$$

where κ and ω_s are continuous variables with $\kappa \in \mathbb{R}$ and $\omega_s \geq 0$. When CV is maximized, it takes the value of the α -CVaR.

3.10. Objective functions

With the proposed model, we can consider both risk-neutral and risk-averse optimization, which merely differ in their objectives. In risk-neutral optimization, where risk is not taken into account, the objective is the minimization of the total expected cost, hence the objective function is TC .

In risk-averse optimization, one tries to balance two typically conflicting objectives: optimizing the expected outcome and hedging against risk. Since the CVaR is defined as the expected profit over a small scenario subset consisting of the worst scenarios, the financial risk decreases with increasing CVaR. Therefore, the objective is to maximize a weighted sum of the total expected profit and the CVaR:

$$\zeta TP + (1 - \zeta) CV \quad (13)$$

with the total expected profit $TP = \sum_s \varphi_s R_s - TC$ and $\zeta \in [0, 1]$. Because there is usually a trade-off between expected outcome and risk in the sense that no solution can be found that maximizes both TC and CV , the weighting factor ζ can be used to specify which objective should be emphasized more. In this context, ζ can also be seen as a parameter that sets the desired level of risk aversion. The smaller ζ , the more risk-averse is the solution, since more weight is assigned to the CVaR. Note that the revenue term, $\sum_s \varphi_s R_s$, is constant; hence, if $\zeta = 1$, we obtain the risk-neutral formulation because maximizing TP is equivalent to minimizing TC . If $\zeta = 0$, only the CVaR is maximized.

4. Scenario generation and reduction

In the proposed two-stage stochastic programming framework, each scenario corresponds to a time series, i.e. a sequence of values assigned to each time period of the scheduling horizon. Scenarios can be obtained by sampling from a suitable stochastic forecasting model. Most electricity price forecasting tools are based on univariate time series models, such as autoregressive integrated moving average (ARIMA) models (Nogales et al., 2002). Aggarwal et al. (2009) present a review of methodologies for electricity price forecasting, which is not the focus of this work. This section merely provides a brief description of the ARIMA and ARIMAX (ARIMA with exogenous inputs) methods that are used to generate the scenarios in our case studies. Furthermore, we outline the scenario reduction technique that is used to reduce the scenario set to a manageable size.

4.1. Scenario generation using ARIMA/ARIMAX models

An ARIMA model is a linear univariate time series model, which expresses the output at time t , y_t , as a function of observed output values in previous time periods. An ARIMA model can be formulated as follows:

$$A(L)y_t = \frac{1}{(1-L)^d} B(L)\epsilon_t \quad (14a)$$

$$A(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p \quad (14b)$$

$$B(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q \quad (14c)$$

where L is the lag operator, i.e. $L^k x_t = x_{t-k}$ and $(1-L)^d x_t = x_t - x_{t-d}$.

An ARIMA model consists of three parts: the autoregressive part of order p given by $A(L)$, the integrated part of order d given by the integrator $1/(1-L)^d$, and the moving average part of order q given by $B(L)$. The parameter p defines how many past output values are considered, whereas q defines how many past error terms ϵ_t are taken into account. ϵ_t , also referred to as white noise, are generally assumed to be independent and identically distributed random variables sampled from a normal distribution with zero mean. The parameter d defines the integrator and is used to remove the nonstationarity in the data. There are $(p+q)$ parameters, namely the coefficients ϕ_1, \dots, ϕ_p and $\theta_1, \dots, \theta_q$. An ARIMA model is built by estimating these parameters using historical time series data.

An ARIMA model only considers time series data of the output variable. However, if we have information about exogenous inputs that are correlated with the output, we may want to make use of this information to help predicting future outputs. For instance, when predicting electricity price, useful exogenous

inputs could be the temperature and the price of natural gas. The ARIMAX model is an extension to the ARIMA model that takes input time series data into account. An ARIMAX model is given by the following formulation:

$$A(L)y_t = \frac{1}{(1-L)^d} B(L)\epsilon_t + C(L)u_t \quad (15)$$

with $C(L) = \eta_1 L^k + \eta_2 L^{k+1} + \eta_3 L^{k+2} \dots + \eta_b L^{k+b-1}$. Eq. (15) consists of the expression given in Eqs. (14) and an additional exogenous input term $C(L)u_t$. The nonnegative integer k is the input-output lag time expressed as number of time periods. Here, the $(p+q+b)$ parameters to be estimated are ϕ_1, \dots, ϕ_p , $\theta_1, \dots, \theta_q$, and η_1, \dots, η_b .

4.2. Scenario reduction using probability distance metrics

With an ARIMA or ARIMAX model, Monte Carlo simulation can be applied to sample scenarios. If the uncertain parameter can change significantly over time, which is certainly true in the case of spot electricity price, a large number of scenarios may be required to accurately characterize the uncertainty, resulting in a computationally intractable optimization problem. To reduce the computational effort yet still obtain good results, we seek to select a manageable number of scenarios that still represent the main features of the uncertainty. To achieve this, we apply the scenario reduction technique proposed by Dupacova et al. (2003), which selects a subset of scenarios from a given set of scenarios such that the probability distribution represented by the reduced scenario set is close to the one represented by the original scenario set. Here, the closeness of two distributions is measured in terms of a so-called probability distance.

A commonly used probability distance is the Kantorovich distance. Given two discrete probability distributions, Q and \tilde{Q} , with scenarios $\{\omega_1, \dots, \omega_N\}$ and $\{\tilde{\omega}_1, \dots, \tilde{\omega}_M\}$, and probability weights $\{q_1, \dots, q_N\}$ and $\{\tilde{q}_1, \dots, \tilde{q}_M\}$, respectively, the Kantorovich distance between Q and \tilde{Q} , here denoted by $D^K(Q, \tilde{Q})$, can be defined as follows:

$$D^K(Q, \tilde{Q}) = \min_{\eta_{ij}} \left\{ \sum_{i=1}^N \sum_{j=1}^M c(\omega_i, \tilde{\omega}_j) \eta_{ij} : \eta_{ij} \geq 0, \sum_{j=1}^M \eta_{ij} = q_i, \sum_{i=1}^N \eta_{ij} = \tilde{q}_j \quad \forall i, j \right\} \quad (16)$$

where $c(\omega, \tilde{\omega})$ is a nonnegative, continuous, symmetric cost function. The minimum is taken over all possible joint probabilities.

In scenario reduction, $M < N$, and the reduced scenario set $\{\tilde{\omega}_1, \dots, \tilde{\omega}_M\}$ is a subset of the original scenario set $\{\omega_1, \dots, \omega_N\}$. The problem then becomes

which M scenarios to select and what values to assign to the new probabilities \tilde{q}_j for $j = 1, \dots, M$ such that D^K is minimized. To solve the scenario reduction problem, Römisch and coworkers (Dupacova et al., 2003; Heitsch and Römisch, 2003) propose two heuristic algorithms—forward selection and backward reduction—which are fast but do not guarantee optimality. Optimality can be achieved by solving an MILP formulation recently proposed by Li and Floudas (2014). In this work, we apply the backward reduction method implemented in the scenario reduction routine SCENRED2 in GAMS (GAMS Development Corporation, 2015).

5. Multicut Benders decomposition

Even with a smaller number of scenarios, the optimization problem can still be very large and difficult to solve. In order to reduce the solution time, Benders decomposition (Benders, 1962) is applied. In this decomposition framework, the optimal solution is found by iteratively converging lower and upper bounds on the optimal objective function. In a minimization problem, upper bounds are obtained by fixing the first-stage variables and optimizing the second-stage decisions for each scenario. Lower bounds are obtained by solving a master problem in the space of the first-stage variables, which incorporates lower bounds on the second-stage costs. Convergence of the algorithm is achieved by improving the lower bound with the master problem which adds successively dual information of the second-stage costs obtained from the upper-bounding subproblems, which in turn potentially leads to first-stage decisions that improve the upper bound.

Lower bounds on the second-stage costs are added to the master problem in the form of cuts determined by the dual multipliers of the subproblems. In the classical Benders decomposition algorithm, one cut is generated at every iteration. To provide stronger lower bounds, Birge and Louveaux (1988) propose to generate multiple cuts at every iteration, namely one cut per scenario. This leads to a master problem that grows faster in size; however, the solution time is often shorter because fewer iterations are required. Multicut Benders decomposition has been successfully applied to various two-stage stochastic programming problems (You and Grossmann, 2013; Garcia-Herreros et al., 2014; Skar et al., 2014). For this specific problem, computational experiments have shown that multicut Benders decomposition is superior to the traditional Benders decomposition approach.

In the following, the multicut Benders decomposition approach is outlined. Consider the original problem expressed in the following general form:

$$\min \quad c^T x + \sum_s \varphi_s q_s^T y_s \quad (17a)$$

$$\text{s.t. } Ax \geq b \quad (17b)$$

$$T_s x + W_s y_s \geq h_s \quad \forall s \quad (17c)$$

$$y_s \geq 0 \quad \forall s \quad (17d)$$

where x denotes the vector of first-stage variables, which may be continuous or integer, and y_s are the continuous second-stage variables for scenario s . Parameter matrices A , b , and c are independent of the scenarios, while T_s , W_s , h_s , and q_s are scenario-specific; φ_s denotes the probability of scenario s .

The subproblem is formulated by fixing the first-stage variables x in (17). After omitting the constant term in the objective function and the pure first-stage constraints $Ax \geq b$, the dual of the subproblem at the k -th iteration can be written as follows:

$$\max \sum_s (h_s - T_s x^k)^\top \mu_s \quad (18a)$$

$$\text{s.t. } W_s^\top \mu_s \leq \varphi_s q_s \quad \forall s \quad (18b)$$

$$\mu_s \geq 0 \quad \forall s \quad (18c)$$

where μ_s denotes the vector of dual multipliers. Note that (18) can be decomposed by scenario allowing solving multiple separate smaller subproblems in parallel.

The multicut master problem at the k -th iteration is formulated as follows:

$$\min c^\top x + \sum_s \xi_s \quad (19a)$$

$$\text{s.t. } Ax \geq b \quad (19b)$$

$$(h_s - T_s x)^\top \mu_s^k \leq \xi_s \quad \forall s, k \quad (19c)$$

where ξ_s is a continuous variable representing the approximate cost of scenario s . So-called optimality cuts are accumulated in Eq. (19c) at each iteration. Note that no feasibility cuts are considered because the problem has complete recourse.

A flowchart for the multicut Benders decomposition algorithm is shown in Figure 3. The algorithm terminates when the gap between upper and lower bounds becomes smaller than a prespecified tolerance ϵ or the time limit is reached.

6. Illustrative example

To demonstrate the main features of the model, we first apply it to an illustrative example in which a plant is considered that produces two products,

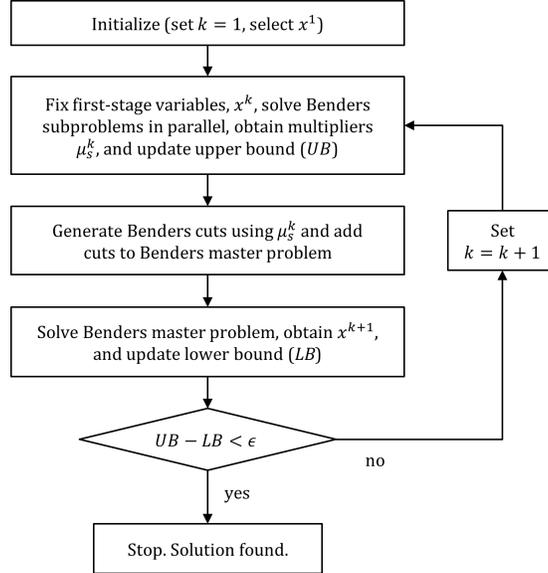


Figure 3: Flowchart for Benders decomposition algorithm.

P1 and P2. The plant can operate in three different operating modes: off, startup, and on. The vertices of the polyhedral feasible regions of the modes are listed in Table 1, while the electricity consumption coefficients associated with each mode are shown in Table 2. The possible mode transitions are off→startup, startup→on, and on→off, for which the minimum stay times after transition are 2 h, 6 h, and 8 h, respectively. In fact, the startup process takes exactly 2 h, i.e. for the sequence on→startup→on, the fixed stay time in the startup mode is 2 h. At the start of the scheduling horizon, the plant is in the on mode. Also, it is assumed that no mode switching has occurred in the previous 8 h.

Table 1: Vertices associated with each operating mode of the plant from the illustrative example.

Mode	Vertex	P1 [kg]	P2 [kg]
off	1	0	0
startup	1	5	5
on	1	10	10
	2	50	10
	3	30	40
	4	70	40

Table 2: Coefficients for linear electricity consumption correlations associated with each operating mode.

Mode	δ [kWh]	γ_{P1} [kWh/kg]	γ_{P2} [kWh/kg]
off	0	0	0
startup	500	0	0
on	800	20	30

Data regarding the inventory of each product are given in Table 3. We consider a two-day scheduling horizon, which starts at 8 AM of the first day. An hourly time discretization is applied, resulting in 48 time periods. One power contract is available with the following characteristics:

- Daily TOU periods: two on-peak (6 AM – 12 AM, 12 AM – 6 PM), two off-peak (0 AM – 6 AM, 6 PM – 0 AM) TOU periods
- Time-dependent price component: \$20/kWh during on-peak, \$15/kWh during off-peak periods
- Amount-dependent price component: \$16/kWh if total electricity purchase less than 30 MWh, \$15/kWh if between 30 MWh and 80 MWh, \$14/kWh if greater than 80 MWh
- Maximum electricity purchase in every time period: 3 MWh

Table 3: Inventory bounds and initial inventory levels for each product.

	IV^{\min} [kg]	IV^{\max} [kg]	IV^{ini} [kg]	IV^{fin} [kg]
P1	600	6,000	1,000	1,000
P2	300	3,000	500	500

Three sets of spot electricity price scenarios with different levels of uncertainty—low, medium, and high—are considered. Each set consists of 40 equiprobable scenarios, where each scenario is associated with a specific price profile over the 48 hours of the scheduling horizon. As an example, the price profiles along with the expected price for the price scenario set with medium level of uncertainty are shown in Figure 4. All price scenario data for this illustrative example are provided in the supplementary material. No limit is imposed on the amount of electricity that can be purchased from the spot market.

Demand is assumed to be constant over time. This is a reasonable assumption since the products in this example are storable and small variations in

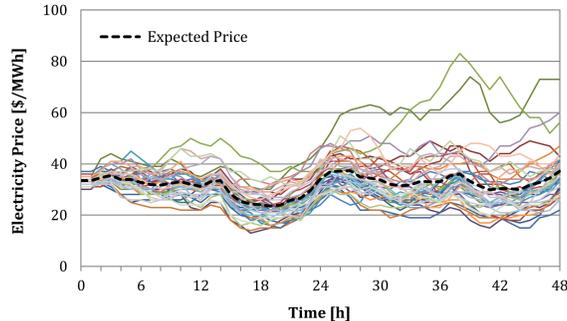


Figure 4: Spot electricity price profiles associated with the scenario set with medium level of price uncertainty.

demand over time can be made up by the available inventory. This assumption simplifies the generation of scenarios since no time series are needed. Also, the uncertainty can be described by considerably fewer scenarios. By further assuming that the demands for the two products are perfectly positively correlated, we generate three sets of demand scenarios with different levels of uncertainty—low, medium, and high—with the same expected demands. The data for the three scenario sets including the probabilities for each scenario are given in Table 4, where Var^D denotes the variance which is a measure for the level of uncertainty.

Table 4: Demand values and probabilities for each scenario of the three different demand scenario sets.

\bar{s}	$\varphi_{\bar{s}}$	low Var^D		medium Var^D		high Var^D	
		D_{P1} [kg/h]	D_{P2} [kg/h]	D_{P1} [kg/h]	D_{P2} [kg/h]	D_{P1} [kg/h]	D_{P2} [kg/h]
1	0.25	57	33.25	54	31.5	48	28
2	0.50	60	35	60	35	60	35
3	0.25	63	36.75	66	38.5	72	42

The prices for purchasing Products P1 and P2 are \$3/kg and \$4/kg, respectively. The revenues for each scenario can be computed by applying the selling prices \$1.5/kg and \$2/kg to the demands of Products P1 and P2, respectively.

In the following, various instances are created by combining different price and demand scenario sets. Risk-neutral optimization is performed before risk-averse optimization is applied to the same instances. In the analysis of the results, we pay special attention to the assessment of the added value obtained from explicitly accounting for uncertainty in the model.

6.1. Risk-neutral optimization

In risk-neutral optimization, the objective is to minimize the total expected cost. To compare the deterministic and stochastic solutions, the value of stochastic solution (VSS) is used as a measure for the difference in the solutions. Here, the VSS is defined as

$$\text{VSS} = TC^{\text{det}} - TC^{\text{sto}} \quad (20)$$

where TC^{sto} is the total expected cost at the optimal solution of the stochastic problem, and TC^{det} is obtained by solving the stochastic problem with the first-stage variables fixed to the values at the optimal solution of the deterministic problem. Hence, the VSS expresses the cost savings that can be expected from implementing the stochastic solution instead of the deterministic solution. The relative VSS is defined as

$$\overline{\text{VSS}} = \frac{TC^{\text{det}} - TC^{\text{sto}}}{TC^{\text{det}}} \quad (21)$$

which is only meaningful if $TC^{\text{det}} > 0$. TC^{det} could become negative if there are scenarios with negative electricity prices, which is unlikely but possible.

First, we consider three cases in which uncertainty only exists in the electricity price. The total expected costs resulting from the deterministic and stochastic solutions as well as the corresponding VSS are listed in Table 5. Here, $|S^{\text{D}}|$ and $|S^{\text{P}}|$ denote the numbers of demand and price scenarios, respectively. Since there is no demand uncertainty, $|S^{\text{D}}| = 1$. One can see that the VSS is essentially zero in the low- and medium-Var^P cases. In the high-Var^P case, the VSS is noticeable but very small.

Table 5: Expected costs and VSS resulting from risk-neutral optimization with only electricity price uncertainty.

$ S^{\text{D}} $	$ S^{\text{P}} $	Var ^P	TC^{det} [\$]	TC^{sto} [\$]	VSS [\$]	$\overline{\text{VSS}}$ [%]
1	40	low	4,351	4,351	0	0.0
1	40	medium	4,422	4,420	2	0.0
1	40	high	4,401	4,329	72	1.6

Table 6 shows the results for three cases in which uncertainty only exists in the product demand. Here, significant VSS can be observed, which grows with increasing level of uncertainty. For the medium-Var^D case, Figure 5 shows the electricity purchase profiles from the deterministic and stochastic solutions. Note that the shown electricity purchase from the spot market is the expected

value computed over all scenarios. The contract price comprises both the time-dependent and the amount-dependent components at the chosen purchasing amount. The comparison between the two solutions shows the impact of accounting for other scenarios besides the expected one. In the stochastic solution, the plant operates longer in order to accommodate for the high-demand scenario. By doing so, more flexibility is provided for load shifting such that the electricity consumption during the high price peak can be reduced; this in turn lowers the need for electricity purchase from the power contract.

Table 6: Expected costs and VSS from risk-neutral optimization with only demand uncertainty.

$ S^D $	Var^D	$ S^P $	TC^{det} [\$]	TC^{sto} [\$]	VSS [\$]	$\overline{\text{VSS}}$ [%]
3	low	1	4,528	4,431	97	2.1
3	medium	1	4,674	4,459	215	4.6
3	high	1	4,967	4,621	346	7.0

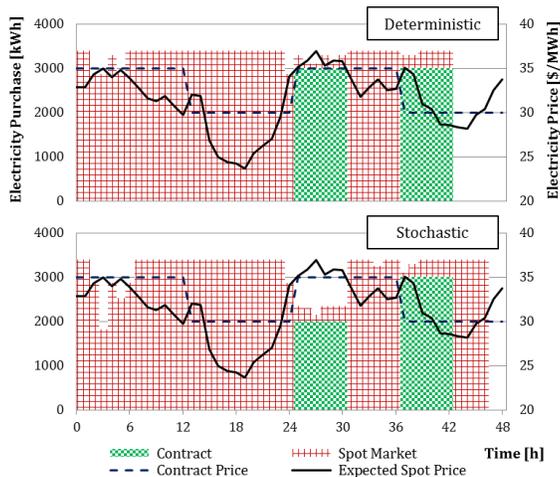


Figure 5: Electricity purchase profiles for the medium- Var^D case.

The results from the first six cases indicate that accounting for demand uncertainty can lead to significant added value, while this does not necessarily hold true for electricity price uncertainty. This raises the following question: If uncertainty exists in both electricity price and product demand, is there a benefit from considering price uncertainty in the model in addition to accounting for demand uncertainty? To answer this question, we consider nine cases involving

both price and demand uncertainty. The results are shown in Table 7, where $|\tilde{S}^P|$ denotes the number of scenarios in the reduced scenario set, which is used in the stochastic optimization. $|\tilde{S}^P|$ indicates whether price uncertainty is considered in the model or not. Each of the nine cases is solved once only considering the expected price profile ($|\tilde{S}^P| = 1$, ignoring price uncertainty) and once with all possible price scenarios ($|\tilde{S}^P| = 40$, accounting for price uncertainty).

Table 7: Expected costs and VSS from risk-neutral optimization with both electricity price and demand uncertainty.

$ S^D $	Var^D	$ S^P $	Var^P	$ \tilde{S}^P $	TC^{det} [\$]	TC^{sto} [\$]	VSS [\$]	$\overline{\text{VSS}}$ [%]
3	low	40	low	1	4,460	4,356	104	2.3
				40		4,356	104	2.3
3	low	40	medium	1	4,528	4,490	38	0.8
				40		4,426	102	2.3
3	low	40	high	1	4,501	4,455	46	0.1
				40		4,333	168	3.7
3	medium	40	low	1	4,605	4,413	192	4.2
				40		4,383	222	4.8
3	medium	40	medium	1	4,670	4,502	168	3.6
				40		4,450	220	4.7
3	medium	40	high	1	4,637	4,464	173	3.7
				40		4,344	293	6.3
3	high	40	low	1	4,898	4,548	350	7.1
				40		4,548	351	7.2
3	high	40	medium	1	4,960	4,673	287	5.8
				40		4,591	369	7.4
3	high	40	high	1	4,923	4,618	305	6.2
				40		4,494	429	8.7

Table 7 lists some significant VSS, especially in the cases with high level of demand uncertainty. In most cases, a comparison of the two instances ($|\tilde{S}^P| = 1$ and $|\tilde{S}^P| = 40$) shows that the impact of price uncertainty on the solution is relatively small. In other words, the added benefit from considering price uncertainty in addition to demand uncertainty is only moderate. Exceptions are the two low- Var^D cases with medium and high Var^P .

6.2. Risk-averse optimization

In risk-averse optimization, the objective is to maximize a weighted sum of the expected profit and the CVaR as given by Eq. (13). The VSS is defined as

$$\text{VSS} = \zeta (TP^{\text{sto}} - TP^{\text{det}}) + (1 - \zeta) (CV^{\text{sto}} - CV^{\text{det}}) \quad (22)$$

where TP^{sto} and CV^{sto} denote the total expected profit and CVaR at the optimal solution of the stochastic problem, and TP^{det} and CV^{det} are computed by solving the stochastic problem with fixed first-stage decisions obtained from the deterministic problem. The relative VSS is defined as

$$\overline{\text{VSS}} = \frac{\zeta (TP^{\text{sto}} - TP^{\text{det}}) + (1 - \zeta) (CV^{\text{sto}} - CV^{\text{det}})}{\zeta TP^{\text{det}} + (1 - \zeta) CV^{\text{det}}} \quad (23)$$

which is only meaningful if the denominator is positive.

Risk-averse optimization is performed on the same instances presented in the previous subsection. We choose $\alpha = 0.9$ and $\zeta = 0.5$, i.e. equal weights are assigned to the total expected profit and the CVaR. The results are shown in Tables 8, 9, and 10. It is remarkable how much the deterministic and stochastic solutions can differ when risk is considered. The VSS reported here are considerably higher than those obtained in risk-neutral optimization. Furthermore, unlike in risk-neutral optimization, accounting for price uncertainty can lead to much improved solutions. In the cases with only electricity price uncertainty, the VSS increases dramatically with the level of uncertainty (see Table 8). This can also be observed in the cases in which uncertainty exists in both price and demand. Moreover, by comparing the results for $|\tilde{S}^P| = 1$ and $|\tilde{S}^P| = 40$ in each case, we see that considering electricity price uncertainty can greatly improve the solution. In fact, ignoring price uncertainty can be detrimental, to an extent that the stochastic solution is even worse than the deterministic solution (indicated by negative VSS).

Table 8: Expected profits, CVaRs, and VSS from risk-averse optimization with only electricity price uncertainty.

$ S^D $	$ S^P $	Var^P	TP^{det} [\$]	CV^{det} [\$]	TP^{sto} [\$]	CV^{sto} [\$]	VSS [\$]	$\overline{\text{VSS}}$ [%]
1	40	low	3,329	2,981	3,328	3,006	12	0.4
1	40	medium	3,258	2,169	3,155	2,668	198	7.3
1	40	high	3,278	1,334	3,242	2,369	499	21.7

For the medium- Var^D medium- Var^P case, Figure 6 shows the electricity purchase profiles for the stochastic solutions with $|\tilde{S}^P| = 1$ and with $|\tilde{S}^P| = 40$. In the

Table 9: Expected profits, CVaRs, and VSS from risk-averse optimization with only demand uncertainty.

$ S^D $	Var^D	$ S^P $	TP^{det} [\$]	CV^{det} [\$]	TP^{sto} [\$]	CV^{sto} [\$]	VSS [\$]	$\overline{\text{VSS}}$ [%]
3	low	1	3,152	3,015	3,249	3,058	70	2.3
3	medium	1	3,006	2,631	3,221	2,831	207	7.4
3	high	1	2,713	1,863	3,059	2,384	434	19.0

first instance, in which price uncertainty is ignored, the solution suggests purchasing all electricity from the spot market. This solution provides maximum flexibility, and is therefore good for dealing with demand uncertainty, which is accounted for in the model. In the second instance, the model considers both price and demand uncertainty, which results in a very different solution. In this solution, more than half of the electricity is purchased from the power contract. This solution results in a slightly lower expected profit, but significantly increases the CVaR, i.e. it reduces the risk of low-profit scenarios.

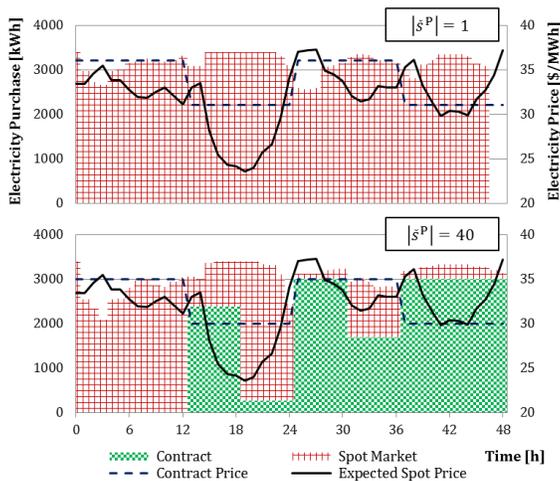


Figure 6: Electricity purchase profiles for the medium- Var^D medium- Var^P case.

All models were implemented in GAMS 24.4.1 ([GAMS Development Corporation, 2015](#)), and the commercial solver CPLEX 12.6.1 was applied to solve the MILPs. Each model has up to approximately 133,000 continuous variables, 315 binary variables, and 128,000 constraints. All models were solved to zero integrality gap in less than 20 seconds wall-clock time on an Intel[®] Core[™] i7-2600

Table 10: Expected profits, CVaRs, and VSS from risk-averse optimization with both electricity price and demand uncertainty.

$ S^D $	Var^D	$ S^P $	Var^P	$ \tilde{S}^P $	TP^{det} [\$]	CV^{det} [\$]	TP^{sto} [\$]	CV^{sto} [\$]	VSS [\$]	$\overline{\text{VSS}}$ [%]
3	low	40	low	1	3,220	2,833	3,323	2,918	94	3.1
				40			3,321	2,921	94	3.1
3	low	40	medium	1	3,152	2,072	3,199	1,360	-333	-12.7
				40			3,163	2,595	267	10.2
3	low	40	high	1	3,179	1,252	3,321	752	-179	-8.1
				40			3,240	2,331	570	25.7
3	medium	40	low	1	3,075	2,496	3,296	2,713	219	7.9
				40			3,297	2,713	220	7.9
3	medium	40	medium	1	3,010	1,876	3,178	1,293	-207	-8.5
				40			3,158	2,495	383	15.7
3	medium	40	high	1	3,043	1,084	3,304	727	-48	-2.3
				40			3,232	2,255	680	33.0
3	high	40	low	1	2,782	1,732	3,132	2,282	450	20.0
				40			3,130	2,296	457	20.2
3	high	40	medium	1	2,720	1,332	3,007	1,057	6	0.3
				40			3,012	2,167	563	27.8
3	high	40	high	1	2,757	624	3,164	635	209	12.4
				40			3,072	1,994	843	49.9

machine at 3.40 GHz with eight processors and 8 GB RAM running Windows 7 Professional.

7. Industrial case study

We now apply the proposed model to a real-world industrial case study provided by Praxair. Here, we consider an air separation plant that produces liquid oxygen (LO2) and liquid nitrogen (LN2). Two power contracts, which differ in price and availability, are considered. The scheduling horizon is one week, to which an hourly time discretization is applied resulting in 168 time periods. Note that due to confidentiality reasons, we cannot disclose information about the plant specifications as well as the actual product demand. Therefore, all results presented in the following are given without units and the values are normalized if necessary.

To model the uncertainty in electricity price, an ARIMAX model with temperature as exogenous input is created in R (R Core Team, 2014) using data from four consecutive weeks as training data. Monte Carlo simulation is then

applied to generate 1,000 equiprobable price scenarios; the corresponding price profiles are shown in Figure 7 along with the expected price profile. Because lines are shown in the same color, the individual scenarios are indistinguishable in the diagram; however, the picture depicts the large spread in the price distribution. One can see that the level of uncertainty increases with time. We apply scenario reduction to the 1,000 price scenarios and obtain a set of 50 scenarios, which are shown in Figure 8. Notice that the scenario reduction process assigns different probabilities to the scenarios in the reduced set.

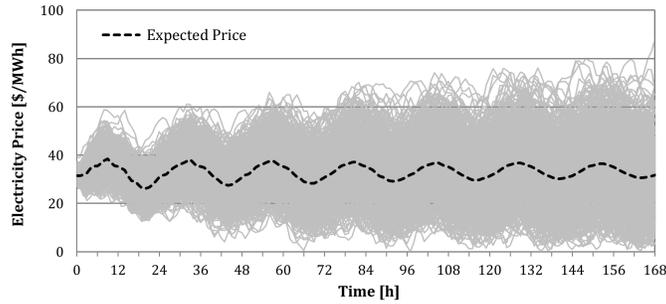


Figure 7: Spot electricity price profiles associated with the full set of scenarios and the expected price profile.

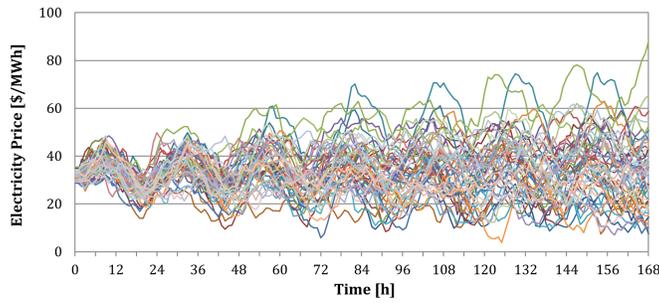


Figure 8: Spot electricity price profiles associated with the reduced set of scenarios.

Product demand occurs every six hours and is the total amount of product that needs to be drawn from the tank over the course of these six hours. The assumption is that there is sufficient capacity in the inventory to handle flows into and out of the inventory tank during this period of time. By further assuming that the demands for LO2 and LN2 are correlated, the uncertainty in product demand is characterized by five scenarios that resemble a normal distribution.

Again, we use the VSS to quantify the difference between the deterministic

and stochastic solutions. However, note that since the stochastic optimization is performed on a reduced scenario set, but the VSS is computed over the full scenario set, only an approximate VSS (AVSS) can be obtained. The AVSS would take the value of the true VSS if the obtained first-stage decisions were the same as the ones that one would obtain from solving the stochastic problem with the full set of scenarios.

7.1. Risk-neutral vs. risk-averse optimization

We create three cases with different levels of demand uncertainty: low, medium, and high. For each case, two instances are created, one neglecting electricity price uncertainty ($|\tilde{S}^P| = 1$), the other incorporating the 50 price scenarios from the reduced scenario set ($|\tilde{S}^P| = 50$). Risk-neutral and risk-averse ($\zeta = 0.5$, $\alpha = 0.9$) optimization are performed on all instances for which the results are shown in Tables 11 and 12, respectively.

The most notable observation is that in risk-neutral optimization, accounting for price uncertainty does not seem to provide any added value at all. In every of the three cases, the VSS does not change when $|\tilde{S}^P|$ is increased from 1 to 50. In contrast, in risk-averse optimization, there is a clear increase in VSS when price uncertainty is considered in the stochastic optimization. Note that a consistent increase in VSS can be observed although some of the stochastic problems with $|\tilde{S}^P| = 50$ were not solved to optimality (see Section 7.4). This result is very similar to what is observed in the illustrative example, which further supports the hypothesis that accounting for electricity price uncertainty in risk-neutral optimization may be unnecessary, while in risk-averse optimization it leads to significant additional benefit.

Table 11: Expected cost and VSS from risk-neutral optimization.

$ S^D $	Var^D	$ S^P $	$ \tilde{S}^P $	TC^{det}	TC^{sto}	AVSS	$\overline{\text{AVSS}}$ [%]
5	low	1,000	1	12.49	12.49	0.00	0.0
			50		12.49	0.00	0.0
5	medium	1,000	1	13.40	13.08	0.32	2.4
			50		13.08	0.32	2.4
5	high	1,000	1	15.98	14.82	1.16	7.2
			50		14.83	1.15	7.2

In Figure 9, we compare for the high- Var^D case the solutions obtained from the deterministic, risk-neutral, and risk-averse optimization. One can see that these three approaches lead to very different decisions. In deterministic optimization, uncertainty is ignored such that decisions are primarily driven by the

Table 12: Expected profits, CVaRs, and VSS from risk-averse optimization.

$ S^D $	Var^D	$ S^P $	$ \hat{S}^P $	TP^{det}	CV^{det}	TP^{sto}	CV^{sto}	AVSS	$\overline{\text{AVSS}} [\%]$
5	low	1,000	1	15.51	12.84	15.51	12.84	0.00	0.0
			50	15.27	13.68	0.30	2.1		
5	medium	1,000	1	14.60	11.15	14.69	11.39	0.17	1.3
			50	14.80	11.59	0.32	2.5		
5	high	1,000	1	12.02	5.78	12.82	6.34	0.68	7.6
			50	12.71	6.57	0.74	8.3		

differences between the power contract prices and the expected spot price. A significant amount of electricity is procured from Contract 2 because the price discount at this purchasing amount makes it less expensive than purchasing from the spot market during on-peak hours. One issue with the deterministic solution is only implicitly shown in the diagram, namely that there is a very high expected cost of purchasing additional products because the selected operating modes do not have sufficient production capacities in the high-demand scenarios.

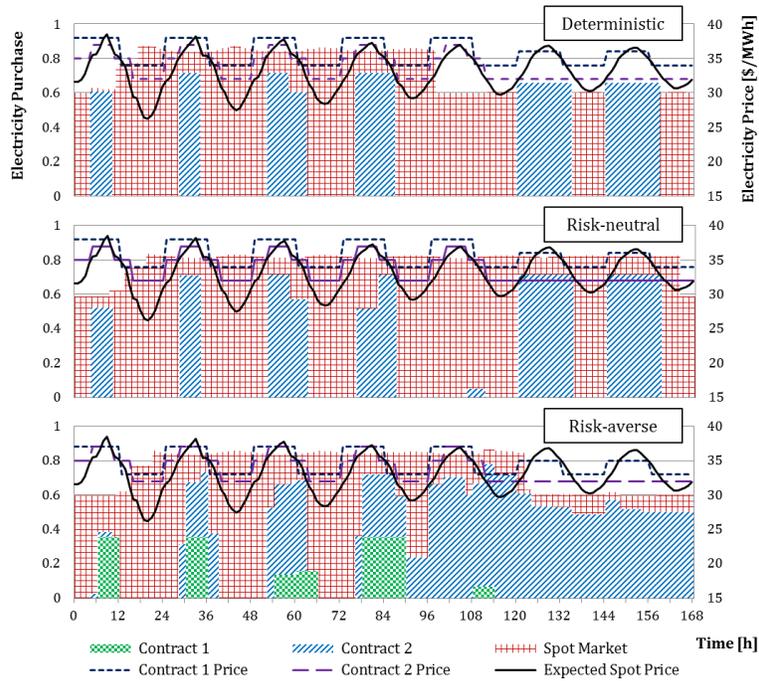


Figure 9: Electricity purchase profiles for the high- Var^D case.

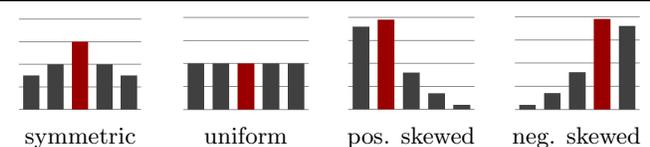
In risk-neutral stochastic optimization, first-stage decisions are to a great extent driven by the need for flexibility that has to be maintained in the second stage in order to react to different scenarios. Deterministic and risk-neutral optimization lead to similar schedules for the electricity procurement from power contracts. However, the risk-neutral solution suggests selecting operating modes with higher production capacities in order to be able to accommodate high-demand scenarios. This strategy leads to considerably lower costs for purchasing additional products, especially during the last three days of the week.

The electricity procurement decisions resulting from risk-averse optimization are very different from the deterministic and risk-neutral solutions. Here, more than half of the required electricity is purchased from power contracts. Because no more electricity can be purchased from Contract 2 due to the specified purchase limit, electricity is also procured from Contract 1. Here, contracts are effectively used to hedge against the risk of low-profit scenarios. In particular, considerable amount of electricity is purchased from contracts toward the end of the week when the level of uncertainty in the spot electricity price is highest.

7.2. Impact of product demand distribution

Next, the impact of different probability distributions describing the demand uncertainty is examined. We create four different demand distributions: symmetric, uniform, positively skewed, and negatively skewed. Each distribution is represented by five discrete scenarios as depicted by the figures in Table 13. The symmetric demand distribution is the one used in the instances presented in the previous subsection. The expected demand is the same for all four distributions.

Table 13: VSS [%] obtained from risk-averse optimization for different demand distributions and levels of demand uncertainty.



Demand Distribution	symmetric	uniform	pos. skewed	neg. skewed
low Var^D	2.1	2.5	3.5	1.8
medium Var^D	2.5	3.4	6.0	1.4
high Var^D	8.3	9.7	19.3	4.9

Table 13 shows the VSS obtained from risk-averse optimization ($\zeta = 0.5$, $\alpha = 0.9$) applied to the three cases with different levels of demand uncertainty. The highest VSS are achieved for the positively skewed demand distribution, while the VSS are lowest for the negatively skewed distribution. Evidently, the

more probability weight is assigned to low-demand scenarios, the greater is the benefit gained from stochastic optimization.

7.3. Trade-off between expected outcome and risk

Optimizing the expected outcome and minimizing risk are usually two conflicting objectives. In our case, it means that one cannot maximize the total expected profit and maximize the CVaR at the same time. To show this trade-off, we solve the risk-averse model with different values for the weighting factor ζ . The results are shown for three different values of α —0.8, 0.9, and 0.95—in Figure 10, where the CVaR is plotted against the total expected profit. The left endpoint of each curve depicts the solution of the case in which $\zeta = 0$, and the right endpoint is achieved when $\zeta = 1$. The CVaR decreases with increasing profit, which means that in order to achieve a high CVaR (reduce risk), one has to accept a lower expected profit.

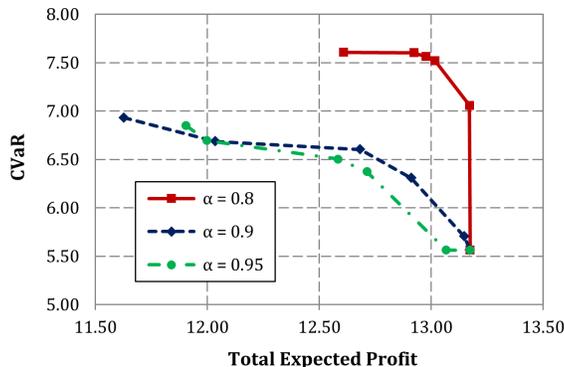


Figure 10: Solutions obtained by changing the weights in the objective function, showing the trade-off between total expected profit and CVaR.

Note that the obtained solutions are not necessarily Pareto-optimal due to multiple reasons: (1) the stochastic problem is only solved for a subset of scenarios, (2) the problem may not be solved to optimality within the given time, and (3) since the model is nonconvex, maximizing the weighted sum does not guarantee Pareto optimality. The last limitation could be circumvented by applying the ϵ -constraint approach (Hwang and Masud, 1979); however, the downside would be that the Benders decomposition algorithm could then not be used to solve the model.

7.4. Computational results

With 5 demand scenarios and 50 price scenarios in the reduced scenario set, the stochastic problem is solved for 250 scenarios. Each model has approxi-

mately 3.6 million continuous variables, 3,700 binary variables, and 2.7 million constraints. Here, we apply the proposed multicut Benders decomposition since the problems cannot be solved in fullspace in a reasonable amount of time. All models were implemented in GAMS 24.4.1 (GAMS Development Corporation, 2015), and the commercial solver CPLEX 12.6.1 was applied to solve the MILPs on an Intel® Core™ i7-2600 machine at 3.40 GHz with eight processors and 8 GB RAM running Windows 7 Professional.

The computational results for solving the risk-neutral models and a selected set of the risk-averse models are shown Tables 14 and 15, respectively. The wall-clock times and optimality gaps for solving each model in fullspace and with multicut Benders decomposition are reported. We specify a computational budget of 7,200 s (two hours). Note that the time reported for multicut Benders decomposition is the time after solving the master problem in the last iteration, which may exceed 7,200 s. However, solving this last master problem is only needed for obtaining the final lower bound on the objective function and computing the final gap; the best feasible solution is obtained before solving the last master problem, at which point the computational time usually has not exceeded 7,200 s. If an optimality gap is not available (n/a), it means that no feasible solution can be found within the time limit.

Table 14: Computational results for risk-neutral optimization

$ S^D $	Var ^D	$ \tilde{S}^P $	Fullspace		Multicut Benders	
			Time [s]	Gap [%]	Time [s]	Gap [%]
5	low	50	7,200	n/a	7,204	0.0
5	medium	50	7,200	100	4,871	0.0
5	high	50	7,200	100	502	0.0

From the results in Tables 14 and 15, one can see that in almost all cases, the solver cannot even find a feasible solution within the given time if the problem is to be solved in fullspace. When solving the fullspace risk-neutral models for the medium- and high-Var^D cases, feasible but very poor solutions are found (100% gap). In contrast, when the proposed multicut Benders decomposition is applied, all problems can be solved to optimality or close to optimality. In general, solving the risk-neutral models requires less computing time than solving the risk-averse models. Furthermore, the computational effort decreases with increasing level of demand uncertainty, which can be explained by the reduction of the degree of symmetry when the scenarios are more easily distinguishable.

Table 15: Computational results for risk-averse optimization with $\zeta = 0.5$ and $\alpha = 0.9$

$ S^D $	Var^D	$ \tilde{S}^P $	Demand Distribution	Fullspace Time [s]	Fullspace Gap [%]	Multicut Benders Time [s]	Multicut Benders Gap [%]
5	low	50	symmetric	7,200	n/a	7,531	5.1
5	medium	50	symmetric	7,200	n/a	7,830	2.2
5	high	50	symmetric	7,200	n/a	7,531	0.7
5	low	50	uniform	7,200	n/a	7,247	3.0
5	medium	50	uniform	7,200	n/a	7,499	1.8
5	high	50	uniform	7,200	n/a	7,550	1.5
5	low	50	pos. skewed	7,200	n/a	7,275	5.7
5	medium	50	pos. skewed	7,200	n/a	7,604	2.2
5	high	50	pos. skewed	7,200	n/a	7,371	0.4
5	low	50	neg. skewed	7,200	n/a	7,231	5.2
5	medium	50	neg. skewed	7,200	n/a	7,286	0.9
5	high	50	neg. skewed	7,200	n/a	5,165	0.0

8. Discussion

The results from the illustrative example and the industrial case study show the difference between deterministic, risk-neutral, and risk-averse optimization, as well as the impact of electricity price and product demand uncertainty. The most remarkable insight drawn from the observations is that in risk-neutral optimization, explicitly modeling electricity price uncertainty does not lead to any significant added value. In the following, we provide an explanation for this phenomenon.

First, we notice that in the risk-neutral model, electricity prices only appear in the objective function, not in the constraints; thus, uncertainty in the price does not affect the feasible space. We can therefore restrict our analysis to the objective function given by Eq. (11), which can also be written as follows:

$$TC = \underbrace{\sum_c BC_c + \sum_{t \in \bar{T}} \sum_c \alpha_{ct}^{EC} EC_{ct}}_{TC^1} + \underbrace{\sum_{\bar{s} \in S^D} \varphi_{\bar{s}}^D \sum_{\hat{s} \in S^P} \varphi_{\hat{s}}^P \sum_{t \in \bar{T}} \left(\alpha_{t\bar{s}\hat{s}}^{ES} ES_{t\bar{s}\hat{s}} + \sum_i \alpha_{it}^{PC} PC_{it\bar{s}\hat{s}} \right)}_{TC^2} \quad (24)$$

where we distinguish between demand scenarios ($\bar{s} \in S^D$) and price scenarios

($\hat{s} \in S^P$), and the general scenario index s is replaced by the corresponding (\bar{s}, \hat{s}) -pair. The first-stage and second-stage costs are denoted by TC^1 and TC^2 , respectively, where only the term expressing TC^2 involves price scenarios.

Now we make the following assumption: Given fixed first-stage decisions and fixed product demand, the optimal solution yields the same second-stage decisions for each price scenario \hat{s} , in particular the same electricity purchase from the spot market, $ES_{t\bar{s}\hat{s}}$, and product purchase, $PC_{it\bar{s}\hat{s}}$, i.e. $ES_{t\bar{s}\hat{s}} = \overline{ES}_{t\bar{s}} \forall t, \bar{s}$ and $PC_{it\bar{s}\hat{s}} = \overline{PC}_{it\bar{s}} \forall i, t, \bar{s}$. With this assumption, TC^2 can be rewritten as

$$\begin{aligned}
TC^2 &= \sum_{\bar{s} \in S^D} \varphi_{\bar{s}}^D \sum_{\hat{s} \in S^P} \varphi_{\hat{s}}^P \sum_{t \in \overline{T}} \left(\alpha_{t\bar{s}\hat{s}}^{\text{ES}} \overline{ES}_{t\bar{s}} + \sum_i \alpha_{it}^{\text{PC}} \overline{PC}_{it\bar{s}} \right) \\
&= \sum_{\bar{s} \in S^D} \varphi_{\bar{s}}^D \sum_{t \in \overline{T}} \left[\left(\sum_{\hat{s} \in S^P} \varphi_{\hat{s}}^P \alpha_{t\bar{s}\hat{s}}^{\text{ES}} \right) \overline{ES}_{t\bar{s}} + \sum_i \left(\sum_{\hat{s} \in S^P} \varphi_{\hat{s}}^P \right) \alpha_{it}^{\text{PC}} \overline{PC}_{it\bar{s}} \right] \quad (25) \\
&= \sum_{\bar{s} \in S^D} \varphi_{\bar{s}}^D \sum_{t \in \overline{T}} \left[\mathbb{E} \left(\alpha_{t\bar{s}}^{\text{ES}} \right) \overline{ES}_{t\bar{s}} + \sum_i \alpha_{it}^{\text{PC}} \overline{PC}_{it\bar{s}} \right]
\end{aligned}$$

with $\mathbb{E}(\alpha_{t\bar{s}}^{\text{ES}})$ denoting the expected value of $\alpha_{t\bar{s}}^{\text{ES}}$. Since the constraints do not change for different price scenarios, Eq. (25) implies that under the given assumption, the stochastic model considering both demand and price uncertainty is equivalent to the model considering only demand uncertainty. In this case, both formulations will result in the same optimal first-stage decisions, which means that accounting for price uncertainty is not necessary.

The important question remains whether the assumption of equal second-stage decisions is really valid. It turns out that the assumption approximately holds true because of a particular characteristic of the electricity price uncertainty. From Figure 8, one can see that the electricity prices in different scenarios may differ considerably in magnitude; however, all price profiles follow essentially the same trend. In other words, the times during which the price is low or high compared to the rest of the price curve are approximately the same in all scenarios. For fixed first-stage decisions and fixed demand (i.e. for a specific demand scenario \bar{s}), the total required electricity purchase from the spot market and additional product purchase are approximately constant, i.e. $\sum_{t \in \overline{T}} ES_{t\bar{s}\hat{s}} \approx \text{const.} \forall \hat{s}$ and $\sum_{t \in \overline{T}} PC_{it\bar{s}\hat{s}} \approx \text{const.} \forall i, \hat{s}$. Then the second-stage cost only depends on the distribution of electricity and product purchases over time, which is mainly affected by the price trend rather than the price value. Hence, $ES_{t\bar{s}\hat{s}}$ tend to take the same values in all price scenarios, which have almost identical price trends. The same applies trivially to $PC_{it\bar{s}\hat{s}}$ since the product purchasing price does not depend on the electricity price scenario.

The analysis described above explains why no significant benefit is gained by

considering electricity price uncertainty in risk-neutral optimization. However, this does not apply to risk-averse optimization since the objective function here also includes the CVaR, which only considers the most unfavorable scenarios. In this case, no equivalence between the formulation using the expected electricity price and the formulation considering different price scenarios can be deduced. In fact, the results from the case studies show that in risk-averse optimization, accounting for price uncertainty is essential for obtaining a good solution.

9. Conclusions

This work addresses the simultaneous optimization of short-term production scheduling and electricity procurement under uncertainty for continuous power-intensive processes. A discrete-time MILP model has been developed that applies a mode-based formulation to represent the flexibility of the plant and a block contract formulation to model power contracts, from which electricity can be purchased besides the spot market. Two-stage stochastic programming has been applied to model both uncertainty in spot electricity price and product demand. Risk is taken into account by incorporating the CVaR into the model. Furthermore, to reduce the computational effort when solving large-scale problems, scenario reduction and multicut Benders decomposition have been applied.

An illustrative example and a real-world industrial air separation case demonstrate the capability of the proposed model and solution approach. Both risk-neutral optimization (minimization of total expected cost) and risk-averse optimization (maximization of weighted sum of total expected profit and CVaR) have been considered. The case studies show how significant the differences can be between the solutions obtained from deterministic, risk-neutral, and risk-averse optimization. Especially the electricity procurement decisions highly depend on the choice of the model. Also, in the analysis of the results, the quantification of the value of stochastic solution has been emphasized, which has led to the following remarkable insight: In risk-neutral optimization, accounting for electricity price uncertainty in the stochastic model does not result in significant additional benefit. In contrast, in risk-averse optimization, modeling price uncertainty is crucial for obtaining good solutions.

Nomenclature

Indices

b, b' contract blocks

c	power contracts
i	products
j	vertices
m, m', m''	operating modes
p	time-of-use (TOU) periods
s	general scenarios
\bar{s}	product demand scenarios
\hat{s}	electricity price scenarios
t	time periods

Sets

B_c	blocks in contract c
C	power contracts
I	products
J_m	vertices of polytope describing mode m
M	operating modes
P_c	TOU periods for contract c
S	general scenarios
S^D	product demand scenarios
S^P	electricity price scenarios
SQ	predefined sequences of mode transitions
T	time periods, $T = \{-\theta^{\max} + 1, -\theta^{\max} + 2, \dots, 0, 1, \dots, t^{\text{fin}}\}$
\bar{T}	time period in the scheduling horizon, $\bar{T} = \{1, 2, \dots, t^{\text{fin}}\}$
\hat{T}_{cp}	time periods in TOU period p of contract c
TR	possible mode transitions
TR_m^f	modes from which mode m can be directly reached
TR_m^t	modes which can be directly reached from mode m

Parameters

D_{its}	demand for product i in time period t in scenario s [kg]
$\overline{EC}_{cp}^{\max}$	maximum amount of electricity that can be purchased from contract c in one time period within TOU period p [kWh]
\widehat{EC}_{cb}^{\max}	maximum amount of electricity that can be purchased from block b of contract c [kWh]
ES_t^{\max}	maximum amount of electricity that can be purchased from the spot market in time period t [kWh]
IV_i^{fin}	minimum final inventory of product i [kg]
IV_i^{ini}	initial inventory of product i [kg]

IV_i^{\min}	minimum inventory of product i [kg]
IV_i^{\max}	maximum inventory of product i [kg]
R_s	total revenue in scenario s [\$]
v_{mji}	amount of product i produced in one time period at vertex j of mode m [kg]
y_m^{ini}	1 if plant was operating in mode m in the time period before the start of the scheduling horizon
$z_{mm't}^{\text{ini}}$	1 if operation switched from mode m to mode m' at time t before the start of the scheduling horizon
α	confidence level at which the CVaR is defined, $\alpha \in (0, 1)$
α_{ct}^{EC}	unit electricity price for purchasing electricity from contract c in time period t [\$/kWh]
α_{ts}^{ES}	unit electricity price for purchasing electricity from the spot market in time period t in scenario s [\$/kWh]
α_{it}^{PC}	unit cost for purchasing product i in time period t [\$/kg]
β_{cb}^{EC}	unit electricity price for purchasing electricity from block b of contract c [\$/kWh]
γ_{mi}	unit electricity consumption corresponding to product i if plant operates in mode m [kWh/kg]
δ_m	fixed electricity consumption if plant operates in mode m [kWh]
Δt	length of one time period [h]
ζ	weight factor for total expected cost, $\zeta \in [0, 1]$
$\theta_{mm'}$	minimum stay time in mode m' after switching from mode m to m' [Δt]
$\bar{\theta}_{mm'm''}$	fixed stay time in mode m' of the predefined sequence (m, m', m'') [Δt]
θ^{\max}	maximum minimum or predefined stay time in a mode [Δt]
φ_s	probability of scenario s
$\varphi_{\bar{s}}^{\text{D}}$	probability of demand scenario \bar{s}
$\varphi_{\hat{s}}^{\text{P}}$	probability of price scenario \hat{s}

Continuous Variables

BC_c	base cost for purchasing electricity from contract c [\$]
EC_{ct}	amount of electricity purchased from contract c in time period t [kWh]
\overline{EC}_{cp}	amount of electricity purchased in each time period during TOU period p of contract c [kWh]
\widehat{EC}_c	amount of electricity purchased from contract c [kWh]
\overline{EC}_{cb}	amount of electricity purchased from block b of contract c [kWh]
ES_{ts}	amount of electricity purchased from the spot market in time period t in scenario s [kWh]
EU_{ts}	amount of electricity used by the plant in time period t in scenario

	s [kWh]
EW_{ts}	amount of electricity wasted in time period t in scenario s [kWh]
CV	conditional value-at-risk (CVaR) [\$]
IV_{its}	inventory of product i at time t in scenario s [kg]
PC_{its}	amount of product i purchased in time period t in scenario s [kg]
PD_{its}	amount of product i produced in time period t in scenario s [kg]
\overline{PD}_{mits}	amount of product i produced in mode m in time period t in scenario s [kg]
PW_{its}	amount of product i discarded in time period t in scenario s [kg]
SL_{its}	amount of product i sold in time period t in scenario s [kg]
TC	total expected cost [\$]
κ	auxiliary variable for modeling CVaR
$\lambda_{m,jts}$	coefficient for vertex j in mode m in time period t in scenario s
ω_s	auxiliary variable for modeling CVaR associated with scenario s

Binary Variables

x_{cb}	1 if electricity is purchased from block b of contract c
y_{mt}	1 if plant operates in mode m in time period t
$z_{mm't}$	1 if plant operation switched from mode m to mode m' at time t

Boolean Variables

X_{cb}	true if electricity is purchased from block b of contract c
----------	---

Acknowledgements

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

References

- Aggarwal, S. K., Saini, L. M., Kumar, A., 2009. Electricity price forecasting in deregulated markets: A review and evaluation. *International Journal of Electrical Power & Energy Systems* 31 (1), 13–22.
- Ashok, S., 2006. Peak-load management in steel plants. *Applied Energy* 83 (5), 413–424.
- Babu, C., 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.
- Benders, J. F., 1962. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik* 4 (1), 238–252.
- Beraldi, P., Violi, A., Scordino, N., Sorrentino, N., 2011. Short-term electricity procurement: A rolling horizon stochastic programming approach. *Applied Mathematical Modelling* 35 (8), 3980–3990.
- 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.
- Carrión, M., Conejo, A. J., Arroyo, J. M., 2007a. Forward Contracting and Selling Price Determination for a Retailer. *IEEE Transactions on Power Systems* 22 (4), 2105–2114.
- Carrión, M., Philpott, A. B., Conejo, A. J., Arroyo, J. M., 2007b. A Stochastic Programming Approach to Electric Energy Procurement for Large Consumers. *IEEE Transactions on Power Systems* 22 (2), 744–754.
- 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., Carrión, M., 2006. Risk-constrained electricity procurement for a large consumer. *IEE Proceedings - Generation, Transmission and Distribution* 153 (4), 407.

- Conejo, A. J., Fernández-González, J. J., Alguacil, N., 2005. Energy procurement for large consumers in electricity markets. *IEE Proceedings-Generation, Transmission and Distribution* 152 (3), 357–364.
- Conejo, A. J., García-bertrand, R., Carrión, M., Caballero, A., de Andrés, A., 2008. Optimal Involvement in Futures Markets of a Power Producer. *IEEE Transactions on Power Systems* 23 (2), 703–711.
- 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.
- GAMS Development Corporation, 2015. GAMS version 24.4.1.
- Garcia-Herreros, P., Wassick, J. M., Grossmann, I. E., 2014. Design of Resilient Supply Chains with Risk of Facility Disruptions. *Industrial & Engineering Chemistry Research*.
- Hand, M. M., Baldwin, S., DeMeo, E., Reilly, J. M., Mai, T., Arent, D., Porro, G., Meshek, M., Sandor, D., 2012. Renewable Electricity Futures Study. Tech. rep., National Renewable Energy Laboratory.
- Hatami, A. R., Seifi, H., Sheikh-El-Eslami, M. K., 2009. Optimal selling price and energy procurement strategies for a retailer in an electricity market. *Electric Power Systems Research* 79 (1), 246–254.
- Heitsch, H., Römisch, W., 2003. Scenario Reduction Algorithms in Stochastic Programming. *Computational Optimization and Applications* 24, 187–206.
- Hwang, C.-L., Masud, A. S. M., 1979. Multiple objective decision making - methods and applications. Springer.
- 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., Kebulis, M. F., 2007. Operations planning with real time pricing of a primary input. *Computers & Operations Research* 34 (3), 848–867.
- Li, Z., Floudas, C. a., 2014. Optimal scenario reduction framework based on distance of uncertainty distribution and output performance: I. Single reduction via mixed integer linear optimization. *Computers & Chemical Engineering* 70, 50–66.

- Lima, R. M., Novais, A. Q., Conejo, A. J., 2015. Weekly self-scheduling, forward contracting, and pool involvement for an electricity producer. An adaptive robust optimization approach. *European Journal of Operational Research* 240 (2), 457–475.
- Merkert, L., Harjunkski, I., Isaksson, A., Säynevirta, S., Saarela, A., Sand, G., 2014. Scheduling and energy – Industrial challenges and opportunities. *Computers & Chemical Engineering* 72, 183–198.
- Mitra, S., Grossmann, I. E., Pinto, J. M., Arora, N., 2012. Optimal production planning under time-sensitive electricity prices for continuous power-intensive processes. *Computers & Chemical Engineering* 38, 171–184.
- 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.
- Nogales, F. J., Contreras, J., Conejo, A. J., Espínola, R., 2002. Forecasting Next-Day Electricity Prices by Time Series Models. *IEE* 17 (2), 342–348.
- Paulus, M., Borggreffe, F., 2011. The potential of demand-side management in energy-intensive industries for electricity markets in Germany. *Applied Energy* 88 (2), 432–441.
- R Core Team, 2014. R: A Language and Environment for Statistical Computing.
- Rockafellar, R. T., 2007. Coherent Approaches to Risk in Optimization Under Uncertainty. *INFORMS Tutorials in Operations Research* (March 2015), 38–61.
- Rockafellar, R. T., Uryasev, S., 2000. Optimization of Conditional Value-at-Risk. *Journal of risk* 2, 21–42.
- Samad, T., Kiliccote, S., 2012. Smart grid technologies and applications for the industrial sector. *Computers & Chemical Engineering* 47, 76–84.
- Sarykalin, S., Serraino, G., Uryasev, S., 2008. Value-at-Risk vs . Conditional Value-at-Risk in Risk Management and Optimization. *Tutorials in Operations Research* (February 2015), 270–294.
- Skar, C., Doorman, G., Tomasgard, A., 2014. Large-scale power system planning using enhanced Benders decomposition. In: *Power Systems Computation Conference (PSCC)*, 2014. IEEE, Wroclaw.

- You, F., Grossmann, I. E., 2013. Multicut Benders decomposition algorithm for process supply chain planning under uncertainty. *Annals of Operations Research* 210, 191–211.
- Zare, K., Moghaddam, M. P., Sheikh El Eslami, M. K., 2010. Electricity procurement for large consumers based on Information Gap Decision Theory. *Energy Policy* 38 (1), 234–242.
- Zhang, Q., Heuberger, C. F., Grossmann, I. E., Sundaramoorthy, A., Pinto, J. M., 2014. Air Separation with Cryogenic Energy Storage: Optimal Scheduling Considering Electric Energy and Reserve Markets (to appear in *AIChE Journal*).
- Zhang, Q., Sundaramoorthy, A., Grossmann, I. E., Pinto, J. M., 2015. A discrete-time scheduling model for continuous power-intensive process networks with various power contracts (submitted for publication).