WU
WIRTSCHAFTS
UNIVERSITÄT
WIEN VIENNA
UNIVERSITY OF
ECONOMICS
AND BUSINESS

EFMD
EQUIS
ACCREDITED

ePub^{WU} Institutional Repository

Nils Löhndorf and David Wozabal and Stefan Minner

Optimizing Trading Decisions for Hydro Storage. Systems using Approximate Dual Dynamic Programming

Article (Draft)

http://epub.wu.ac.at/

Optimizing Trading Decisions for Hydro Storage Systems using Approximate Dual Dynamic Programming

Nils Löhndorf[1], David Wozabal[2], Stefan Minner[2]

[1]Vienna University of Economics and Business, Vienna, Austria
[2]Technische Universität München, Munich, Germany
{nils.loehndorf@wu.ac.at, david.wozabal@wu.ac.at, stefan.minner@tum.de}

We propose a new approach to optimize operations of hydro storage systems with multiple connected reservoirs whose operators participate in wholesale electricity markets. Our formulation integrates short-term intraday with long-term interday decisions. The intraday problem considers bidding decisions as well as storage operation during the day and is formulated as a stochastic program. The interday problem is modeled as a Markov decision process of managing storage operation over time, for which we propose integrating stochastic dual dynamic programming with approximate dynamic programming. We show that the approximate solution converges towards an upper bound of the optimal solution. To demonstrate the efficiency of the solution approach, we fit an econometric model to actual price and inflow data and apply the approach to a case study of an existing hydro storage system. Our results indicate that the approach is tractable for a real-world application and that the gap between theoretical upper and a simulated lower bound decreases sufficiently fast.

Key words: OR in Energy, Stochastic Programming, Markov Decision Processes, Approximate Dynamic Programming

History:

1. Introduction

The steady increase of electricity from intermittent sources of renewable energy poses challenges for the electrical grid. A key component of a more flexible, smarter grid is the ability to store electricity and thereby to decouple electricity generation from electricity consumption. The most common large-scale storage technology for electricity is hydro storage. A hydro storage power plant either stores the natural flow of water or pumps water into an elevated reservoir to be able to release the water and produce electricity when it is needed. Hydro storage systems thereby offer the

ability to buffer the intermittent supply of electricity from renewable power sources such as wind, solar, or run-of-river. The European electricity mix, for example, consists of 15 percent hydropower with a total capacity of 260 gigawatts of which 45 gigawatts are pumped-hydro storage (Auer 2011, Zuber 2011). The growing share of renewable energies increases Europe's demand for storage, and generating companies are currently investing about 26 billion Euros into new pumped-hydro storage plants with a total capacity of 27 gigawatts (Zuber 2011).

Today, a large share of electricity is sold in wholesale electricity markets. Most generating companies in Austria, Germany, and Switzerland, for example, sell electricity at the European Power Exchange (EPEX SPOT), which is one of the largest European electricity markets. Since supply and demand have to be synchronized in advance, EPEX SPOT provides different types of forward markets, of which the day-ahead market and the intraday market are the most important markets for owners of storage plants. At the day-ahead market, producers place supply bids and consumers place demand bids for each hour of the following day, i.e., one day ahead of delivery. After the day-ahead market is closed, the intraday market allows market participants to clear imbalances that arise during the day up to 45 minutes before delivery. In case actual volumes deviate from day-ahead or intraday bids, all remaining imbalances are automatically cleared at the balancing market with a high risk of additional cost. A generating company with pumped hydro storage capacities tries to buy electricity at the market when the price is low and sell electricity when the price is high, while trying to mitigate the risk of positive and negative imbalances.

Trading with a system of hydro storage plants in a wholesale electricity market involves many decisions as well as a great deal of uncertainty. In particular, there exist two major challenges to solve the problem efficiently. First, not only are day-ahead and intraday prices uncertain, but also the development of electricity prices over time as well as the inflow of water into the reservoirs. Second, a system of hydro storage plants with multiple reservoirs requires a coordinated water release policy, since upstream releases influence downstream reservoir levels. In addition to day-ahead and intraday bidding decisions, a generating company has to decide about water releases from multiple reservoirs over time. Future decisions and states of the system as well as their

probabilities therefore have to be considered in the decision-making process, which turns storage operation into a complex stochastic-dynamic decision problem.

In the literature, the day-ahead bidding problem is typically modeled as a two-stage stochastic program, with bidding decisions at the first stage and price realizations as well as operational decisions at the second stage (Baíllo et al. 2004, Fleten and Kristoffersen 2007, García-González et al. 2007). A two-stage approach is well-suited for optimizing bidding decisions in the short-term, but does not anticipate future storage states and decisions. To optimize bidding decisions over a longer planning horizon, we have to solve a multi-stage stochastic programming problem. For this class of problems, two basic solution strategies have emerged in the literature. One strategy is to construct a scenario tree to represent uncertainty and solve the problem as one large mathematical program (Heitsch and Römisch 2003, Eichhorn et al. 2009, Hochreiter and Wozabal 2010). This strategy can handle discrete decisions as well as any type of exogenous stochastic process, but is limited to problems with a small number of stages. Fleten and Kristoffersen (2008) and Matevosyan et al. (2009) propose mixed-integer programs of hydro storage operation where a scenario tree is used to model uncertainty over a weekly planning horizon. A comparison of solution methods for a tree-based stochastic unit commitment problem is given in Cerisola et al. (2009).

Another strategy is based on formulating the problem as a dynamic program and then applying Benders' decomposition to recursively construct the value function at each stage around a set of sample decisions (Pereira and Pinto 1991). This strategy, also known as *stochastic dual dynamic programming* (SDDP), can handle problems with a large number of stages as long as the optimization problem at each stage is convex and the stochastic process stagewise independent. Most SDDP formulations of hydro storage operation only consider inflow or demand uncertainty, e.g., Flach et al. (2010), Philpott and de Matos (2012). To the best of our knowledge, the only SDDP formulation that also considers price uncertainty is given in Gjelsvik et al. (2010). However, the authors only model weekly price averages, which keeps the problem tractable but also underrates the short-term value of storage.

In contrast to most previous approaches, the model proposed in this work decomposes the multi-stage problem into an *intrastage* and an *interstage* problem (Pritchard et al. 2005). Day-ahead bidding decisions as well as hourly reservoir operations are modeled as part of the intrastage problem which is formulated as a stochastic mixed-integer quadratic program with randomness in the objective function. Decisions about reservoir contents at the end of the day, on the other hand, are modeled as part of the interstage problem which is formulated as a Markov decision process (MDP). The proposed decomposition severely reduces the complexity of the problem. Since day-ahead prices realize simultaneously, we can view day-ahead price vectors as realizations of intraday randomness. Day-ahead price distributions can thus be modeled as separate random variables conditioned on a small number of explanatory variables which are defined as the state of an exogenous Markov process. This allows us to describe the dynamics of the hourly electricity prices by a discrete state transition process which is represented by a probability lattice.

To solve the problem efficiently, we integrate SDDP with ideas from approximate dynamic programming (ADP). ADP algorithms simulate the state transition process of an MDP and use the sampled information to approximate the high-dimensional value function by a function of much lower complexity (Powell 2011). An ADP algorithm to optimize day-ahead bidding and storage decisions is also proposed in Löhndorf and Minner (2010), but only for a single storage unit and a single bidding decision per stage. By contrast, the present approach allows us to model storage systems with multiple units as well as hundreds of decision variables per stage.

In the same way as SDDP, the proposed solution method iteratively solves the decision problem using forward simulation to sample candidate decisions and backward recursion to construct an approximation of the value function. Unlike SDDP, however, the method does not require stagewise independence of the stochastic process, but rather assumes that randomness follows a Markov process.

To solve the problem numerically, the solution approach pursues a three-fold strategy to approximate the value function of the MDP. First, the continuous-state Markov process that describes

the evolution of environmental variables over time is reduced to a probability lattice. Second, candidate decisions that do not improve the approximation quality by more than a given epsilon are discarded. Third, the approach uses a relaxed version of the problem to approximate the value function, but evaluates the decision policy based on the original problem formulation. To emphasize the focus on approximation, we refer to the solution method as *approximate dual dynamic programming* (ADDP).

For the problem at hand, the approximated value function is more optimistic than the true value function with respect to the future value of water that remains in the reservoirs at the end of the day. However, if the resulting gap is negligible, the relaxation provides an efficient solution to a complex optimization problem.

The paper is organized as follows. In Section 2, we present the model formulation of the multi-stage decision problem. In Section 3, we introduce a relaxed version of the problem, describe the solution algorithm, prove its convergence, and derive an error bound of the approximation. In Section 4, we propose an econometric electricity price model for the EPEX SPOT market. In Section 5, we apply the model and the solution algorithm to a case study of an existing hydro storage system in Austria. Finally, in Section 6, we summarize the results and discuss possible directions for future research.

2. Model Formulation

2.1. Assumptions

We consider the stochastic unit commitment problem of a power generating company that operates a network of hydro storage plants and participates in a wholesale electricity market. The objective of the company is to maximize expected profits from buying and selling electricity while operating its hydro resources efficiently.

We assume that the company is planning storage operation over an entire year, but schedules its resources on an hourly basis. Uncertainty enters the planning problem through stochastic natural inflows into the reservoirs as well as through stochastic electricity prices. We assume that the

dynamics of the random variables can be separated into an interday and an intraday process. The interday process is characterized by a state transition model of environmental variables, e.g., temperature, renewable power production, fuel prices, natural inflows, as well as calendar information, i.e., day of the year, day of the week. This process is assumed to be a Markov chain, possibly non-homogeneous, which is represented by a probability lattice. The intraday process describes random hourly electricity prices which depend on the realization of interday randomness.

In line with the case study presented in Sections 4 and 5, all assumptions regarding the electricity market are made with the EPEX SPOT market in mind. We assume that the electricity market implements a multi-settlement system with a day-ahead, an intraday (hour-ahead) and a balancing (real-time) market. The company makes the majority of its trades in the day-ahead market where it places price-dependent supply and demand bids by submitting piecewise-linear bidding curves for each hour of the following day. After the day-ahead market is closed, the system operator announces a clearing price for each hour using a uniform auction mechanism. Day-ahead bidding therefore takes place under price and volume uncertainty.

In case produced volumes deviate from day-ahead bids, the company clears all foreseeable imbalances at the intraday market and does not deliberately use the balancing market. We do not explicitly model the cost of balancing in case of unplanned outages.

We assume that the generating company is a price-taker in the day-ahead market, but a price-setter in the intraday market. Actual EPEX SPOT sales volumes in 2012 support this assumption (mean day-ahead: 28 gigawatts, mean intraday: 1.3 gigawatts).

Moreover, we assume that expected day-ahead prices equal expected intraday prices. Price data from 2011 and 2012 also supports this assumption, with the mean day-ahead price being €46.90, and the volume weighted mean intraday price being €47.21 with mean low and high prices of €37.39 and €57.52, respectively. This implies that, although the storage operator is aware of the price effect on the intraday market, she cannot exert market power, since the price always turns against the operator in expectation, i.e., the expected intraday price is higher than the realized day-ahead price when electricity is being bought and lower when it is being sold. Introducing a price

response moreover reflects the propensity of the company to trade day-ahead without introducing additional bias. In this setup, the company has little incentive to trade in the intraday market, so that we assume that the company makes all intraday trading decisions at once when day-ahead prices are known.

The topology of the network of hydro storage plants is convergent, so that each reservoir may have multiple inflows but only a single outflow and is associated with a single turbine and possibly a pump. These assumptions could easily be relaxed and are made to keep the notation simple. Moreover, connected reservoirs are close, so that there are no significant delays regarding the flow of water from one reservoir to another. We assume that head effects can be ignored so that the power conversion function only depends on water release per time unit but not on reservoir levels. The natural inflow of water into a reservoir is state-dependent and remains constant throughout the day.

2.2. Markov Decision Process

We model the interday decision process of storage operation as a finite-horizon Markov decision process (MDP) with decision epoch of one day. Denote t as the time index for a day of the year. Randomness is separated into a process of environmental variables $(S_t)_{t=1}^T$ and a process of hourly intraday electricity prices $(P_t)_{t=1}^T$. We assume that the state of the MDP $(S_t)_{t=1}^T$ influences the electricity prices and that the distributions $P_t|S_t$ are known and such that $S_{t+1}|S_t$ is independent of $P_t|S_t$.

The objective of the generating company is to maximize its discounted expected profits for a given environmental state $S_t \in \mathcal{S}_t$ and initial storage states $R_{t-1} \in \mathcal{R}$ in stage $t \in \{1, \ldots, T\}$, with $\mathcal{S}_t$ being the set of environmental states in t and $\mathcal{R}$ being the set of all possible reservoir states. Denote $\mathbb{P}(S_{t+1}|S_t)$ as the state transition probability of the Markov process. Let $\pi = \{\pi_1, \ldots, \pi_T\}$ be a decision policy that encompasses all operational decision variables, i.e., all bidding and dispatch decisions, subject to the state-dependent feasible set $\Pi_t(S_t, R_{t-1})$ (see Section 2.3), and define $C(S_t, R_{t-1}, \pi_t)$ as the random intraday profit (contribution) and γ as discount factor. Note that,

given the state of the MDP, the intraday profit is a random variable, since it depends on the random prices $P_t | S_t$. For R_0 and V_{T+1} fixed, the value of being in state S_t with initial reservoir states R_{t-1} is given by the optimality equations

$$
V_t(S_t, R_{t-1}) = \max_{\pi_t \in \Pi_t(S_t, R_{t-1})} \Big\{ \mathbb{E}\Big[C(S_t, R_{t-1}, \pi_t)
$$
$$
+ \gamma \sum_{S_{t+1} \in \mathcal{S}_{t+1}} \mathbb{P}(S_{t+1}|S_t) \, V_{t+1}\big(S_{t+1}, R_t(\pi_t)\big) \Big] \Big\}, \quad (1)
$$

for $S_t \in \mathcal{S}_t$, $R_{t-1} \in \mathcal{R}$ and $t = 1, \ldots, T$. Since π_t assigns a decision to every realization of intraday randomness, it results in random reservoir states $R_t(\pi_t)$. An optimal decision policy maximizes the sum of expected intraday profits and expected future profits. Future profits depend on the random state transition from S_t to S_{t+1} as well as the (random) final reservoir state $R_t = R_t(\pi_t)$ in t which is the initial reservoir state in $t + 1$. For notational convenience, we suppress the dependence of R_t on π_t and Π_t on S_t, R_{t-1} where no confusion can arise.

In line with Powell (2011), let us reformulate (1) using the post-decision state. Denote $\bar{V}_t$ as value function around the post-decision state, which gives us the value of being in state S_t at the end of the day after realization of R_t but before a random transition to the next state. For a fixed function $\bar{V}_T$, the post-decision value function is

$$
\bar{V}_t(S_t, R_t) = \sum_{S_{t+1} \in \mathcal{S}_{t+1}} \mathbb{P}(S_{t+1}|S_t) \, V_{t+1}(S_{t+1}, R_t)
$$
$$
= \sum_{S_{t+1} \in \mathcal{S}_{t+1}} \mathbb{P}(S_{t+1}|S_t) \max_{\pi_{t+1} \in \Pi_{t+1}} \Big\{ \mathbb{E}\Big[C(S_{t+1}, R_t, \pi_{t+1}) + \gamma \bar{V}_{t+1}(S_{t+1}, R_{t+1}) \Big] \Big\}, \quad (2)
$$

for $S_t \in \mathcal{S}_t$, $R_t \in \mathcal{R}$ and $t = 1, \ldots, T - 1$. This formulation of the optimality equations is equivalent to (1) but provides us with a computational advantage, because the expectation operator associated with the state transition is now outside of the maximization problem. For now, let us assume that $\bar{V}_t(S_t, R_t)$ is known. In Section 3, we are going to show how to recursively build an approximation of the post-decision value function.

2.3. Stochastic Programming Formulation

For a given post-decision value function, the intraday problem can be formulated as a stochastic program with recourse, with the objective to maximize the expected profit for a given state S_t
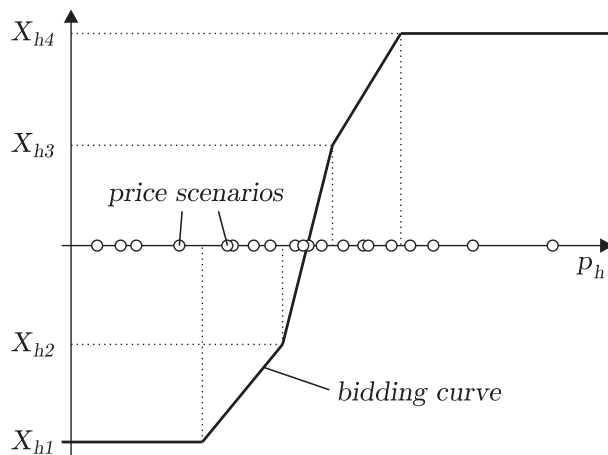
Figure 1 Bidding curve with four breakpoints and four price scenarios per segment

and initial reservoir state R_{t-1}. Although all variables and random parameters of the intraday problem depend on S_t, we suppress this dependence to streamline the presentation. Furthermore, most parameters and decision variables depend on an hour h and a scenario s. We indicate this by adding the corresponding subscripts.

2.3.1. First-Stage Decision: Day-Ahead Bidding The first stage of the stochastic program involves fixing price-dependent bidding curves (see Figure 1). For each hour $h \in \mathscr{H} = \{1, \dots, 24\}$ of the following day, the generating company submits I price-volume pairs (ρ_{hi}, X_{hi}), with $\rho_{hi} < \rho_{h,i+1}$ and $X_{hi} \leq X_{h,i+1}$ for $i < I$. A linear interpolation of these pairs yields a monotone increasing, piecewise-linear function that maps price realizations to day-ahead sales volumes. However, choosing prices and volumes simultaneously yields a non-convex decision problem. In line with Fleten and Kristoffersen (2007), we therefore fix the price points in advance and only decide the day ahead volumes for each price (in megawatt hours).

Denote $p_s \in \mathbb{R}^{24}$, $s \in \mathscr{S} = \{1, \dots, K\}$ as the finite set of realizations of the price process $P_t|S_t$ and p_{sh} as a day-ahead price realization in hour h. We assume $\mathbb{P}(P_t = p_s|S_t) = 1/K$, so that each price scenario has equal probability. The realized day-ahead sales volume x_{sh}^d in scenario s depends on the bidding curve as well as the realized day-ahead prices,

$$
x_{sh}^d = \begin{cases} X_{h1} & \text{if } p_{sh} < \rho_{h1}, & \forall\, s \in \mathscr{S}, h \in \mathscr{H}, \\ X_{h,i-1} + \frac{X_{hi} - X_{h,i-1}}{\rho_{hi} - \rho_{h,i-1}}(p_{sh} - \rho_{h,i-1}) & \text{if } \rho_{h,i-1} \leq p_{sh} < \rho_{hi},\ 1 < i \leq I, & \forall\, s \in \mathscr{S}, h \in \mathscr{H}, \quad (3) \\ X_{hI} & \text{if } p_{sh} \geq \rho_{hI}, & \forall\, s \in \mathscr{S}, h \in \mathscr{H}. \end{cases}
$$

Depending on the sign of the bid, each sales volume either represents an offer $(X_{hi} \geq 0)$ or a purchase $(X_{hi} < 0)$ of electricity (in megawatt hours). To additionally enforce monotonicity, we include the following constraint,

$$X_{h,i-1} \leq X_{hi} \quad \forall\, h \in \mathscr{H}, i \in \{2,\ldots,I\}. \tag{4}$$

Instead of spreading the price breakpoints ρ_{hi} evenly over the price domain, we assign each line segment approximately the same number of price scenarios. More specifically, we compute the breakpoints by first partitioning the set of sorted prices into $I+1$ subsets and then calculate ρ_{hi} as the mean of the adjacent prices that are not of the same subset, i.e.,

$$\rho_{hi} = \frac{p'_{g(i)-1,h} + p'_{g(i)h}}{2}, \quad g(i) = \left\lfloor \frac{iK}{I+1} + 1 \right\rfloor \quad \forall\, h \in \mathscr{H}, i \in \mathscr{I}, \tag{5}$$

with $p'_{g(i)h}$ as the day-ahead prices sorted in ascending order. The resulting bidding curve is smoother in areas where the probability mass is high and coarser where the probability mass is low. Note that the number of scenarios K has to satisfy $K \geq 2I+2$ to ensure non-anticipativity of the bids.

2.3.2. Second-Stage Decision: Short-Term Unit Commitment   Short-term operational decisions are modeled at the second stage of the stochastic program. At this stage, day-ahead sales volumes have realized and the generating company uses either storage capacities or the intraday market to close its positions. In line with other authors (e.g., Fleten and Kristoffersen 2007, García-González et al. 2007), we model the unit commitment problem as a mixed-integer program.

Denote $\mathscr{J} = \{1,\ldots,J\}$ as the set of reservoirs, c_{sjh} as charge into reservoir j, and d_{sjh} as discharge from reservoir j (in metric tons). The topology of the reservoir network is defined by matrix $A = (A_{jk}) \in \{-1,0,1\}^{J \times J}$, with $A_{jk} = 1$ if water can be released from j into k and $A_{kj} = -1$ if water can be pumped from k into j. The hourly natural inflow of water into reservoir j is given by IN_{tj} on day t. In contrast to electricity prices, natural inflows are assumed to be deterministic given the state of the MDP and constant in all hours of a day. Denote r_{sjh} as the reservoir state

with $r_{sj0} = R_{t-1,j}$ and o_{sjh} as the overflow (or spill). Then, the storage state at the end of hour h is given by the following balance equation

$$r_{sjh} = r_{sj,h-1} - d_{sjh} + c_{sjh} - o_{sjh}$$

$$+ \sum_{k \in \mathscr{J}: A_{kj}=1} (d_{skh} + o_{skh}) - \sum_{k \in \mathscr{J}: A_{jk}=-1} c_{skh} + \mathrm{IN}_{tj} \quad \forall\, s \in \mathscr{S}, j \in \mathscr{J}, h \in \mathscr{H}. \tag{6}$$

The generating company must balance realized day-ahead and intraday bids with power generation and consumption. Denote x_{sh}^b as the amount of power (in megawatt hours) sold or purchased at the intraday market in hour h and scenario s. All open positions are closed if

$$x_{sh}^d + x_{sh}^b = \sum_{j \in \mathscr{J}} (\eta_j^+ d_{sjh} - \eta_j^- c_{sjh}) \quad \forall\, s \in \mathscr{S}, h \in \mathscr{H}, \tag{7}$$

with η_j^+ and η_j^- as constant power conversion factors (in megawatts per metric ton of water) which relate flow volume to power quantity.

Charge and discharge decisions are constrained by minimum and maximum capacities of pumps and turbines. Denote $[\mathrm{LB}_j^R, \mathrm{UB}_j^R]$ as the allowed reservoir content range j (in metric tons), $[\mathrm{LB}_j^+, \mathrm{UB}_j^+]$ as power limits of the j-th turbine, and $[\mathrm{LB}_j^-, \mathrm{UB}_j^-]$ as power limits of the j-th pump (in megawatts). Then,

$$\mathrm{LB}_j^R \le r_{sjh} \le \mathrm{UB}_j^R \qquad\qquad \forall\, s \in \mathscr{S}, j \in \mathscr{J}, h \in \mathscr{H}, \tag{8}$$

$$z_{sjh}^+ \mathrm{LB}_j^+ \le \eta_j^+ d_{sjh} \le z_{sjh}^+ \mathrm{UB}_j^+ \qquad\qquad \forall\, s \in \mathscr{S}, j \in \mathscr{J}, h \in \mathscr{H}, \tag{9}$$

$$z_{sjh}^- \mathrm{LB}_j^- \le \eta_j^- c_{sjh} \le z_{sjh}^- \mathrm{UB}_j^- \qquad\qquad \forall\, s \in \mathscr{S}, j \in \mathscr{J}, h \in \mathscr{H}, \tag{10}$$

with binary variables, z_{sjh}^+ and z_{sjh}^-, to model the on/off status of turbines and pumps, respectively.

2.3.3. Objective Function The objective of the generating company is to maximize its expected intraday profits through efficient bidding and storage operation while considering the expected future value of storage as defined by the post-decision value function. Following Pereira and Pinto (1991), we model the post-decision value v_s as a concave, piecewise-linear function of the final reservoir states r_{sj24} at the end of the day. Note that r_{sj24} is a realization of the j-th

element of R_t in (1) and (2). For a given state S_t, the post-decision value function is defined as the minimum of a set of hyperplanes $\mathcal{N} = \{1, \ldots, N\}$ with intercepts $a_n(S_t)$ and slopes $b_{nj}(S_t)$ so that the future value of storage is given by

$$v_s = \min_{n \in \mathcal{N}} \left\{ a_n(S_t) + \sum_{j \in \mathcal{J}} b_{nj}(S_t) r_{sj24} \right\} \qquad \forall\, s \in \mathcal{S}. \tag{11}$$

If we add v_s to the objective function of a maximization problem, we can reformulate (11) by the following set of linear constraints,

$$v_s \leq a_n(S_t) + \sum_{j \in \mathcal{J}} b_{nj}(S_t) r_{sj24} \qquad \forall\, n \in \mathcal{N}, s \in \mathcal{S}. \tag{12}$$

Denote $p_{sh} - \beta x_{sh}^b$ as the expected intraday price, with $\beta \geq 0$ as the slope of the price-response function. For a given state S_t and an initial reservoir state R_{t-1}, we formulate the optimization problem as the following stochastic mixed-integer quadratic program

$$V_t(S_t, R_{t-1}) = \max \frac{1}{K} \sum_{s=1}^{K} \sum_{h=1}^{24} \left(\left(p_{sh} x_{sh}^d + (p_{sh} - \beta x_{sh}^b) x_{sh}^b \right) + \gamma v_s \right) \tag{13}$$

$$\text{s.t.} \quad (3), (4), (6), (7), (8), (9), (10), (12);$$

$$X_{hi} \in \mathbb{R}\ \forall\, h \in \mathcal{H}, i \in \{1, \ldots, I\};\ x_{sh}^d, x_{sh}^b \in \mathbb{R}\ \forall\, s \in \mathcal{S}, h \in \mathcal{H};$$

$$r_{sjh}, c_{sjh}, d_{sjh}, o_{sjh} \geq 0\ \forall\, s \in \mathcal{S}, j \in \mathcal{J}, h \in \mathcal{H};$$

$$z_{sjh}^+, z_{sjh}^- \in \{0, 1\}\ \forall\, s \in \mathcal{S}, j \in \mathcal{J}, h \in \mathcal{H};$$

$$v_s \in \mathbb{R}\ \forall\, s \in \mathcal{S}.$$

3. Solution Methods

To obtain the hyperplanes required in (11) or (12), we integrate stochastic dual dynamic programming (SDDP) with ideas from approximate dynamic programming (ADP). The method referred to as approximate dual dynamic programming (ADDP) constructs a polyhedral approximation of the post-decision value function defined in Section 2.2 by sampling the state transitions of the Markov decision process.

3.1. Approximate Value Function

To be able to construct a polyhedral approximation of the post-decision value function, we relax certain requirements of the original model formulation.

DEFINITION 1. (Relaxed Problem) Define the relaxed problem as the linear relaxation of the stochastic mixed-integer quadratic program in (13) with $\beta = 0$, and denote $V_t'(S_t, R_{t-1})$ as its optimal objective value and $\bar{V}_{t-1}'$ as the corresponding post-decision value function.

By modeling the post-decision value function as the minimum of a set of linear functions (12), we tacitly assume that the true function admits a tight concave approximation. While we cannot make this assertion for the post-decision value function associated with the original model formulation, we can show that concavity holds for the post-decision value function of the relaxed problem.

PROPOSITION 1. *The objective value $V_t'(S_t, R_{t-1})$ as well as the post-decision value $\bar{V}_t'(S_t, R_t)$ are both concave in the reservoir levels.*

Proof. With the binary variables relaxed to $z_{sjh}^+, z_{sjh}^- \in [0, 1]$ and $\beta = 0$, the maximization problem in (13) is an ordinary linear program. From the theory of linear programming we know that a problem of this type is jointly concave in the right-hand sides of its constraints, e.g., by Proposition 2.22 in Rockafellar and Wets (1998). The vector R_t enters the right-hand side of equation (6). Therefore, $V_t'(S_t, R_{t-1})$ is concave in R_{t-1}. Denote $V_T'(S_T, R_{T-1})$ as the objective value of the relaxed problem in the final stage and $\bar{V}_T'$ as an arbitrary piecewise-linear function which is assumed to be concave in R_T. Since $V_T'(S_T, R_{T-1})$ is concave in R_{T-1}, the expected value $\sum_{S_T \in \mathcal{S}} \mathbb{P}(S_T | S_{T-1}) V_T'(S_T, R_{T-1})$ is concave in R_{T-1}. Hence, $\bar{V}_{T-1}'(S_{T-1}, R_{T-1})$ is concave in R_{T-1}. Concavity of $\bar{V}_t'(S_t, R_t)$ for $t = 1, \ldots, T - 2$ follows by backward induction. $\square$

With $\beta = 0$ in the relaxed problem, we assume that the generating company is price-taker in both markets, day-ahead and intraday. Without an intraday price response, however, the risk-neutral company has no incentive to trade in the day-ahead market as long as we assume that the mean intraday price is identical to the realized day-ahead price. Instead, the company could move all of

its trades to the intraday market. In that case, the relaxed version of the stochastic program can

be decomposed into K linear programs, which is supported by the following proposition.

PROPOSITION 2. *Without a price response, i.e., $\beta = 0$, the relaxed problem is the sample average*

of K linear programs.

Proof. For each feasible decision $(x^b_{sh}, x^d_{sh}) = (\bar{x}^b, \bar{x}^d)$, the decision $(x^b_{sh}, x^d_{sh}) = (\bar{x}^b + \bar{x}^d, 0)$ is

feasible with respect to (7), $\forall\ s \in \mathscr{S}, h \in \mathscr{H}$. With $\beta = 0$, the marginal prices of x^d_{sh} and x^b_{sh} in

(13) are identical and the decisions $(\bar{x}^b, \bar{x}^d)$ and $(\bar{x}^b + \bar{x}^d, 0)$ have the same objective values. Hence,

there exists an optimal decision, where $X_{hi} = 0 \ \forall\ h \in \mathscr{H}, i \in \mathscr{I}$. The non-anticipativity constraints

(3) can then be dropped, and the relaxed problem can be decomposed, such that

$$V'_t(S_t, R_{t-1}) = \frac{1}{K} \sum_{s=1}^{K} V'_{ts}(S_t, R_{t-1}),$$

where $V'_{ts}(S_t, R_{t-1})$ is defined as $V'_t(S_t, R_t)$ for a single scenario $s \in \mathscr{S}$.   □

The objective value associated with the relaxed problem is an upper bound of the optimal objec-

tive value, i.e., $V'_t(S_t, R_{t-1}) \geq V_t(S_t, R_{t-1})$. An operational policy, where the generating company

does not bid in the day-ahead market, however, is of little practical use. Nevertheless, as long as

the difference between upper bound and optimum is reasonably small, we can use the optimal solu-

tion of the relaxed problem to construct an approximation of the post-decision value function. We

then use this function inside the original problem formulation to compute near-optimal intraday

decisions. As we will see in Section 5.2, the difference is small for the actual problem considered in

this work.

Let us briefly outline how a polyhedral approximation of the post-decision value function can be

constructed. Since $V'_t(S_t, \ \cdot\ )$ is the optimal objective value of a linear program, the post-decision

value function of the relaxed problem can be described by a concave, piecewise-linear function, i.e.,

by a polyhedral function. We can construct an approximation $\hat{\bar{V}}'_{t-1}(S_{t-1}, R)$ of the post-decision

value function by first defining a set of sample reservoir states, $\{\hat{R}_1, \ldots, \hat{R}_N\}$, with $\hat{R}_n \in \mathcal{R}$, and

then deriving the corresponding hyperplanes going through

$$(\hat{R}_{11}, \ldots, \hat{R}_{1J}, V'_t(S_t, \hat{R}_1)), \ldots, (\hat{R}_{N1}, \ldots, \hat{R}_{NJ}, V'_t(S_t, \hat{R}_N)) \ \forall\ S_t \in \mathcal{S}_t.$$

To obtain the hyperplanes, let $\partial_R V_t(S_t, R_t)$ be the set of super-gradients of the function $R_t \mapsto V'_t(S_t, R_t)$. From this set, we select a super-gradient, $b(S_t) \in \partial_R V'_t(S_t, \hat{R})$, which is the slope of the supporting hyperplane of $V'_t(S_t, \cdot)$ going through $(\hat{R}_1, \ldots, \hat{R}_J, V'_t(S_t, \hat{R}))$. The hyperplane is given by the linear function

$$H(S_t, R; \hat{R}_i) = a(S_t) + b(S_t)^\top R, \quad a(S_t) = V'(S_t, \hat{R}) - \sum_{j \in \mathscr{J}} b_j(S_t) \hat{R}_{ij}, \tag{14}$$

with $a(S_t) \in \mathbb{R}$ as the intercept and $b(S_t) \in \mathbb{R}^J$ as the vector of slopes. Since we are dealing with linear programs, the slopes can be obtained from the dual variables λ associated with constraints (6) for $h = 1$,

$$b_j(S_t) = \sum_{s \in \mathscr{S}} \lambda_{sj1}. \tag{15}$$

The resulting approximate post-decision value function is then given by

$$\hat{\bar{V}}'_{t-1}(S_{t-1}, R) = \min \left\{ \sum_{S_t \in \mathcal{S}_t} \mathbb{P}(S_t | S_{t-1}) \left(a_n(S_t) + b_n(S_t)^\top (R - \hat{R}_n) \right), \ n = 1, ..., N \right\}, \tag{16}$$

where the hyperplane going through $\hat{R}_n$ is the weighted sum of all hyperplanes $H(S_t, R; \hat{R}_n)$ over all successor states. For a given set of sample reservoir states at each stage, a polyhedral approximation of the post-decision value function can be easily constructed by solving the dynamic program using backward recursion.

3.2. Approximate Dual Dynamic Programming

Although the number of supporting hyperplanes of $V'_t(S_t, \cdot)$ is finite, computing all hyperplanes is prohibitive for larger problems. Like SDDP, the ADDP algorithm therefore uses Monte Carlo simulation to define a set of sample reservoir states, thereby finding those hyperplanes that are necessary to obtain an optimal decision policy.

The ADDP algorithm is outlined in Figure 2. The algorithm is initialized with an environmental state S_1, a reservoir state R_0, initial value functions $\hat{\bar{V}}'_t$, and the sets $\mathcal{M}_t = \emptyset$, $t = 1, \ldots, T$. Over N iterations, ADDP alternates between a forward and a backward pass. During the forward pass, the algorithm generates new states by sampling the state transition function, S^M. For a sampled state,

---

Input arguments: initial states S_1 and R_0, initial value function $\left(\hat{\hat{V}}_t'\right)_{t=1}^T$

Do for $n = 1, 2, \ldots, N$

   Forward Pass

  (1) Do for $t = 1, 2, \ldots, T-1$

      (1.1) Sample p_s from $P_t | S_t$

      (1.2) Solve $\hat{R}_{nt} \leftarrow \arg\ \max_{\pi_t} \left\{ C(S_t, R_{t-1}, \pi_t) + \gamma \hat{\hat{V}}_t'(S_t, R_t(\pi_t)) \right\}$ for the single scenario p_s

      (1.3) Sample $S_{t+1} \leftarrow S^M(S_t)$

   Backward Pass

  (2) Do for $t = T, T-1, \ldots, 2$

      (2.1) Do for all $S_t \in \mathcal{S}_t$

          (2.1.1) Do for $m \in \mathcal{M}_t \cup \{n\}$

             (2.1.1.1) Get hyperplane $(a_m(S_t), b_m(S_t)) \leftarrow H_{mt}(S_t, R; \hat{R}_{mt-1}) \in \partial_R V_t'(S_t, \hat{R}_{mt-1})$

      (2.2) If $\exists\ S_t \in \mathcal{S}_t : |\hat{V}_t'(S_t, \hat{R}_{nt-1}) - V_t'(S_t, \hat{R}_{nt-1})| > \varepsilon$ then $\mathcal{M}_t \leftarrow \mathcal{M}_t \cup \{n\}$

      (2.3) Do for all $S_{t-1} \in \mathcal{S}_{t-1}$

          (2.3.1) $\hat{\hat{V}}_{t-1}'(S_{t-1}, R) \leftarrow \min \left\{ \sum_{S_t \in \mathcal{S}_t} P(S_t | S_{t-1}) \left(a_m(S_t) + b_m(S_t)^\top (R - \hat{R}_{mt}) \right), m \in \mathcal{M}_t \right\}$

Return post-decision value functions $\hat{\hat{V}}_t'$ $(t = 1, \ldots, T-1)$

Figure 2   Approximate dual dynamic programming for Markov decision processes

the algorithm solves the relaxed version of problem (13) for a single (random) price scenario using the current approximation of the value function, i.e., maximizing $C(S_t, R_{t-1}, \pi_t) + \gamma \hat{\hat{V}}_t'(S_t, R_t(\pi_t))$, and then stores the final reservoir state that is a subset of the solution to the linear program (Step 1.2). During the backward pass, in each stage, the algorithm loops over all environmental states and previously stored reservoir states and computes the supporting hyperplanes (Step 2.1). For each predecessor state, we compute the weighted sum of all hyperplanes over all successor states and then update the approximation of the post-decision function (Step 2.3).

In conventional SDDP, the size of the set of sample reservoir states increases by one during each iteration of the outer loop. Some hyperplanes around the set of reservoir states, however, may be redundant or at least *similar* to existing hyperplanes. For ADDP, we therefore propose that hyperplanes which do not improve the approximation quality by more than ε should be omitted (Step 2.2). Denote $\hat{V}'$ as the approximate (pre-decision) value function constructed from a set of hyperplanes H_{mt}, with $m \in \mathcal{M}_t$, such that

$$\hat{V}_t'(S_t, R) = \min\left\{\sum_{S_t \in \mathcal{S}_t} \left(a_m(S_t) + b_m(S_t)^\top (R - \hat{R}_{mt})\right), m \in \mathcal{M}_t\right\}. \tag{17}$$

A new hyperplane H_{nt} is added to $\mathcal{M}_t$ if

$$\exists\, S_t \in \mathcal{S}_t : \hat{V}_t'(S_t, \hat{R}_{nt-1}) - V_t'(S_t, \hat{R}_{nt-1}) > \varepsilon. \tag{18}$$

In this way, ADDP converges to an upper bound of the solution to the relaxed problem, since the approximate value function in general remains an approximation and never converges to the true value function. Note that we also obtain an upper bound if we stop ADDP before an optimal solution is found, as this is often done in the literature, e.g., Flach et al. (2010), Philpott and de Matos (2012). A practical advantage of using $\varepsilon > 0$ instead of $\varepsilon = 0$ is that omitting new hyperplanes accelerates computation of the outer loop, which allows a larger number of states to be sampled in the same amount of time.

Existing convergence results for SDDP algorithms require that randomness is stagewise independent and enters only the right-hand side of the constraints of the linear program at each stage (Philpott and Guan 2008, Shapiro 2011). Both assumptions are necessary if the linear program is only being solved for a subset of scenarios during the backward pass. Right-hand side randomness guarantees that the optimal dual solutions for scenarios in the subset are also dual feasible for all other scenarios, which significantly accelerates the generation of new hyperplanes. Stagewise independence, in turn, enables sharing hyperplanes among different scenarios at the previous stage, since the post-decision value function is identical for all scenarios. Algorithms that exploit these properties can be found in Higle and Sen (1991) and Chen and Powell (1999). Although in our model these assumptions are not fulfilled, the algorithm still converges almost surely. First, dual solutions are always feasible because the linear program is being solved for the entire set of scenarios during the backward pass. Second, hyperplanes are not shared among scenarios, since we can construct a separate post-decision value function for each scenario by using the transition matrix of the Markov process.

PROPOSITION 3. *Denote π^ε as the policy obtained by ADDP for $\varepsilon > 0$ and π^* as the optimal policy of the relaxed problem. For a given initial reservoir level R_0, the policies obtained by ADDP for $\varepsilon = 0$ converge to the optimal policy in a finite number of steps. The values obtained from following π^ϵ are at most $\varepsilon(T-1)$ worse than the optimal values.*

Proof. We first consider the case $\varepsilon = 0$. It follows from Lemma 1 in Philpott and Guan (2008), that for fixed S_t the functions

$$R \mapsto V_t(S_t, R) \tag{19}$$

are the pointwise maxima of finitely many linear functions, i.e., are piecewise linear for all $1 \leq t \leq T$. Note that each possible sequence of states $(S_1, \ldots, S_T)$ has positive probability and therefore by the Borel-Cantelli Lemma occurs infinitely often in the forward pass. Since we add a hyperplane in each iteration, the finiteness of the set of hyperplanes implies that there exists an $\bar{n} \in \mathbb{N}$ such that no further hyperplanes are added after iteration $\bar{n}$. We denote the approximate value function after that state by $\hat{V}_t'$ for $1 \leq t \leq T$.

Suppose that the policy $\hat{\pi}$ found by using $(\hat{V}_t')_{1 \leq t \leq T}$ is suboptimal in period $T-1$ and some $(S_1, \ldots, S_T)$, i.e., $\max_\pi C(S_T, R_{T-1}, \pi) + \bar{V}_T(S_T, R_T) < \hat{V}_T'(S_T, R_{T-1})$. Since $(S_1, \ldots, S_T)$ is sampled in iterations $n > \bar{n}$, the value function approximation would be updated in these iterations – a contradiction to the choice of $\bar{n}$. Hence, $\hat{V}_T'(S_T, R_{T-1})$ coincides with V_T' for all R that can be reached by $\hat{\pi}$ and for all $S_T \in \mathcal{S}_T$. The same holds for the post-decision value function $\hat{\bar{V}}_{T-1}'$. Having established the accuracy of $\bar{V}_{T-1}'$, we can inductively show the accuracy of all $\hat{V}_t'$ and $\hat{\bar{V}}_t'$ for all S_t and t. Hence, the solutions obtained with $\hat{V}_t'$ coincide with the optimal solutions of the relaxed problem.

The finite convergence property carries over to the case $\varepsilon > 0$. To prove the second part of the proposition, we begin with the last period T and note that by definition

$$\hat{V}_T'(S_T, R_{T-1}) - V_T'(S_T, R_{T-1}) \leq \varepsilon, \quad \forall\, S_T \in \mathcal{S}_T, \tag{20}$$

for all states of the system R_T that can be reached from R_0 by following π^ε. This inequality also holds for the respective post-decision value functions.

Let $\Delta \equiv C(S_{T-1}, R_{T-2}, \pi^*) - C(S_{T-1}, R_{T-2}, \pi^\varepsilon)$. Since π^ε is optimal for $\hat{V}_T'$ it follows that

$$\hat{\bar{V}}_{T-1}'(S_{T-1}, R_{T-1}, \pi^\varepsilon) \geq \hat{\bar{V}}_{T-1}'(S_{T-1}, R_{T-1}, \pi^*) + \Delta. \tag{21}$$

Therefore for all $S_{T-1} \in \mathcal{S}_{T-1}$,

$$0 \leq C(S_{T-1}, R_{T-2}, \pi^*) + \bar{V}_{T-1}'(S_{T-1}, R_{T-1}(\pi^*))$$

$$- C(S_{T-1}, R_{T-2}, \pi^\varepsilon) - \bar{V}_{T-1}'(S_{T-1}, R_{T-1}(\pi^\varepsilon)) \tag{22}$$

$$= \Delta + \bar{V}_{T-1}'(S_{T-1}, R_{T-1}(\pi^*)) - \bar{V}_{T-1}'(S_{T-1}, R_{T-1}(\pi^\varepsilon)) \tag{23}$$

$$\leq \Delta + \bar{V}_{T-1}'(S_{T-1}, R_{T-1}(\pi^*)) - \hat{\bar{V}}_{T-1}'(S_{T-1}, R_{T-1}(\pi^\varepsilon)) + \varepsilon \tag{24}$$

$$\leq \Delta + \hat{\bar{V}}_{T-1}'(S_{T-1}, R_{T-1}(\pi^*)) - \hat{\bar{V}}_{T-1}'(S_{T-1}, R_{T-1}(\pi^\varepsilon)) + \varepsilon \tag{25}$$

$$\leq \varepsilon, \tag{26}$$

where (24) follows from (21) and (25) from $\bar{V}_{T-1}' \leq \hat{\bar{V}}_{T-1}'$. Since $\hat{V}_{T-1}'$ is an ε-approximation of the function

$$R \mapsto \max_{\pi^\varepsilon} C(S_{T-1}, R, \pi^\varepsilon) + \hat{\bar{V}}_{T-1}'(S_{T-1}, R_{T-1}(\pi^\varepsilon)), \tag{27}$$

we have $\hat{V}_{T-1}'(S_{T-1}, R_{T-2}) \leq 2\varepsilon + V_{T-1}'(S_{T-1}, R_{T-2})$. Since the above holds for all $S_{T-1} \in \mathcal{S}_{T-1}$, the property carries over to the post-decision value function $\bar{V}_{T-2}'$ and the error bound follows by induction. $\square$

4. Econometric Model

Consistent with our model formulation, we propose an econometric model that separates randomness into a process of environmental variables with daily time increments and a process of electricity prices with hourly time increments. The objective of the econometric model is to accurately describe the dynamics of electricity prices and natural inflows by a small number of explanatory variables that fit into this modeling framework.

As with every commodity, the price of electricity is determined by supply and demand. In the short term, supply is primarily driven by seasonal variations of intermittent power sources, such as wind, solar, and run-of-river, or by power plant outages. In particular, wind and solar power

production drives down the electricity price, since more expensive technologies are forced out of the market. In the medium and long term, it is mainly the prices for primary energy such as coal and gas that influence the price for electricity. Electricity demand, on the other side, can be largely explained by temperatures and deterministic seasonal factors. The temperature affects electricity prices due to higher demand for heating and cooling. To verify this relationship, we ran a linear regression of the mean demand for electricity per day in Austria and Germany on the mean day temperature, the squared mean temperature, the day length (i.e., the time from sunrise to sunset), as well as dummy variables for national holidays in Germany and Austria. Based on a sample of 730 observations from 2010 to 2011, the model explains 76% of the variance in electricity demand. Accordingly, we model electricity prices dependent on those variables that influence supply and demand.

To meet the requirements of a finite-horizon Markov decision process, we decompose the dynamics of the environmental variables into a time-dependent trend and a state transition process which has the Markov property. The state of the Markov process on day t is defined by the weekday (DAY), the mean day temperature (TEMP), the total wind power generation during that day (WP), the total solar power generation (SP), the natural inflow (IN), and the gas price (GAS)

$$S_t = (\text{DAY}_t, \text{TEMP}_t, \text{WP}_t, \text{SP}_t, \text{IN}_t, \text{GAS}_t). \tag{28}$$

For a given realization of the state and a given day of the year, we can then model the hourly conditional expectations of the electricity prices, $\mathbb{E}(p_1, \ldots, p_{24}|S_t)$.

For model estimation, we used hourly day-ahead and intraday spot prices from 2009 to 2011 as published by EPEX SPOT. Hourly data on wind and solar power forecasts are published by E.ON, EnBW, RWE, and Vattenfall for the four major German transmission zones. We used forecasts instead of realized generation because spot prices are fixed one day in advance so that forecasts have a greater explanatory power than actual generation data. Since the price effect of temperature is a function of population density and local temperatures, we define the mean day temperature as a population weighted index over all Austrian and German cities with a population of more

than 10,000. The index has been calculated using Mathematica 7 *CityData* and *WeatherData*.
Data on natural inflows have been provided by an Austrian generating company for a system of
seven interconnected hydro plants in the Alps (see Section 5 for further details). As the inflow
patterns exhibit pairwise correlations of around $\rho = 0.8$, we aggregate these inflows into a single
state variable by summing up the inflows for all reservoirs. To estimate the dynamics of the gas
price, we use daily closing prices from NetConnect Germany.

4.1. State Transition Model

To separate stationary from non-stationary state variables, we divide the state S_t into two separate
sub states: one state that only contains the gas price S_t^1 and another state that contains all other
state variables S_t^2. The gas price is modeled as non-stationary geometric Brownian motion (GBM),
representing the long term market trends, while the other state variables follow a stationary Markov
process, capturing short term variations in the electricity price.

To describe the dynamics of the gas price S_t^1, we fit a GBM to the data and discretize the resulting
log-normal price distributions for every day t of the planning horizon. Specifically, we choose
gas price states and probabilities such that the Kantorovich distances between the corresponding
discrete distributions and the log-normal distributions are minimal (Graf and Luschgy 2000). The
Kantorovich distance is suitable for the use in stochastic programming and, in a certain sense,
ensures an optimal discretization (Graf and Luschgy 2000, Pflug 2001). The number of gas price
states is chosen such that the Kantorovich distance does not exceed 0.5 with a cap of 30 gas
price states per day. The transition probabilities between nodes in consecutive stages can easily be
computed using the conditional distributions resulting from the specification of the GBM. In this
way we obtain a probability lattice, representing the stochastic evolution of the gas price.

The other state $(S_t^2)_{1 \leq t \leq T}$ is decomposed into a deterministic trend component $(D_t)_{1 \leq t \leq T}$ and a
random error $(E_t)_{1 \leq t \leq T}$ which follows a time-homogeneous Markov chain,

$$S_t^2 = D_t + E_t, \quad t = 1, \ldots, T \tag{29}$$

For the TEMP state variable, we used the temperature index for the years 2009 to 2011 and estimated the parameters δ, α and u of the trigonometric regression model,

$$\text{TEMP}_t = \delta + \alpha \sin\left(\frac{365 - t}{2\pi} - u\right) + \epsilon_{t1}, \tag{30}$$

which yields an $R^2=80.09\%$. The total wind power generation per day exhibits annual seasonality because of higher wind speeds during winter as well as an upward trend over time due to the ongoing installation of new wind power units. To capture these two trends, we include a quadratic term as well as an interaction term in the trigonometric regression model of wind power production,

$$\text{WP}_t = \delta_1 + \delta_2 t + \delta_3 t^2 + \alpha_1 \sin\left(\frac{365 - t}{2\pi} - u_1\right) + \alpha_2 \sin\left(\frac{365 - t}{2\pi} - u_2\right) t + \epsilon_{t2}, \tag{31}$$

with an $R^2=9.46\%$. Production of solar power exhibits a strong seasonal component as well as a trend in time and is modeled in an analogous way with residuals ϵ_{t3} ($R^2=71.67\%$). We do not model inflows using trigonometric models, since the observed inflow peaks in spring as well as the long dry period in winter cannot be captured by a sine function. Instead, based on 18 years of daily inflow data, we estimated the trend w_t in natural inflows for each day of the year by a Nadaraja-Watson non-parametric regression ($R^2=73.65\%$). The inflow model is given by

$$\text{IN}_t = w_t \epsilon_{t4}. \tag{32}$$

Residuals ϵ_{t4} are obtained by dividing the inflow realizations by their respective estimated means.

To estimate a model of E_t, we used the detrended state variables $\hat{e}_t = (\hat{\epsilon}_{t1}, \hat{\epsilon}_{t2}, \hat{\epsilon}_{t3}, \hat{\epsilon}_{t4})$, i.e., the residuals from (30) to (32) for 2009 to 2011. All residuals show a strong autocorrelation supporting the hypothesis of stagewise dependence. By modeling the transition from one state to another as a Markov process, we capture autocorrelation up to the first lag. Note that modeling inflow randomness through a Markov chain of geometric errors does not fully capture long-term variability, so that the coefficient of variation of total inflows over one year is 5.2% in the model compared to 7.7% in the data. To ensure parsimony of the model, however, we do not include higher order lags.

We estimated the transition probabilities of the Markov chain for E_t by first fixing a number of states M and applying k-means clustering to organize the observations $(\hat{e}_t)_{t=1}^T$ into M clusters,

$Z_1, \ldots, Z_M$. In a second step, we estimated the transition probabilities by counting the number of transitions between clusters as they occur in the sample. Accordingly, the transition probability matrix is given by

$$\mathbb{P}(E_{t+1} = Z_j | E_t = Z_i) = \frac{|\{t : \hat{e}_{t+1} \in Z_j, \hat{e}_t \in Z_i\}|}{|\hat{e}_t \in Z_i|}, \quad \forall \, i, j. \tag{33}$$

For our implementation, we chose $M = 30$ to obtain the cluster centers $Z_1, \ldots, Z_{30}$ and end up with partitions where the smallest of the clusters contains 14 of the original data points, while the largest cluster represents 143 observations.

To obtain a joint probability lattice, we formed the product of the gas price lattice S_t^1 and the Markov chain S_t^2, under the assumption that both processes are independent.

4.2. State-Dependent Price Models

Day-ahead prices are represented by linear models. We estimated one model for every hour and distinguish between working days and weekends, i.e., a total of 48 models. The regressors consist of all state variables, daily demand for electricity, and the day length in minutes. We included all interactions of the regressors up to the second order. Note that, we do not transform the price data, as is done for example in semilog models, since our sample contains negative prices, which cannot be handled in these settings. Furthermore, the corresponding reverse transformations of such models would have introduced instability in the simulation.

In order to ensure parsimony of the model, we performed stepwise combined forward-backward elimination as described in Draper and Smith (1998), Section 15.2, and used the Bayesian information criterion (BIC) for model selection. For the day-ahead price models, the selection routine chooses 11.67 regressors on average, but at most 20 out of 45 regressors. The number of regressors is reasonable, considering that we used around 312 observations for the weekend models and 782 observations for the working day models. The overall in-sample fit of the linear models for the day-ahead prices is R²=65.81%, which is satisfactory, considering the varying economic conditions as well as structural changes on the market for electricity. Although autocorrelation is still present in

the residuals of the linear models, due to long term trends in the power prices not captured by our model, we do not consider these influences to circumvent a further increase of the dimensionality of the MDP.

To estimate the price response on the intraday market, we regressed the difference of intraday and day-ahead price on the hourly demand for electricity. The linear regression yielded an intercept of -20.3 (€) and a slope coefficient of 0.0011 (€ per megawatt hour), with an in-sample fit of R^2=45.07%. The negative intercept reflects the fact that the true price response function is non-linear. Based on the data, we set the slope of the price-response function in the objective function of the stochastic program to $\beta = 0.0011$.

4.3. Simulation

To simulate price trajectories over one year, we began by sampling a state from the steady-state distribution of the Markov chain. The state transition process is simulated using the probabilities in (33). To obtain the inflow for a single reservoir, we multiply the sampled ϵ_{t4} with the average inflows of the respective reservoir on the chosen day. Based on the realization of the state variable, and the day of the year, we first simulated a demand for electricity using the linear model for demands described above and an error term sampled from a normal distribution fitted to the residuals of that model. In a second step, using the demand and the state variables of the MDP, we generated hourly day-ahead prices using the linear models for electricity prices. Random noise was added by sampling the error term of location scale t-distributions fitted to the residuals of the linear models. The approach is supported by the Kolomogorov-Smirnov goodness of fit test, which does not reject the null hypothesis of a t-distribution in any of the linear models ($\alpha = 0.05$). Using the t-distribution yields heavy-tailed prices, as they are often observed in electricity markets.

To generate day-ahead price scenarios for the stochastic program, we resorted to *Latin Hypercube* sampling (LHS) as a variance reduction technique (Shapiro 2003). Denote $F_h^{-1}(S_t, U_l)$ as the inverse CDF (t-distribution) of the day-ahead price during hour h for a given S_t, and denote U_l as a uniform random variable. Then, we can generate K day-ahead prices using

Reservoir		j	1	2	3	4	5	6	7
Max reservoir content	1000m^3	UB$_j^R$	84 941	83 000	30	10	2000	120	95
Initial reservoir content	1000m^3	R_{0j}	34 333	37 000	15	5	1000	60	47
Average hourly inflow	1000m^3	$\bar{w}_j$	23.0	6.9	134.9			45.2	66.8
Max pumping capacity	MW	UB$_j^-$	600						
Min pumping capacity	MW	LB$_j^-$	20						
Max generating capacity	MW	UB$_j^+$	592	220	0.5	120	9	16	16.4
Min generating capacity	MW	LB$_j^+$	15	10					
Pump efficiency	MW/1000m^3	η_j^-	4.23						
Turbine efficiency	MW/1000m^3	η_j^+	3.17	7.51	0.006	1.32	0.06	0.27	0.10

Table 1   System specifications

$$p_{sh} = p'_{\Theta(s)h}, \quad p'_{lh} = F_h^{-1}(S_t, U_l) \quad \text{with} \quad U_l \sim U[(l-1)/K, l/K] \quad \forall \quad l \in \{1, \ldots, K\}, \quad (34)$$

where Θ is defined as a mapping from s to l such that p_h is a random permutation of p'_h, i.e., we shuffled the price scenarios.

5. Results

To test the efficiency of the proposed algorithm, we conducted a numerical analysis based on data from a generating company in Austria. The company operates a large hydro storage system in the Austrian Alps, which consists of an upper ($j = 1$) and a lower reservoir ($j = 2$). Both reservoirs are fed by natural inflows of two glacier rivers, and water can be pumped from the lower into the upper reservoir. In 2011, the system received a capacity upgrade which increased the pumping and generating capacities at the upper reservoir by a factor of five. Another capacity upgrade by the same amount is planned for 2016. All releases from the lower reservoir flow into a hydro cascade which consists of several small hydro plants that line up along downstream rivers. System specifications of all plants are given in Table 1.

Based on a default model configuration, consisting of reservoirs $\mathscr{J} = \{1, 2\}$ at their current stage of expansion, we investigated the influence of changing selected parameters ceteris paribus on the performance of the algorithm as well as the behavior of the optimal policy. As variations, we considered a version of the model with its former capacity until the 2011 expansion (small capacity) as well as a version of the model with its future capacity after a possible expansion in 2016 (large capacity). To study the effect of a larger number of reservoirs, we also included five downstream

Model	$\mathscr{J}$	ε	UB_1^+	LB_1^+	UB_1^-	LB_1^-
Default	$\{1,2\}$	10^4	592	15	600	20
Small Capacity	$\{1,2\}$	10^4	112	15	120	20
Large Capacity	$\{1,2\}$	10^4	1072	15	1080	20
Reservoir Chain	$\{1,2,3,4,5,6,7\}$	10^4	592	15	600	20
50% Min Capacity	$\{1,2\}$	10^4	592	296	600	300
Epsilon Zero	$\{1,2\}$	0	592	15	600	20

Table 2   Model configurations used for the numerical analyses

hydro plants (reservoir chain). Moreover, to study the effect of more difficult integer requirements on approximation quality, we increased the minimum capacity of turbines and pumps to 50% of maximum capacity (50% Min Capacity). Finally, since we used a default $\varepsilon = 10^4$ for all models, we studied the effect of setting $\varepsilon = 0$ (Epsilon Zero). All model configurations are summarized in Table 2.

5.1. Implementation

The algorithm and the electricity price model were implemented in Java. The linear relaxation used with ADDP was modeled using the Java API of Google's OR-Tools and solved using Sulum Optimization's linear programming solver. The stochastic quadratic mixed-integer program used for the simulations was modeled and solved using the Java API of the Xpress Optimization Suite. All computations were executed on Amazon EC2 'm2.4xlarge' instances which correspond to Intel Xeons E5-2665 with 8 cores at 2.4Ghz with 68G memory. The implementation of the algorithm makes heavy use of multithreading which led to a linear speed up in the number of cores.

For our numerical analyses, we generated states and electricity price scenarios using the econometric model described in Section 4. The full problem formulation has $T = 365$ stages, $|\mathcal{S}| = 282,211$ states, and $44,765,192$ transition probabilities. Moreover, for each state, we generated $K = 20$ scenario paths, each containing $|\mathscr{H}| = 24$ price realizations. An equivalent scenario tree that reproduces all possible price paths would require $4.3 \cdot 10^{1234}$ terminal nodes.

The discount factor was set to $\gamma = 1.0$ and all bidding curves had four segments with $I = 3$ breakpoints. Initial experiments showed that using a larger number of breakpoints or a larger number of scenarios did not significantly change the objective value.

Model	Iter	Exp Profit	Sim Mean (SE)	RH (SE)	Gap	CV	Time	Hyp Count
Default	13	164.3	162.4 (2.1)	154.8 (1.7)	0.012	0.21	6.9	2006569
Small Capacity	11	134.0	132.2 (1.8)	131.3 (1.6)	0.013	0.22	4.6	1509970
Large Capacity	17	188.3	188.2 (3.1)	173.5 (2.0)	0.000	0.26	11.9	2512240
Reservoir Chain	17	229.2	229.2 (3.2)	219.3 (2.6)	-0.002	0.22	38.8	2244477
50% Min Capacity	13	164.7	163.4 (2.6)	154.5 (1.9)	0.008	0.25	6.8	1892744
Zero Epsilon	13	164.3	162.6 (2.2)	154.2 (1.7)	0.012	0.21	7.8	3498359

RH = rolling horizon benchmark, SE = standard error of the mean, CV = coefficient of variation of the simulated profit, Iter = iterations to convergence, Time = computing hours to convergence, Hyp Count = final number of hyperplanes

Table 3   Summary of the results for different problem configurations

5.2. Computational Performance

To test the convergence of the algorithm, we ran ADDP for 25 iterations with all six configurations and compared expected first-stage profits with simulated profits. To avoid reservoirs ending up being emptied at the end of the year, the value function of the final stage $\bar{V}_T$ is such that it sufficiently penalizes any reservoir content below the initial reservoir levels. To obtain the expected profits, after each iteration, we computed the objective values of the stochastic program at $T = 1$ using the most recent approximate post-decision value function. To assess the actual value of the current approximation, we simulated the decision process by solving the non-relaxed version of the stochastic program 1000 times for different scenario paths over the whole planning horizon of one year.

Furthermore, as a benchmark, we formulated the deterministic counterpart of the relaxed problem as one large linear program. We then simulated the planning process that results from using this model on a rolling horizon. The model takes the current point estimate of future prices and inflows to make decisions and then evaluates these decisions using the actual price and inflow realizations.

Table 3 summarizes the results of the numerical study. We recorded the iteration in which the upper bound of the 95% confidence interval of the simulated profit exceeds the expected profit found by ADDP. Note that the true profit lies in between the expected profit and the lower bound of the 95% confidence interval of the simulated profit with 97.5% probability (Shapiro 2011).

The figures in all other columns refer to the corresponding recorded iteration. The third and

fourth columns show the expected and the simulated profits (in €1.0M). Column five summarizes the rolling horizon solution. Column six shows the gap, i.e. one minus the ratio of expected to simulated profit. Column seven shows the coefficient of variation of the simulated profit. Column eight shows the computational time in hours until convergence. The last column shows the total number of hyperplanes that were used to construct the post-decision value function.

The small gap between expected and simulated profits observed for all problem instances demonstrates that the relaxed problem is sufficient to find a tight value function approximation for the original two-stage mixed integer problem.

We find that the convergence behaviour is largely affected by the size of the system, both in terms of storage power capacity, i.e., the capacity of turbines and pumps, as well as the number of storage units. Increasing the number of reservoirs from 2 to 7 increases the computational time due to the larger complexity of the optimization problem. However, the increase in the number of iterations as well as the number of hyperplanes does not reflect the magnitude of the increase in problem size. While a larger number of reservoirs has a direct effect on computational times, it apparently has only a minor effect on the number of required hyperplanes and thereby on the required number of iterations. Problems that require more hyperplanes seem to be those where the optimal value function reaches a larger number of sufficiently different reservoir states, which is the case when the power capacity of the storage plant is relatively large.

With respect to the rolling horizon benchmark, the value of following the optimal policy also largely depends on the power capacity of the plant. While the gap between rolling horizon and optimal policy is low for the small capacity system (+0.7%), it is much higher for the large capacity system (+8.5%). This indicates that a flexible storage system that possesses the ability to quickly change the reservoir content benefits more from a stochastic solution than less flexible systems.

The convergence of expected and simulated profits for $\varepsilon = 0$ and $\varepsilon = 10^4$ is shown in Figure 3 along with the final number of hyperplanes. Despite using the relaxed problem to approximate the value function, the gap between expected and simulated profits is less or equal to 1.2% (even if the minimum capacity is at 50% of the maximum capacity as seen in Table 3). We can see from the
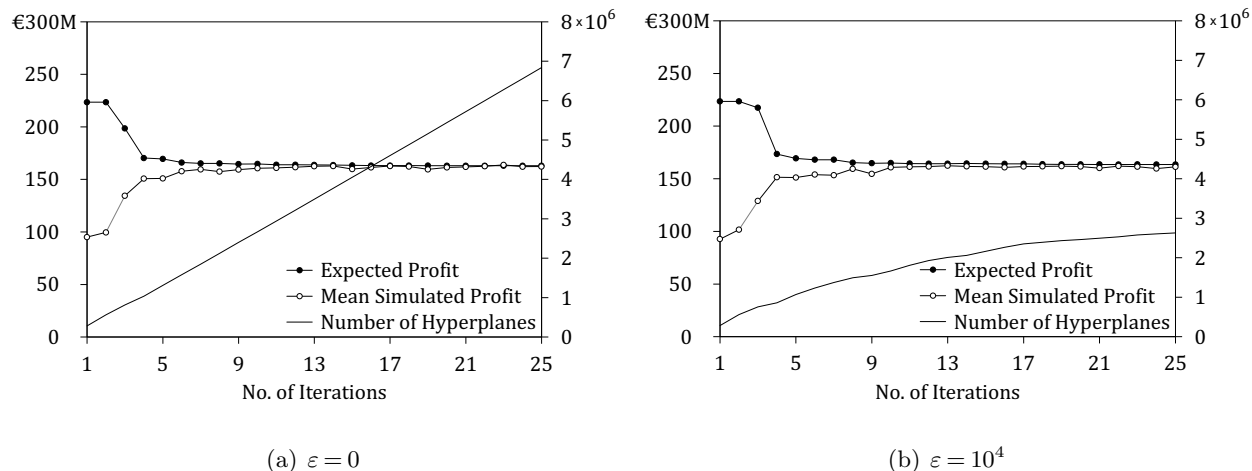
(a) $\varepsilon = 0$ (b) $\varepsilon = 10^4$

Figure 3   Convergence of ADDP for different approximation bounds for the case 50% Min Capacity.

plot that the gap closes after 25 iterations if $\varepsilon = 0$ but not if $\varepsilon = 10^4$. However, using a lower ε is paid for by a higher computational burden, which is reflected in the number of hyperplanes needed to approximate the value function. Setting $\varepsilon = 10^4$ requires solution of $3.7 \cdot 10^8$ linear programs until convergence is reached (as opposed to $4.9 \cdot 10^8$ linear programs if $\varepsilon = 0$), which reduces the computational time by about 12%. Note that we deliberately used a large default ε to illustrate this aspect.

5.3. Structural Insights

So far, the analyses showed that the optimal policy as well as the added value of following it is highly sensitive towards the power capacity of the storage system, i.e., the size of the turbines and pumps relative to the size of the reservoir. For a constant reservoir size, a higher power capacity decreases the time that the storage plant can run at full capacity. For example, if we ignore inflows, it would take 134 days to empty the upper reservoir with 112 MW turbine capacity (small), but only 14 days with 1,073 MW turbine capacity (large). This has a tremendous effect on which reservoir states could possibly be reached by the optimal policy.

Figure 4 shows the reservoir contents of the upper reservoir over the course of the year for the small and the large setup. While the variation in the reservoir content curves over the year is low for the small capacity system, the variation is high for the high capacity system. Due to the large
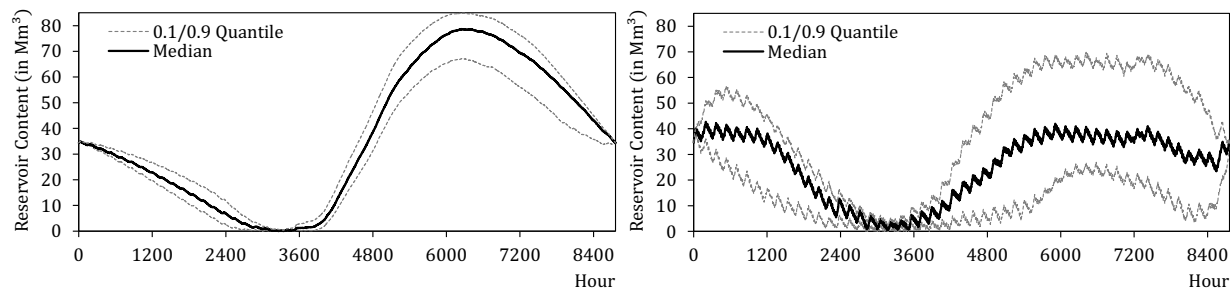
(a) with former capacity until the 2011 expansion   (b) with future capacity after the 2016 expansion

Figure 4   Reservoir content curves

amount of inflow arriving during the summer months, the small system pumps 5.0% of the time in winter and 20.0% of the time in summer on average. The large system, on the other hand, pumps 30.0% of the time regardless of the season. This operational flexibility in addition to the lower number of days at which the plant can run at full capacity leads to a higher variability in reservoir contents and thereby increases the value of a stochastic solution.

An analysis of the simulated decisions revealed that the plants operate in three states 99% of the time. The plants either turbinate at full capacity, or they pump at full capacity, or neither. Such a decision policy seems to be sufficiently represented by bidding curves with only four segments, which is also reflected by the small gap between expected and simulated profits.

6. Conclusion

We modeled the bidding problem of a generating company that operates a network of hydro storage plants as a multi-stage stochastic program and proposed a solution strategy that integrates stochastic dual dynamic programming with ideas from approximate dynamic programming. We divided the annual planning horizon into daily stages with hourly bidding decisions as part of the intrastage bidding problem. Accordingly, we separated intrastage from interstage randomness, which enabled us to model price uncertainty at each stage dependent on a state variable that evolves over time following a Markov process. To solve the multi-stage decision problem, we proposed a solution strategy that computes an approximation of the value function of the interstage process. The algorithm, referred to as approximate dual dynamic programming (ADDP), uses a probability

lattice to represent the Markov process and iteratively constructs a polyhedral approximation of the value function. This approximation can be used inside the original, more complicated intraday bidding problem to derive near-optimal bidding decisions.

We showed that the algorithm converges and derived an error bound of the polyhedral approximation. Tailored to the modeling framework, we developed an econometric model of electricity prices and stochastic inflows fitted to data from the EPEX SPOT wholesale electricity market as well as actual inflow data. We then carried out a case study based on different configurations of a hydro storage system in Austria. We find that approximating the continuous Markov process by a discrete probability lattice provides a good model fit. Numerical results indicate that the algorithm converges to a near-optimal solution, despite using a relaxed version of the original problem to approximate the value function of the interstage problem. Furthermore, we find that computational complexity as well as the value of the stochastic solution depends on the ratio of reservoir size to installed power capacity.

Future work should focus on models that additionally consider the market for reserve electricity as well as the market for future contracts. It would also be interesting to test the approach on the unit commitment problem of a thermal power system.

Acknowledgements

Biographies

Nils Löhndorf is an assistant professor at the Vienna University of Economics and Business, Vienna, Austria. His research interests are in stochastic optimization, energy systems planning, machine learning, and production planning.

David Wozabal is an assistant professor at the Technische Universität München, Munich, Germany. His research interests are in robust and stochastic optimization, risk management, and energy markets modeling.

Stefan Minner is a full professor at the Technische Universität München, Munich, Germany. His research interests are in logistics network design, supply chain coordination, and inventory management.

References

Auer, J. 2011. Hydropower in Europe. Market Study. Deutsche Bank Research.

Baíllo, A., M. Ventosa, M. Rivier, A. Ramos. 2004. Optimal offering strategies for generation companies operating in electricity spot markets. *IEEE Transactions on Power Systems* **19**(2) 745–753.

Cerisola, S., Á. Baíllo, J. M. Fernández-López, A. Ramos, R. Gollmer. 2009. Stochastic power generation unit commitment in electricity markets: A novel formulation and a comparison of solution methods. *Operations Research* **57**(1) 32–46.

Chen, Z. L., W. B. Powell. 1999. Convergent cutting-plane and partial-sampling algorithm for multistage stochastic linear programs with recourse. *J. Optim. Theory Appl.* **102** 497–524.

Draper, N.R., H. Smith. 1998. *Applied regression analysis*. Wiley.

Eichhorn, A., H. Heitsch, W. Römisch. 2009. Scenario tree approximation and risk aversion strategies for stochastic optimization of electricity production and trading. *Optimization in the Energy Industry*. Springer, 321–346.

Flach, B.C., L.A. Barroso, M.V.F. Pereira. 2010. Long-term optimal allocation of hydro generation for a price-maker company in a competitive market: latest developments and a stochastic dual dynamic programming approach. *IET Generation, Transmission and Distribution* **4**(2) 299–314.

Fleten, S.-E., T. K. Kristoffersen. 2007. Stochastic programming for optimizing bidding strategies of a Nordic hydropower producer. *Eur J Oper Res* **181**(2) 916–928.

Fleten, S.-E., T. K. Kristoffersen. 2008. Short-term hydropower production planning by stochastic programming. *Computers and Operations Research* **35**(8) 2656–2671.

García-González, J., E. Parrilla, A. Mateo. 2007. Risk-averse profit-based optimal scheduling of a hydro-chain in the day-ahead electricity market. *Eur J Oper Res* **181**(3) 1354–1369.

Gjelsvik, A., B. Mo, A. Haugstad. 2010. Long- and medium-term operations planning and stochastic modelling in hydro-dominated power systems based on stochastic dual dynamic programming. *Handbook of Power Systems I*. Springer, 33–55.

Graf, S., H. Luschgy. 2000. *Foundations of Quantization for Probability Distributions*. Springer.

Heitsch, H., W. Römisch. 2003. Scenario reduction algorithms in stochastic programming. *Computational Optimization and Applications* **24** 187–206.

Higle, J. L., S. Sen. 1991. Stochastic decomposition: An algorithm for two-stage linear programs with recourse. *Mathematics of Operations Research* **16**(3) 650–669.

Hochreiter, R., D. Wozabal. 2010. A multi-stage stochastic programming model for managing risk-optimal electricity portfolios. *Handbook of Power Systems II*. Springer, 383–404.

Löhndorf, N., S. Minner. 2010. Optimal day-ahead trading and storage of renewable energies - an approximate dynamic programming approach. *Energy Systems* **1** 61–77.

Matevosyan, J., M. Olsson, L. Söder. 2009. Hydropower planning coordinated with wind power in areas with congestion problems for trading on the spot and the regulating market. *Electr Pow Syst Res* **79**(1) 39–48.

Pereira, M.V.F., L.M.V.G. Pinto. 1991. Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming* **52**(2) 359–375.

Pflug, G.Ch. 2001. Scenario tree generation for multiperiod financial optimization by optimal discretization. *Mathematical Programming, Series B* **89**(2) 251–271.

Philpott, A.B., V.L. de Matos. 2012. Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion. *European Journal of Operational Research* **218**(2) 470 – 483.

Philpott, A.B., Z. Guan. 2008. On the convergence of stochastic dual dynamic programming and related methods. *Operations Research Letters* **36**(4) 450–455.

Powell, W. B. 2011. *Approximate dynamic programming. Solving the curses of dimensionality*. Wiley.

Pritchard, G., A. B. Philpott, P. J. Neame. 2005. Hydroelectric reservoir optimization in a pool market. *Mathematical Programming* **103**(3) 445–461.

Rockafellar, R. Tyrrell, R. J.-B. Wets. 1998. *Variational analysis*, *Fundamental Principles of Mathematical Sciences*, vol. 317. Springer.

Shapiro, A. 2003. Monte Carlo sampling methods. *Stochastic programming*, *Handbooks in Operations Research and Management Science*, vol. 10. Elsevier, 353–425.

Shapiro, A. 2011. Analysis of stochastic dual dynamic programming method. *Eur J Oper Res* **209**(1) 63–72.

Zuber, M. 2011. The European market for pumped-storage power plants. Market Study. Ecoprog GmbH, Cologne.