

Weekly Two-Stage Robust Generation Scheduling for Hydrothermal Power Systems

Hossein Dashti, *Student Member, IEEE*, Antonio J. Conejo, *Fellow, IEEE*, Ruiwei Jiang, *Member, IEEE*, and Jianhui Wang, *Senior Member, IEEE*

Abstract—As compared to short-term forecasting (e.g., 1 day), it is often challenging to accurately forecast the volume of precipitation in a medium-term horizon (e.g., 1 week). As a result, fluctuations in water inflow can trigger generation shortage and electricity price spikes in a power system with major or predominant hydro resources. In this paper, we study a two-stage robust scheduling approach for a hydrothermal power system. We consider water inflow uncertainty and employ a vector autoregressive (VAR) model to represent its seasonality and accordingly construct an uncertainty set in the robust optimization approach. We design a Benders' decomposition algorithm to solve this problem. Results are presented for the proposed approach on a real-world case study.

Index Terms—Hydrothermal coordination, robust optimization, unit commitment, vector autoregressive model.

NOMENCLATURE

A. Sets and Indices

$\mathcal{B} = \{1, \dots, B\}$	Set of buses.
$\mathcal{I} = \{1, \dots, I\}$	Set of hydro and thermal generators.
$\mathcal{I}_H \subseteq \mathcal{I}$	Set of hydro generators.
$\mathcal{G}_b \subseteq \mathcal{I}$	Set of hydro/thermal generators at bus b .
$\mathcal{K} = \{1, \dots, K\}$	Set of hydro reservoirs.
$\mathcal{I}_k \subseteq \mathcal{I}$	Set of hydro generators at reservoir k .
\mathcal{L}	Set of transmission lines.
$\mathcal{T} = \{1, \dots, T\}$	Set of operating time intervals.

Manuscript received May 27, 2015; revised September 13, 2015 and November 07, 2015; accepted December 17, 2015. The work of H. Dashti was supported in part by the University of Arizona Renewable Energy Network. The work of A. J. Conejo was supported in part by the National Science Foundation (NSF) under grant 60050502. The work of R. Jiang was supported in part by the NSF under grant CMMI-1555983. The work of J. Wang was supported by the U.S. Department of Energy Office of Electricity Delivery and Energy Reliability. Paper no. TPWRS-00750-2015.

H. Dashti is with the Department of Systems and Industrial Engineering, University of Arizona, Tucson, AZ 85721 USA (e-mail: hdashti@email.arizona.edu).

A. J. Conejo is with the Department of Integrated Systems Engineering and the Department of Electrical & Computer Engineering, The Ohio State University, Columbus, OH 43210 USA (e-mail: conejonavarro.1@osu.edu).

R. Jiang is with the Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI, 48109 USA (e-mail: ruiwei@umich.edu).

J. Wang is with the Argonne National Laboratory, Lemont, IL 60439 USA (e-mail: jianhui.wang@anl.gov).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TPWRS.2015.2510628

\mathcal{U}_k	Set of immediate up-stream reservoirs of hydro reservoir k .
\mathcal{W}	Uncertainty set of the volume of natural water inflow.

B. Parameters

α_{ip}	Slope coefficients of linear piece p in approximating the performance curve of hydro reservoir i (MW/(m ³ /s)).
β_{ip}	Coefficients of linear piece p in approximating the performance curve of hydro reservoir i (MW).
c_{iq}^G	Unit cost of electricity generation from generator i in the q^{th} piece of its approximate fuel cost function (\$/MW).
d_{iq}^G	Constant term in the q^{th} piece of the approximate fuel cost function (\$).
C_{mn}	Capacity of the transmission line linking bus m and bus n (MW).
c_i^U/c_i^D	Unit cost of upward/downward reserve of generator i (\$/MW).
D_{bt}	Load at bus b in time period t (MW).
f_{kt}	Volume of natural water inflow into hydro reservoir k during time period t (Mm ³).
$h_k^{\text{MAX}}/h_k^{\text{MIN}}$	Upper/lower limit of the volume of water in hydro reservoir k (Mm ³).
K_{mn}^b	Line flow distribution factor for the transmission line linking bus m and bus n due to a net injection at bus b .
L_i	Minimum power output if generator i is on (MW).
MD_i	Minimum-down time for generator i (h).
MU_i	Minimum-up time for generator i (h).
NL_i	No-load cost for generator i . (\$)
P	Number of linear pieces used to approximate the performance curve.
RD_i	Ramp-down limit for generator i (MW).
RU_i	Ramp-up limit for generator i (MW).
SD_i	Shut-down cost for generator i (\$).
SU_i	Start-up cost for generator i (\$).
τ_{jk}	Time delay between hydro reservoir j and hydro reservoir k (h).

U_i	Maximum power output if generator i is on (MW).
$w_i^{\text{MAX}}/w_i^{\text{MIN}}$	Upper/lower limit of the volume of turbined water from hydro generator i (m^3/s).

C. Decision Variables

e_{it}	Scheduled amount of electricity for generator i in period t (MW).
g_{it}	Actual amount of electricity generated by generator i in period t (MW).
h_{kt}	Total volume of water in hydro reservoir k at the end of time period t (Mm^3).
$r_{it}^{\text{U}}/r_{it}^{\text{D}}$	Upward/downward reserve amount allocated to generator i in time period t (MW).
s_{kt}	Volume of water spilled from hydro reservoir k during time period t (m^3/s).
u_{it}	Binary decision variable to indicate if generator i is started up at the beginning of period t .
v_{it}	Binary decision variable to indicate if generator i is shut down at the beginning of period t .
w_{it}	Volume of turbined water from hydro reservoir i during time period t (m^3/s).
y_{it}	Binary decision variable to indicate if generator i is on in time period t .

I. INTRODUCTION

THIS paper considers the generation scheduling problem for a power system with predominant or major hydro resources (hereafter denoted as a hydrothermal power system). Hydrothermal power systems can be found worldwide, e.g., in Brazil [1], New Zealand [2], and Norway [3]. Unlike conventional energy resources (e.g., coal, oil, and gas), the generation capacity of a hydro plant depends on the amount of stored water in the reservoir and the random natural water inflow. Also, since the water reservoirs are often interconnected along a river basin, the generation capacity can considerably depend on the operations of the upstream hydro plants as well. Because of these distinctive features, the system operator of a hydrothermal power system needs to consider temporal and spacial coordinations of the hydro plant operations. On the one hand, it needs to ensure sufficient capacity for future generation needs. On the other hand, it needs to carefully balance the water stock and power generation among interconnected hydro plants to respect the corresponding operational restrictions and to minimize the operational costs. However, due to the random nature of water inflow, it is challenging for the system operator to accomplish these two tasks in an accurate manner. For example, without correctly predicting a low water inflow in one month, the system operator can deplete the water stock now and consequently resorts to the more expensive thermal generation later. As compared to the short-term forecasting, it is often more challenging to accurately forecast water inflow in a medium-term horizon. As a result, fluctuations in water inflow can trigger generation shortage and electricity price spikes in a hydrothermal power system.

The existing literature on generation scheduling of hydrothermal power systems can be categorized into centralized

and decentralized approaches. In the centralized approaches, the system operator schedules the generation of each hydro plant in the system to optimize the system payoff (e.g., the total operational cost). A number of deterministic and stochastic optimization approaches have been studied in this category. For the deterministic approaches, [4] proposes a hydrothermal scheduling formulation and a primal decomposition algorithm. Also, [5] develops a detailed formulation for the hydrothermal power system scheduling problem and applies a Lagrangian relaxation algorithm to obtain near-optimal solutions. More recently, [6] utilizes integer programming techniques to accurately model the hill chart of a hydro plant, and [7] extends the hill chart modeling of [6] by adopting a three-dimensional interpolation approach describing the relationship among the net head, turbined water volume, and electricity generation amount. In contrast to [6] and [7], [8] approximates the hill charts by using quadratic functions and solves the hydro-thermal coordination problem to global optimality by considering tight semidefinite programming (SDP) relaxations. Later, [9] further extends the study of the SDP relaxations by accounting for the ac power flow constraints. Taking the Itaipú hydroelectric plant as a case study, [10] employs a dynamic programming approach to optimize the commitment and dispatch of the turbines. In contrast, [11] focuses on the Three Gorges Project of China and elaborates a mixed-integer linear programming (MILP) formulation based on [6] and [7]. For the stochastic approaches, [12] considers a multi-stage stochastic problem for scheduling generation and models the random water inflow via a scenario tree. [3] considers a generation scheduling model that emphasizes the electricity price uncertainty in the deregulated Norwegian market. More recently, [13] formulates the generation scheduling problem as an equilibrium problem with equilibrium constraints, and models the water inflow uncertainty by using a set of scenarios together with their probability of occurrence. From the perspective of coordination, [14] proposes a stochastic MILP model to co-optimize the dispatch of wind power and pumped-storage hydro units. In the decentralized approaches, different market players (e.g., hydro generating companies) schedule their own hydro plants and submit supply offer bids to various electricity markets to optimize their own payoff (e.g., profit), while the system operator clears the markets and coordinates the operations of the players. For price-taking players, [15] develops a target-oriented model that guides the generation scheduling and offer bid submissions to meet a target mean and variance of the water release amount, and [16] models the water inflow and electricity price uncertainty based on a set of scenarios, and considers a risk-averse formulation that contains a risk measure of the payoff. Different from the MILP formulations proposed in [6] and [7], [17] adopts a continuous and nonlinear approximation of the hill charts and optimizes the scheduling of a price-taker based on a mixed-integer nonlinear programming formulation. For price-making players, [18] optimizes the generation scheduling by using a detailed model of the step-wise price quota curve, and [19] adopts the modeling techniques of [18] and incorporates the water inflow uncertainty by considering a multi-stage scenario tree. Later, [20] considers both uncertain residual demand curves and the hill charts of hydro generators

and proposes a stochastic MILP model that is successfully applied to a Portuguese hydro system.

In this paper, we propose a two-stage robust optimization approach for scheduling a hydrothermal power system. Our approach focuses on a centralized power system. For deregulated electricity markets, our approach can also be applied by the system operator to the reliability unit commitment (RUC) runs for maintaining the system robustness. In this approach, we consider the unit commitment (UC) and economic dispatch (ED) for both hydro plants and additional thermal generators. UC decisions are determined in the first-stage problem for all generators, while ED decisions are determined in the second-stage problem for each realized water inflow scenario. Meanwhile, we model the random water inflow by using a distribution-free uncertainty set. That is, we do not assume a specific probability distribution of the random water inflow, but require only that it is within a confidence interval (CI) centered at the forecasted value. In this approach, we aim to make good UC decisions in order to accommodate the worst-case water inflow scenario within its uncertainty set. In comparison, the stochastic optimization approach needs a large number of scenarios to approximate the water inflow uncertainty. In a weekly generation scheduling problem as considered in this paper, this approximation results in a large-scale formulation whose computational burden can be significantly higher than that of the robust optimization (RO) approach. In the literature, RO approaches have recently been applied in various problems, including UC (see, e.g., [21], [22], and [23]), system security (see, e.g., [24]), plug-in hybrid electric vehicles (see, e.g., [25]), and offering strategies (see, e.g., [26] and [27]). A key feature separating our approach from the classical RO ones is that we employ a vector autoregressive (VAR) model to analyze the seasonality of the water inflow, and accordingly generate the forecasted value and CI. Time series models, e.g., VAR models and autoregressive integrated moving average (ARIMA) models, have been used for load forecasting (see, e.g., [28] and [29]), electricity price prediction (see, e.g., [30] and [31]), and wind speed forecasting (see, e.g., [32], [33], and [34]). We summarize the main contributions of this paper as follows:

1. We address the weekly hydrothermal scheduling problem using an adaptive robust optimization framework;
2. We provide a physically based methodology to construct realistic uncertainty sets;
3. We carry out a comprehensive out-of-sample simulation spanning one year and using weekly steps.

In particular, the proposed multiple time-scale VAR uncertainty set is tailored to the natural water inflows and drastically different from the classical polyhedral uncertainty sets (e.g., the ones considered in [35]).

The remainder of this paper is organized as follows. Section II provides the mathematical formulation of the two-stage robust generation scheduling problem. Section III presents uncertainty set definitions and proposes a VAR model for analyzing the random water inflow data and generating CIs. Section IV describes a Benders' decomposition framework to solve the proposed two-stage robust optimization formulation. Section V reports the case studies we conduct based on a 60-bus system incorporating six hydro reservoirs. Finally, Section VI concludes our study.

II. MATHEMATICAL FORMULATION

We present the two-stage robust generation scheduling model for hydrothermal power systems as follows.

$$\begin{aligned}
\min_{y,u,v,e,r} & \sum_{t=1}^T \sum_{i \in \mathcal{I}} (\text{SU}_i u_{it} + \text{SD}_i v_{it} + \text{NL}_i y_{it} + c_i^U r_{it}^U + c_i^D r_{it}^D) \\
& + \max_{f \in \mathcal{W}} [Q(e, r, f)] \quad (1a) \\
\text{s.t.} & y_{i(t-1)} - y_{it} + u_{it} \geq 0, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \quad (1b) \\
& v_{it} = y_{i(t-1)} - y_{it} + u_{it}, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \quad (1c) \\
& y_{it} - y_{i(t-1)} \leq y_{i\tau}, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \\
& \quad \forall \tau = t + 1, \dots, \min\{t + \text{MU}_i - 1, T\}, \quad (1d) \\
& y_{i(t-1)} - y_{it} \leq 1 - y_{i\tau}, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \\
& \quad \forall \tau = t + 1, \dots, \min\{t + \text{MD}_i - 1, T\}, \quad (1e) \\
& e_{it} + r_{it}^U \leq U_i y_{it}, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \quad (1f) \\
& e_{it} - r_{it}^D \geq L_i y_{it}, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \quad (1g) \\
& y_{it}, u_{it}, v_{it} \in \{0, 1\}, e_{it}, r_{it}^U, r_{it}^D \geq 0, \\
& \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \quad (1h)
\end{aligned}$$

where $Q(e, r, f)$ represents the operating cost for given scheduled generation e_{it} , upward/downward reserve r_{it}^U/r_{it}^D (note that notation r subsumes r^U and r^D), and realized water inflow f_{kt} , and equals to the optimal objective function value of the following problem

$$\begin{aligned}
Q(e, r, f) &= \min_{g,h,w,s} \sum_{t=1}^T \sum_{i \in \mathcal{I}} F_i^G(g_{it}) \\
\text{s.t.} & e_{it} - r_{it}^D \leq g_{it} \leq e_{it} + r_{it}^U, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \quad (2a) \\
& g_{i(t+1)} - g_{it} \leq \min\{\text{RU}_i, U_i\}, \\
& \quad \forall i \in \mathcal{I}, \forall t = 1, \dots, T - 1, \quad (2b) \\
& g_{it} - g_{i(t+1)} \leq \min\{\text{RD}_i, U_i\}, \\
& \quad \forall i \in \mathcal{I}, \forall t = 1, \dots, T - 1, \quad (2c) \\
& \sum_{i \in \mathcal{I}} g_{it} = \sum_{b=1}^B D_{bt}, \quad \forall t \in \mathcal{T}, \quad (2d) \\
& -C_{mn} \leq \sum_{b=1}^B K_{mn}^b \left(\sum_{i \in \mathcal{G}_b} g_{it} - D_{bt} \right) \leq C_{mn}, \\
& \forall (m, n) \in \mathcal{L}, \forall t \in \mathcal{T}, \quad (2e) \\
& h_{kt} = h_{k(t-1)} + f_{kt} - \left(\sum_{i \in \mathcal{I}_k} w_{it} + s_{kt} \right) \\
& \quad + \sum_{j \in \mathcal{U}_k} \left(\sum_{\ell \in \mathcal{I}_j} w_{\ell(t-\tau_{jk})} + s_{j(t-\tau_{jk})} \right), \\
& \quad \forall k \in \mathcal{K}, \forall t \in \mathcal{T}, \quad (2f) \\
& h_{k0} = h_k^{\text{initial}}, h_{kT} = h_k^{\text{final}}, \quad \forall k \in \mathcal{K}, \quad (2g) \\
& g_{it} \leq \alpha_{ip} w_{it} + \beta_{ip} y_{it}, \\
& \quad \forall i \in \mathcal{I}_H, \forall t \in \mathcal{T}, \forall p = 1, \dots, P, \quad (2h) \\
& h_k^{\text{MIN}} \leq h_{kt} \leq h_k^{\text{MAX}}, \quad \forall k \in \mathcal{K}, \forall t \in \mathcal{T}, \quad (2i) \\
& w_i^{\text{MIN}} y_{it} \leq w_{it} \leq w_i^{\text{MAX}} y_{it}, \quad \forall i \in \mathcal{I}_H, \forall t \in \mathcal{T}, \quad (2j) \\
& g_{it}, h_{kt}, w_{it}, s_{it} \geq 0, \quad \forall i \in \mathcal{I}, \forall k \in \mathcal{K}, \forall t \in \mathcal{T}. \quad (2k)
\end{aligned}$$

In formulation (1)–(2) described above, we seek to minimize the worst-case total cost including the scheduling cost (i.e., costs for startup, shutdown, and upward/downward reserves) and the operating cost $Q(e, r, f)$. The upward/downward reserve costs in centralized power systems are generally low. As a result, it provides relatively large flexibility for generation scheduling. In this paper, we set the costs of reserves based on historical time series of reserve prices. Note that f can run adversely within the uncertainty set \mathcal{W} so that the operating cost is maximized. Constraints (1b) (respectively, (1c)) describe the generator start-up (respectively, shut-down) operations, constraints (1d) (respectively, (1e)) describe the generator minimum-up time (respectively, minimum-down time) restrictions, and constraints (1f) (respectively, (1g)) describe bounds of upward (respectively, downward) reserve amounts and scheduled generation amounts. (Note that both thermal and hydro units are subject to the UC logic.) In formulation (2), $F_i^G(g_{it})$ represents the fuel cost of generator i in period t incurred by generating g_{it} amount of electricity. We use piece-wise linear functions to approximate fuel cost $F_i^G(g_{it})$ such that $F_i^G(g_{it}) \geq c_{iq}^G g_{it} + d_{iq}^G$ for all $q = 1, \dots, Q$, where Q represents the number of pieces. Constraints (2a) describe bounds on actual generation amounts, constraints (2b)–(2c) describe the ramp-rate limit restrictions of generation, constraints (2d) describe the system balance between generation and load, constraints (2e) describe transmission line capacity restrictions based on the dc approximation (see, e.g., [36] and [37]), constraints (2f) describe the system dynamic of the water volume in each hydro reservoir, constraints (2g) designate the total volume of water in hydro reservoirs at the beginning and end of the operating time intervals (note that h_k^{initial} and h_k^{final} can be different), constraints (2h) approximate the performance curve of each hydro generator by using a (concave) piecewise linear function, and constraints (2i)–(2j) describe the upper and lower bounds of total water volume and turbined water volume in each hydro reservoir, respectively.

For the piecewise linear approximation of performance curves as described in constraints (2h), we observe that at least one piece out of all P pieces are active at optimality for all committed hydro generators, i.e., the hydro generators with $y_{it} = 1$. We summarize this observation as follows and present a proof of it in Appendix A.

Observation 1: There exists an optimal solution to formulations (1)–(2), such that for all $i \in \mathcal{I}_H$ and $t \in \mathcal{T}$ with $y_{it} = 1$, there exists a $p^*(i, t) \in \{1, \dots, P\}$ such that $g_{it} = \alpha_{ip^*(i,t)} w_{it} + \beta_{ip^*(i,t)}$, i.e., constraints (2h) are active.

In this paper, we call formulation (1) the first-stage problem, and (2) the second-stage problem. Note that the second-stage problem is a linear program parametrized by the first-stage variables (e, r) and random variables f . For brevity of notation, we present (2) in a compact form as follows:

$$Q(e, r, f) = \min_x c_x^\top x \quad (3a)$$

$$\text{s.t. } Wx \geq T_e e + T_r r + T_f f + z, \quad (3b)$$

where x represents the second-stage decision variables that subsume variables (g, h, w, s) , and c_x represents the corresponding coefficients in the objective function. Constraints (3b) subsume

constraints (2a)–(2k), where matrices W , T_e , T_r , and T_f represent the coefficients of variables x , e , r , and f respectively, and vector z subsumes the given constants.

Before moving on to the discussions on uncertainty sets, we remark that the proposed robust generation scheduling model (1)–(2) can be extended to model the dependency of the generation level with the reservoir head with higher fidelity. In particular, we can introduce additional binary variables and the corresponding mixed-integer linear constraints proposed by [6] and [7] to formulation (2), in order to provide a non-concave piecewise linear approximation of the head dependency. However, this detailed model will lead to a two-stage robust optimization model with mixed-integer recourse and significantly increase the computational efforts. Indeed, as reported in [7], a deterministic generation scheduling model with a single turbine can incorporate more than 2,000 additional binary variables, and can take up to 2 hours to obtain an optimal solution. Similarly, nonlinear approximation of the hill charts (see, e.g., [8], [9], and [17]) results in nonlinear recourse in the proposed model (1)–(2), which can also be computationally challenging. In this paper, we focus on modeling the uncertain natural water inflow and obtaining good UC solutions (i.e., good first-stage solutions) in the two-stage robust generation scheduling model, for which we might accept reasonable approximations of the second-stage problem to better balance the modeling fidelity and the resulting computational burdens. Meanwhile, we emphasize that it is an important future research direction to explore the possibility of extending formulation (2) for higher fidelity while keeping practical computational burdens.

III. UNCERTAINTY SETS AND A VAR MODEL FOR THE NATURAL WATER INFLOW

In order to guarantee operation robustness, formulations (1)–(2) protect against all possible scenarios of the natural water inflow amount f within uncertainty set \mathcal{W} . Uncertainty sets can be conveniently constructed based on the CIs of the corresponding random variables, which can further be statistically inferred from the historical data of water inflow. For example, following [38] and [22], we can define a budget-constrained uncertainty set for f as follows:

$$\mathcal{W}_{\text{budget}} := \left\{ f : f_t^l \leq f_t \leq f_t^u, \forall t \in \mathcal{T}, \quad (4a) \right.$$

$$\left. \xi_t^l \leq \sum_{k \in \mathcal{K}} f_{kt} \leq \xi_t^u, \forall t \in \mathcal{T}, \quad (4b) \right.$$

$$\left. \varepsilon_k^l \leq \sum_{t \in \mathcal{T}} f_{kt} \leq \varepsilon_k^u, \forall k \in \mathcal{K} \right\}, \quad (4c)$$

where f_t represent a vector of water volumes during time period t , i.e., $f_t := [f_{kt}]_{k \in \mathcal{K}}$, and inequalities (4a) describe lower and upper bounds of each f_{kt} based on, e.g., a 95% CI of f_{kt} . Likewise, based on the corresponding CIs, inequalities (4b) bound the total water volume of the system in each time period and inequalities (4c) bound the total water volume of each water reservoir throughout the operating time intervals, respectively.

By bounding total water volumes, the budget-constrained uncertainty set (4) reflect the spacial dependence among water reservoirs (see inequalities (4b)) and the temporal dependence within each water reservoir (see inequalities (4c)). However,

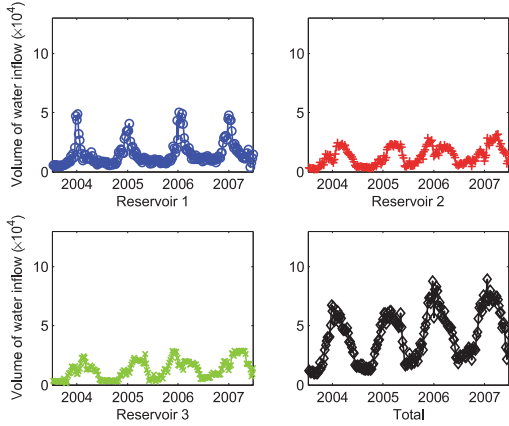


Fig. 1. Weekly total volume of water inflow during years 2004–2007.

such dependence only relies on the total water volumes and is rather implicit. For example, $\mathcal{W}_{\text{budget}}$ allows any patterns of spacial and temporal dependence among f_{kt} as long as they satisfies inequalities (4b)–(4c). This drawback becomes more significant when we model the water inflow uncertainty.

In addition to the budget-constrained uncertainty sets in the form of (4), many other polyhedral uncertainty sets can be formulated and applied in the robust UC framework, including norm-constrained [23] and cardinality-constrained ones ([38] and [26]). Similar to (4), these classical uncertainty sets do not explicitly represent seasonality in their formulations and so can become over-conservative when applied to model water inflows. In this paper, we mention (4) as an example to motivate and benchmark the proposed VAR-based uncertainty set.

For example, Fig. 1 depicts the weekly total volume of water inflow at a series of three water reservoirs along a major river basin in the U.S. during years 2004–2007. From this figure, we can observe (i) a positive correlation among the three water inflow time series, and (ii) a seasonal pattern repeating itself in each year at each water reservoir. Hence, if we know the time of prediction in a year as well as the recent water inflow data, the estimation of water volume can become much more accurate. This observation motivates us to employ a multivariate time series model to analyze the water inflow uncertainty, in order to obtain a tighter uncertainty set as compared to the budget-constrained ones.

Specifically, we construct the uncertainty set \mathcal{W} by using a VAR model. In general, the VAR model assumes that each random variable is a linear function of past lags of itself and past lags of the other random variables. In general, for each $t \in \mathcal{T}$, the VAR model for the hourly water volume can be presented as

$$f_t = f_{\text{const}} + \lambda t + \sum_{\phi=1}^{\phi_{\max}} \Phi_{\phi} f_{t-\phi} + \omega_t,$$

where vector f_{const} represents the constant benchmark, vector λ represents the trend, Φ_{ϕ} represents the matrix describing the dependence of f_t upon up to previous ϕ_{\max} time periods, and vector ω_t represents the white noise in the VAR model. In practice, ω_t is typically modeled as a multivariate zero-mean Normal random vector. A number of statistical toolboxes (e.g., the R package [39]) can be applied to obtain the aforementioned co-

efficients in the VAR model and the covariance matrix of ω_t . Based on the VAR model, we construct the uncertainty set \mathcal{W} as follows:

$$\mathcal{W}_{\text{VAR}} := \left\{ f : f_t = f_{\text{const}} + \lambda t + \sum_{\phi=1}^{\phi_{\max}} \Phi_{\phi} f_{t-\phi} + \omega_t, \right. \\ \left. \forall t \in \mathcal{T}, \right. \quad (5a)$$

$$\sum_{t \in \mathcal{T}_s} f_t = F_{\text{const}} + \Lambda t + \sum_{\psi=1}^{\psi_{\max}} \Psi_{\psi} \left(\sum_{t \in \mathcal{T}_{s-\psi}} f_t \right) + \delta_s, \\ \forall s = 1, \dots, S, \quad (5b)$$

$$\|\omega_t\|_1 \leq \Omega_t^1, \|\omega_t\|_{\infty} \leq \Omega_t^{\infty}, \forall t \in \mathcal{T}, \quad (5c)$$

$$\|\delta_s\|_1 \leq \Delta_s^1, \|\delta_s\|_{\infty} \leq \Delta_s^{\infty}, \\ \forall s = 1, \dots, S \}. \quad (5d)$$

In (5), we incorporate two VAR models on two time scales. First, (5a) follows the hourly VAR model as described above. Second, in (5b), $\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_S$ represent subsets of consecutive time periods in chronological order and form a partition of the operating time intervals, i.e., $\mathcal{T} = \mathcal{T}_1 \cup \mathcal{T}_2 \cup \dots \cup \mathcal{T}_S$. For example, if \mathcal{T} represents all operating hours in a week and $S = 7$, then \mathcal{T}_1 represents operating hours in day 1, \mathcal{T}_2 represents operating hours in day 2, and so on. Hence, (5b) describes another VAR model in a different time scale. Note that (5b) can further narrow the uncertainty set \mathcal{W} and reduce the conservatism of the robust model (1). Other constraints in the form of (5b), possibly defined in other time scales, can also be added to (5) to further reduce the model conservatism. Finally, (5c) and (5d) restrict the white noises in terms of the ℓ_1 - and ℓ_{∞} -norms, where the values of $\Omega_t^1, \Omega_t^{\infty}, \Delta_s^1$, and Δ_s^{∞} can be changed to adjust the robustness level. We note that [40] proposed a time-series-based approach to model the temporal and spatial correlations of wind power. This approach models the seasonal patterns of wind power based on predictions pre-estimated from historical data, while our uncertainty set \mathcal{W}_{VAR} directly models the seasonality of water inflows by using a multiple time-scale VAR model (see (5a)–(5b)).

In practice, the covariance matrices of ω_t and δ_s can be obtained from the R package and the parameters $\Omega_t^1, \Omega_t^{\infty}, \Delta_s^1$, and Δ_s^{∞} can be selected based on numerical simulations of the corresponding Normal random vectors. More specifically, we can generate a set of independent samples of each ω_t and δ_s (denoted as $\{\hat{\omega}_t^n\}_{n=1}^{N_{\omega_t}}$ and $\{\hat{\delta}_s^n\}_{n=1}^{N_{\delta_s}}$, respectively, where N_{ω_t} and N_{δ_s} represent the sample sizes), and select $\Omega_t^1, \Omega_t^{\infty}, \Delta_s^1$, and Δ_s^{∞} to be the 90% quantiles of $\{\|\hat{\omega}_t^n\|_1\}_{n=1}^{N_{\omega_t}}$, $\{\|\hat{\omega}_t^n\|_{\infty}\}_{n=1}^{N_{\omega_t}}$, $\{\|\hat{\delta}_s^n\|_1\}_{n=1}^{N_{\delta_s}}$, and $\{\|\hat{\delta}_s^n\|_{\infty}\}_{n=1}^{N_{\delta_s}}$, respectively.

IV. SOLUTION METHODOLOGY

In this section, we develop a Benders' decomposition (BD) framework to solve the two-stage robust generation scheduling formulation (1)–(2). Our BD framework is similar to the one described in [22] (see Section III for the heuristic separation approach). For brevity of notation, we derive the BD framework based on the compact form (3). First, we note that formulation (3) can be infeasible with some given first-stage decisions (y, u, v, e, r) and realizations of f . In this case, we call the

first-stage decisions (y, u, v, e, r) infeasible and cut them off by identifying feasibility cuts. More specifically, we detect the feasibility of formulation (3) by formulating the following linear program

$$Q^{\text{FC}}(e, r, f) = \min_{x, \zeta} \mathbf{1}^\top \zeta \quad (6a)$$

$$\text{s.t. } Wx + \zeta \geq T_e e + T_r r + T_f f + z, \quad (\pi_{\text{FC}}) \quad (6b)$$

where ζ represent slack variables of constraints (6b), $\mathbf{1}$ represents a vector of all ones, and π^{FC} represent dual variables associated with constraints (6b). We note that formulation (3) is infeasible if and only if $Q^{\text{FC}}(e, r, f) > 0$ at optimality, i.e., $\pi_{\text{FC}}^*(T_e e + T_r r + T_f f + z) > 0$ where π_{FC}^* represent the dual optimal solutions. Hence, we can cut off the infeasible solutions (y, u, v, e, r) by adding the following feasibility cut:

$$\pi_{\text{FC}}^*(T_e e + T_r r + T_f f + z) \leq 0. \quad (7)$$

Second, we take the dual of the linear program in formulation (3) to yield

$$Q(e, r, f) = \max_{\pi} \pi^\top (T_e e + T_r r + T_f f + z) \quad (8a)$$

$$\text{s.t. } \pi \in \Pi := \{\pi \geq 0 : W^\top \pi = c_x\}. \quad (8b)$$

where π represents the dual variables associated with primal constraints (3b), constraints (8b) are associated with primal variables x , and Π represents the feasible region of variables π described by constraints (8b). Note that when the first-stage decisions (y, u, v, e, r) are feasible, the equality in (8a) holds due to the strong duality of linear programs (see, e.g., [41]). Then we obtain a reformulation of the worst-case operating cost as

$$\max_{f \in \mathcal{W}} Q(e, r, f) = \max_{f, \pi} \pi^\top (T_e e + T_r r + T_f f + z) \quad (9a)$$

$$\text{s.t. } \pi \in \Pi, f \in \mathcal{W}. \quad (9b)$$

Third, we substitute the dual formulation (9) into formulation (1) to obtain a reformulation as follows:

$$\begin{aligned} & \min_{y, u, v, e, r} \sum_{t=1}^T \sum_{i \in \mathcal{I}} (\text{SU}_i u_{it} + \text{SD}_i v_{it} + c_i^{\text{U}} r_{it}^{\text{U}} + c_i^{\text{D}} r_{it}^{\text{D}}) \\ & \quad + \max_{f \in \mathcal{W}, \pi \in \Pi} \{ \pi^\top (T_e e + T_r r + T_f f + z) \} \\ & \text{s.t. (1b)-(1h)} \\ & = \min_{y, u, v, e, r, \theta} \sum_{t=1}^T \sum_{i \in \mathcal{I}} (\text{SU}_i u_{it} + \text{SD}_i v_{it} + c_i^{\text{U}} r_{it}^{\text{U}} + c_i^{\text{D}} r_{it}^{\text{D}}) + \theta \\ & \text{s.t. (1b)-(1h),} \\ & \quad \theta \geq \pi^\top (T_e e + T_r r + T_f f + z), \\ & \quad \forall f \in \mathcal{W}, \forall \pi \in \Pi, \end{aligned} \quad (10a)$$

where θ represents the value of $\max_{f \in \mathcal{W}} Q(e, r, f)$, and constraints (10a) follows from formulation (9). For any given pair of $(f, \pi) \in \mathcal{W} \times \Pi$, we call inequality (10a) a Benders' cut pertaining to (f, π) . In our BD framework, we start from a relaxation of the above reformulation by incorporating only a subset of Benders' cuts pertaining to all $f \in \mathcal{W}_0 \subseteq \mathcal{W}$ and $\pi \in \Pi_0 \subseteq \Pi$. The BD framework then iteratively increments \mathcal{W}_0 and Π_0 by adding in more Benders' cuts, and obtains stronger

relaxations till the stopping criterion is satisfied. We summarize the BD framework as follows.

1. Input: lower bound $\text{LB} := -\infty$, upper bound $\text{UB} := +\infty$, optimality gap tolerance ϵ , iteration number limit L , set of Benders' cuts $\text{CUT} := \emptyset$.
2. For $\ell = 1, \dots, L$, repeat the following steps:
 - (a) Solve the master problem

$$\begin{aligned} & \min_{y, u, v, e, r, \theta} \sum_{t=1}^T \sum_{i \in \mathcal{I}} (\text{SU}_i u_{it} + \text{SD}_i v_{it} + c_i^{\text{U}} r_{it}^{\text{U}} + c_i^{\text{D}} r_{it}^{\text{D}}) + \theta \\ & \text{s.t. (1b)-(1h)} \end{aligned}$$

with θ as a decision variable, and the current set of Benders' cuts in CUT as additional constraints. Record optimal solutions $(y^*, u^*, v^*, e^*, r^*, \theta^*)$, and set LB equal to the optimal objective function value.

- (b) Solve formulation (6). Record dual optimal solutions π_{FC}^* and optimal objective function value $Q^{\text{FC}}(e^*, r^*, f^*)$. If $Q^{\text{FC}}(e^*, r^*, f^*) > 0$, then add a feasibility cut (7) into set CUT and go to Step 1.(a). Otherwise, go to Step 1.(c).
- (c) Solve formulation (9). Record optimal solutions (f^*, π^*) and optimal objective function value $V_{2\text{SP}}$, and set UB equal to $\text{LB} - \theta^* + V_{2\text{SP}}$.
- (d) If $|\text{UB} - \text{LB}|/\text{LB} < \epsilon$ or $\theta^* \geq V_{2\text{SP}}$, then return and output $(y^*, u^*, v^*, e^*, r^*)$ as an optimal solution; otherwise, go to the next step.
- (e) Add a Benders' cut

$$\theta \geq (\pi^*)^\top (T_e e + T_r r + T_f f^* + z)$$

into set CUT .

Note that formulation (9) is a bilinear program because of the bilinear product $\pi^\top T_f f$ in the objective function (9a). In general, a bilinear program in the form of problem (9) is NP-hard. In practice, bilinear heuristic (see, e.g., [42] and [22]) has been very helpful to find local optima of problem (9) in a short time, which can then be used in Step (b) above to generate a Benders' cut. In this paper, we apply the alternating direction method in [42] for solving problem (9) and summarize this approach as follows.

1. Input: a starting point $f^0 \in \mathcal{W}$, optimality gap tolerance ϵ . Set iteration counter $j = 1$.
2. Fix variable f to be f^{j-1} and solve problem (9) as a linear program. Record optimal solution π^j and optimal objective function value Q_1^j .
3. Fix variable π to be π^j and solve problem (9) as a linear program. Record optimal solution f^j and optimal objective function value Q_2^j .
4. If $|Q_2^j - Q_1^j|/Q_2^j < \epsilon$, then a local optimum is reached. Fix variable f to be f^j and solve problem (9) as a linear program, record optimal solution π^* , return and output (f^j, π^*) as a locally optimal solution to problem (9). Otherwise, update j to be $j + 1$ and go to Step 1.

V. COMPUTATIONAL RESULTS

In this section, we conduct a computational case study on a hydrothermal power system to demonstrate the proposed robust

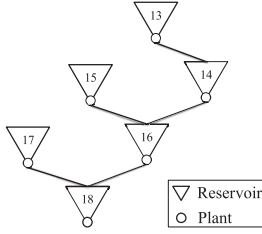


Fig. 2. Topology of the hydro reservoirs.

TABLE I
GENERATOR CHARACTERISTICS – PART I.

Generator	Bus	U_i	L_i	RU_i	RD_i	MU_i	MD_i
1,2	1,31	200	50	75	75	5	5
3,4	3,33	100	10	40	40	1	1
5,6	7,37	120	15	50	50	1	1
7,8	9,39	120	10	50	50	2	2
9,10	12,42	150	20	60	60	2	2
11,12	20,50	300	60	80	80	4	4
13,14	26,56	500	25	400	400	4	4
15,16	11,41	600	25	500	500	5	5
17,18	4,34	600	25	600	600	8	8

TABLE II
GENERATOR CHARACTERISTICS – PART II.

Generator	SU_i	SD_i	NL_i	c_{i1}^G	c_i^U	c_i^D
1,2	20	20	5	10	8	8
3,4	50	50	10	20	15	15
5,6	50	50	5	35	17	17
7,8	50	50	10	40	20	20
9,10	50	50	10	50	25	25
11,12	25	25	5	10	10	10
13,14	10	10	5	5	2	2
15,16	15	15	5	4	2	2
17,18	20	20	10	3	1	1

TABLE III
HYDRO RESERVOIR CAPACITIES.

Reservoir	h_k^{MAX}	h_k^{MIN}
13,14	122	4
15,16	143	5
17,18	240	6

approaches and solution algorithm. We describe the experiment setup in Section V-A and present detailed results in Section V-B.

A. Experiment Setup

The case study is based on the IEEE 30-bus system that contains 30 nodes and 41 transmission lines (see <http://www.ee.washington.edu/research/pstca/> for the data set). First, in this case study, we duplicate the 30-bus system to make a 60-bus one, and place 12 thermal generators and 6 hydro reservoirs in this system (see Table I). We assume that these six hydro reservoirs are interconnected as displayed in Fig. 2. We also assume that each hydro reservoir is equipped with a hydro generator, and that the time delay between each pair of reservoirs is $\tau_{jk} = 2$. We have tested various selections of $\tau_{jk} \in \{1, 2\}$ and the results and conclusions obtained are similar. We only report the results for $\tau_{jk} = 2$ for presentation brevity.

We provide generator characteristics in Tables I–IV. In particular, generators 1–12 are thermal and generators 13–18 are

TABLE IV
PERFORMANCE CURVE COEFFICIENTS.

Generator	α_{i1}	α_{i2}	α_{i3}	β_{i1}	β_{i2}	β_{i3}
13,14	0.8	0.5	0.3	0	31.13	142.37
15,16	0.75	0.6	0.35	0	35.38	154.05
17,18	0.7	0.4	0.2	0	38.51	157.62

hydro. In this case study, we approximate the fuel cost functions of the thermal generators by using linear functions (i.e., $Q = 1$), and assume that hydro generators are predominant in the power system and take up a major portion of the electricity generation capacity as summarized in Tables I–II. Moreover, the hydro reservoir characteristics are summarized in Tables III–IV. Finally, we assume that the initial and final volumes of water in each hydro reservoir are designated as

$$h_{k\ell}^{\text{initial}} = h_k^{\text{MAX}} \left[80\% + 0.5\% \sin \left(\frac{(\ell - 1)(2\pi)}{52} - \frac{\pi}{4} \right) \right],$$

$$h_{k\ell}^{\text{final}} = h_k^{\text{MAX}} \left[80\% + 0.5\% \sin \left(\frac{\ell(2\pi)}{52} - \frac{\pi}{4} \right) \right]$$

for all $k \in \mathcal{K}$, where $\ell = 1, \dots, 52$ represents an index for weeks.

Second, we construct uncertainty sets $\mathcal{W}_{\text{budget}}$ and \mathcal{W}_{VAR} based on the hourly natural water inflow data collected at a series of three hydro reservoirs along a major river basin in the U.S. during years 1999–2007. The hourly water inflow data is reverse engineered from the Supervisory Control and Data Acquisition (SCADA) data. We apply the data analytic results and the obtained uncertainty sets from the first hydro reservoir to hydro reservoirs 13 and 14, the second one to hydro reservoirs 15 and 16, and the third one to hydro reservoirs 17 and 18, respectively. More specifically, we construct $\mathcal{W}_{\text{budget}}$ by computing the empirical 95% CIs of the hourly, daily, and weekly total water inflow for each water reservoir, respectively. Also, we consider the seasonality of the water inflow by focusing on similar relative time windows in previous years. For example, to construct the CI of the weekly total water inflow in week 2 of year 2007, we consider the data in weeks 1–3 of years 1999–2006. Also, we construct \mathcal{W}_{VAR} by analyzing the VAR models of the hourly, daily, and weekly total water inflow in the six hydro reservoirs. All time series analyses are conducted using the R package [39], and the time lags we use for the hourly, daily, and weekly VAR models are 96 hours, 28 days, and 52 weeks respectively.

Finally, we generate loads spanning one whole year based on the IEEE RTS-96 system ([43]), taking into account the load seasonality and the weekday-weekend load difference. All problem instances were solved by using CPLEX 12.3 at a Linux server with Intel Xeon Quad Core 2.93 GHz and 2GB memory.

B. Experiment Results

In this case study, we simulate to schedule the hydrothermal power system for 52 weeks (i.e., 1 year) from Mar. 5th, 2006 to Mar. 4th, 2007. In each week during this year, we make weekly scheduling decisions for the hydrothermal power system by solving the proposed two-stage robust optimization model (1) with $T = 168$ hours. For each weekly scheduling problem,

we (i) conduct data analyses based on the historical data available up to the time point of making schedules, (ii) construct uncertainty sets $\mathcal{W}_{\text{budget}}$ or \mathcal{W}_{VAR} based on the data analyses, and (iii) employ the Benders' decomposition algorithm described in the BD framework to obtain a robust optimal schedule. Based on whether uncertainty set $\mathcal{W}_{\text{budget}}$ or \mathcal{W}_{VAR} is used, we denote the schedule obtained as the budget-robust (BR) schedule or the VAR-robust (VR) schedule, respectively.

To benchmark BR and VR schedules and evaluate the conservativeness of the robust approach, we consider a perfect-information problem by assuming that the actual water inflow in the incumbent week is known before we make the schedule. In other words, the perfect-information problem is a special case of formulation (1) by setting $\mathcal{W} = \{f_{\text{real}}\}$, where f_{real} represents the actual data during the week under investigation. Note that the perfect-information problem provides the actual optimal scheduling decisions one can obtain if the natural water inflow can be perfectly predicted beforehand, and so is guaranteed to outperform any other approaches. We denote the schedule thus obtained as the perfect-information (PI) schedule. In this case study, we compare the BR, VR, and PI schedules by computing the following values:

1. V_{BR} , representing the optimal objective value of formulation (1) with budget-constrained uncertainty set $\mathcal{W}_{\text{budget}}$.
2. V_{VR} , representing the optimal objective value of formulation (1) with VAR uncertainty set \mathcal{W}_{VAR} .
3. V_{PI} , representing the optimal objective value of the perfect-information problem.
4. $\text{RO-Gap}_{\text{BR}} = (V_{\text{BR}} - V_{\text{PI}})/V_{\text{PI}} \times 100\%$, representing the worst-case relative gap between the budget-robust and the perfect-information schedules.
5. $\text{RO-Gap}_{\text{VR}} = (V_{\text{VR}} - V_{\text{PI}})/V_{\text{PI}} \times 100\%$, representing the worst-case relative gap between the VAR-robust and the perfect-information schedules.

Based on V_{PI} , $\text{RO-Gap}_{\text{BR}}$ and $\text{RO-Gap}_{\text{VR}}$ indicate the amount of additional operational cost to be incurred if the worst-case scenario in the corresponding uncertainty set realizes. As the worst-case scenario is not likely to occur, $\text{RO-Gap}_{\text{BR}}$ and $\text{RO-Gap}_{\text{VR}}$ roughly measure how conservative the two robust schedules are when we consider $\mathcal{W}_{\text{budget}}$ and \mathcal{W}_{VAR} , respectively. To provide a more direct and intuitive measure of conservativeness, we also conduct a post-optimization simulation to evaluate the BR and VR schedules. In this simulation, we fix the binary scheduling decisions (i.e., on/off status, start-up, and shut-down) to be the optimal first-stage binary decisions obtained from the robust approaches, and resolve the second-stage problem (2) by fixing the water inflow f to be the real data f_{real} . More specifically, for the BR schedule, we denote the obtained budget-robust optimal binary solution as (y^*, u^*, v^*) . Then we resolve the second-stage problem (2) by replacing constraints (2a) to be $L_i y_{it} \leq g_{it} \leq U_i y_{it}$ for all $i \in \mathcal{I}$ and $t \in \mathcal{T}$, and fixing $(y, u, v) = (y^*, u^*, v^*)$. We let S_{BR} represent the optimal objective value of the resolved second-stage problem, plus the first-stage fixed cost $\sum_{t=1}^T \sum_{i \in \mathcal{I}} (S U_i u_{it}^* + S D_i v_{it}^* + N L_i y_{it}^*)$. We note that S_{BR} reflects how the budget-robust schedules perform in practice using the real water inflows, and we can similarly obtain S_{VR} to evaluate the VAR-robust schedules. In this case study, we

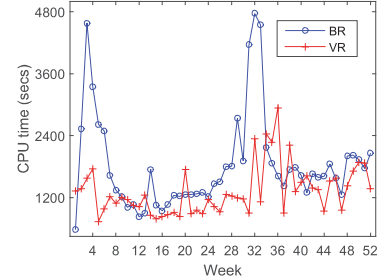


Fig. 3. CPU time used to solve problem instances.

TABLE V
CPU TIME IN SECONDS.

	BR	VR
Average	1832.67	1322.26
St.Dev.	929.40	472.86
Min	580.50	737.49
Max	4767.05	2939.14

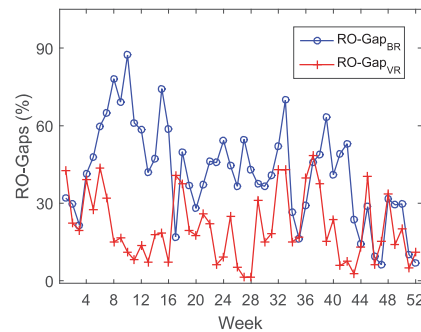


Fig. 4. Worst-case relative gaps between the robust and perfect-information schedules.

TABLE VI
WORST-CASE RELATIVE GAPS OF THE ROBUST SCHEDULES.

	RO-Gap _{BR}	RO-Gap _{VR}
Average	41.67	20.70
St.Dev.	18.57	13.10
Min	6.17	1.26
Max	87.37	48.50

evaluate the conservativeness of the two robust approaches in practice by computing the following values:

1. $S\text{-Gap}_{\text{BR}} = (S_{\text{BR}} - V_{\text{PI}})/V_{\text{PI}} \times 100\%$, representing the actual relative gap between the budget-robust and the perfect-information schedules.
2. $S\text{-Gap}_{\text{VR}} = (S_{\text{VR}} - V_{\text{PI}})/V_{\text{PI}} \times 100\%$, representing the actual relative gap between the VAR-robust and the perfect-information schedules.

First, we report the CPU time needed to solve each problem instance in Fig. 3 and the corresponding statistics in Table V. We observe that the robust weekly schedules can be found within 2 hours for both models and within 50 minutes for the VAR-robust approach, in all 104 problem instances. This demonstrates the effectiveness of the proposed BD framework to solve the two-stage robust optimization formulation.

Second, we report the worst-case relative gaps $\text{RO-Gap}_{\text{BR}}$ and $\text{RO-Gap}_{\text{VR}}$ throughout all 52 weeks in Fig. 4, and the corresponding statistics in Table VI. From Table VI, we observe that the worst-case relative gaps between the BR and the PI schedules

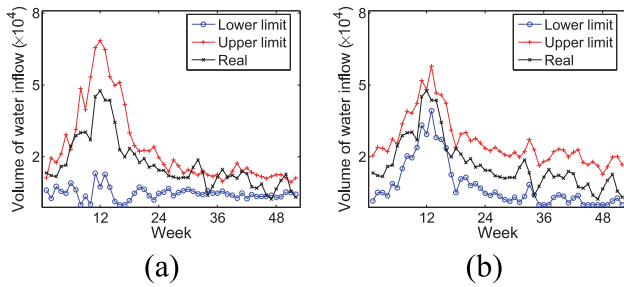


Fig. 5. CIs of the weekly water inflows of a hydro reservoir from $\mathcal{W}_{\text{budget}}$ and \mathcal{W}_{VAR} . (a) CI from $\mathcal{W}_{\text{budget}}$. (b) CI from \mathcal{W}_{VAR} .

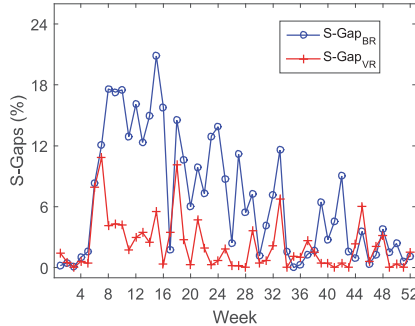


Fig. 6. Actual relative gaps between the robust and perfect-information schedules.

are above 40% on average, and the ones between the VR and the PI schedules are around 20% on average. In other words, the total operational cost of the hydrothermal power system can increase by more than 40% if the worst-case natural water inflow takes place under the BR schedules, and this relative increase drops to around 20% if the VR schedule is implemented instead. From Fig. 4, we observe that the VR schedules tend to perform less conservatively than the BR schedules throughout the 52 weeks. More specifically, we have $\text{RO-Gap}_{\text{VR}} < \text{RO-Gap}_{\text{BR}}$ during 43 weeks, or 83% of the year.

This conclusion can also be confirmed by our observations that (i) \mathcal{W}_{VAR} is often tighter than $\mathcal{W}_{\text{budget}}$, and (ii) \mathcal{W}_{VAR} is more likely to correctly predict the real water inflow f_{real} , i.e., $f_{\text{real}} \in \mathcal{W}_{\text{VAR}}$. For example, we depict in Fig. 5 the CIs of the weekly water inflow of the hydro reservoir at bus 26, predicted by $\mathcal{W}_{\text{budget}}$ and \mathcal{W}_{VAR} respectively. From this figure, we observe that the CIs obtained from \mathcal{W}_{VAR} are often tighter than the ones obtained from $\mathcal{W}_{\text{budget}}$ (compare the two curves with specifier '+' and the two curves with specifier 'o'). Similar observations can be made from the other five hydro reservoirs. Furthermore, the CIs obtained from \mathcal{W}_{VAR} also outperform in correctly predicting f_{real} (e.g., the curve with specifier 'x' should lie in between the other two curves in Fig. 5). More specifically, $f_{\text{real}} \in \mathcal{W}_{\text{VAR}}$ in 50 out of 52 weeks (i.e., 96.15% of the year), as compared to $f_{\text{real}} \in \mathcal{W}_{\text{budget}}$ in 41 out of 52 weeks (i.e., 78.85% of the year).

Third, we report the actual relative gaps S-Gap_{BR} and S-Gap_{VR} throughout all 52 weeks in Fig. 6, and the corresponding statistics in Table VII. We observe that the actual relative gaps between the BR and the PI schedules are within 25% in all 52 weeks, and are 6.72% on average. In comparison, the VR schedules have smaller actual relative gaps (within 15% in all 52 weeks, and 2.20% on average). In particular, the

TABLE VII
ACTUAL RELATIVE GAPS OF THE ROBUST SCHEDULES.

	S-Gap_{BR}	S-Gap_{VR}
Average	6.72	2.20
St.Dev.	5.94	2.53
Min	0.00	0.00
Max	20.86	10.85

actual relative gaps for the VR schedules are within 5% in 46 out of 52 weeks (i.e., 88.46% of the year, see Fig. 6). Again, this observation can be confirmed by the tight uncertainty set \mathcal{W}_{VAR} as depicted in Fig. 5. Hence, we can conclude that the VR schedules are less conservative than the BR schedules in general, and can perform very similarly to the PI schedules in a majority part of the year. Meanwhile, as the PI approach is guaranteed to perform the best in this problem, small gaps S-Gap_{VR} indicate that the VR schedules can compare favorably with other alternatives, e.g., the schedules obtained from stochastic optimization approaches.

VI. CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

In this paper, we proposed a two-stage robust optimization model for scheduling a hydrothermal power system in a weekly horizon. In particular, we focus on the uncertain water inflows and consider two types of robust approaches, one based on the classical budget-constrained uncertainty set and the other based on a VAR time series model. Based on real data, we conducted a 52-week case study on a hydrothermal power system and applied the robust schedules obtained from these two approaches. The experiment results indicated that the VAR-robust schedules can take longer to obtain, but they are less conservative than the budget-robust schedules in general, and perform comparably well as compared with the perfect-information schedules in a majority part of the year.

An important future research direction is to explore the possibility of extending formulation (2) for higher fidelity by adopting a non-concave piecewise linear approximation of the hill charts (such as the ones proposed by [6] and [7]), while keeping practical computational burdens.

APPENDIX

PROOF OF OBSERVATION 1

Proof: Let (y^*, g^*, w^*, s^*) represent a part of an optimal solution to formulations (1)–(2), and suppose that there exists an $i_1 \in \mathcal{I}_H$ and $t_1 \in \mathcal{T}$ with $y_{i_1 t_1}^* = 1$, but $g_{i_1 t_1}^* < \alpha_{i_1 p} w_{i_1 t_1}^* + \beta_{i_1 p}$ for all $p = 1, \dots, P$. We define $\varpi = \min\{w_{i_1 t_1}^* + \beta_{i_1 p}/\alpha_{i_1 p} - g_{i_1 t_1}^*/\alpha_{i_1 p} : p = 1, \dots, P\}$ and let $p_1 \in \{1, \dots, P\}$ represent an index where the minimum is attained. Furthermore, we define $\hat{w}_{i_1 t_1} = w_{i_1 t_1}^* - \varpi$, $\hat{w}_{it} = w_{it}^*$ for all $i \neq i_1$ or $t \neq t_1$, $\hat{s}_{k_1 t_1} = s_{k_1 t_1}^* + \varpi$ for $i_1 \in \mathcal{I}_{k_1}$, and $\hat{s}_{kt} = s_{kt}^*$ for all $k \neq k_1$ or $t \neq t_1$. Now we generate a new solution by replacing w^* and s^* with \hat{w} and \hat{s} , respectively, in the optimal solution. It follows that constraints (2f) remain satisfied under this replacement because

$$\begin{aligned} \sum_{i \in \mathcal{I}_{k_1}} \hat{w}_{it_1} + \hat{s}_{k_1 t_1} &= \sum_{i \in \mathcal{I}_{k_1}} w_{it_1}^* - \varpi + s_{k_1 t_1}^* + \varpi \\ &= \sum_{i \in \mathcal{I}_{k_1}} w_{it_1}^* + s_{k_1 t_1}^*. \end{aligned}$$

Meanwhile, constraints (2h) remain satisfied because

$$\begin{aligned} & \alpha_{i_1 p} \hat{w}_{i_1 t_1} + \beta_{i_1 p} y_{i_1 t_1}^* - g_{i_1 t_1}^* \\ &= \alpha_{i_1 p} w_{i_1 t_1}^* + \beta_{i_1 p} - g_{i_1 t_1}^* - \alpha_{i_1 p} \varpi \\ &\geq \alpha_{i_1 p} w_{i_1 t_1}^* + \beta_{i_1 p} - g_{i_1 t_1}^* \\ &\quad - \alpha_{i_1 p} \left(w_{i_1 t_1}^* + \frac{\beta_{i_1 p}}{\alpha_{i_1 p}} - \frac{g_{i_1 t_1}^*}{\alpha_{i_1 p}} \right) = 0 \end{aligned}$$

for all $p \in \{1, \dots, P\}$, where the first equality is due to the definition of $\hat{w}_{i_1 t_1}$ and $y_{i_1 t_1}^* = 1$, and the inequality is due to the definition of ϖ . In particular, if $p = p_1$, then $\alpha_{i_1 p_1} \hat{w}_{i_1 t_1} + \beta_{i_1 p_1} y_{i_1 t_1}^* - g_{i_1 t_1}^* = 0$ and so constraint (2h) is active. Furthermore, we observe that the objective function value remains the same under this replacement. Therefore, this replacement results in another optimal solution to formulation (1)–(2) with constraint (2h) being active at at least one pieces for $i = i_1$ and $t = t_1$. The proof is completed in view that we can perform similar replacement for all other i, t indices where constraints (2h) are inactive for any pieces. \square

REFERENCES

- [1] Operador Nacional do Sistema Elétrico Brazil [Online]. Available: http://www.ons.org.br/institucional/modelo_setorial.aspx?lang=en
- [2] Electricity Authority, New Zealand, Retail and Wholesale Market [Online]. Available: <http://www.ea.govt.nz/>
- [3] O. B. Fosso, A. Gjelsvik, A. Haugstad, B. Mo, and I. Wangensteen, "Generation scheduling in a deregulated system. The Norwegian case," *IEEE Trans. Power Syst.*, vol. 14, no. 1, pp. 75–81, Feb. 1999.
- [4] A. A. Carneiro, S. Soares, and P. S. Bond, "A large scale of an optimal deterministic hydrothermal scheduling algorithm," *IEEE Trans. Power Syst.*, vol. 5, no. 1, pp. 204–211, Feb. 1990.
- [5] H. Yan, P. B. Luh, X. Guan, and P. M. Rogan, "Scheduling of hydrothermal power systems," *IEEE Trans. Power Syst.*, vol. 8, no. 3, pp. 1358–1365, Aug. 1993.
- [6] A. J. Conejo, J. M. Arroyo, J. Contreras, and F. A. Villamor, "Self-scheduling of a hydro producer in a pool-based electricity market," *IEEE Trans. Power Syst.*, vol. 17, no. 4, pp. 1265–1272, Nov. 2002.
- [7] A. Borghetti, C. D'Ambrosio, A. Lodi, and S. Martello, "An MILP approach for short-term hydro scheduling and unit commitment with head-dependent reservoir," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1115–1124, Aug. 2008.
- [8] Y. Zhu, J. Jian, J. Wu, and L. Yang, "Global optimization of non-convex hydro-thermal coordination based on semidefinite programming," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 3720–3728, Nov. 2013.
- [9] M. Paredes, L. S. A. Martins, and S. Soares, "Using semidefinite relaxation to solve the day-ahead hydro unit commitment problem," *IEEE Trans. Power Syst.*, vol. 30, no. 3, pp. 2695–2705, May 2015.
- [10] A. Arce, T. Ohishi, and S. Soares, "Optimal dispatch of generating units of the Itaipú hydroelectric plant," *IEEE Trans. Power Syst.*, vol. 17, no. 1, pp. 154–158, Feb. 2002.
- [11] X. Li, T. Li, J. Wei, G. Wang, and W. W.-G. Yeh, "Hydro unit commitment via mixed integer linear programming: A case study of the Three Gorges Project, China," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1232–1241, May 2014.
- [12] M. V. F. Pereira and L. M. V. G. Pinto, "Stochastic optimization of a multireservoir hydroelectric system: A decomposition approach," *Water Resources Res.*, vol. 21, no. 6, pp. 779–792, 1985.
- [13] K. C. Almeida and A. J. Conejo, "Medium-term power dispatch in predominantly hydro systems: An equilibrium approach," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2384–2394, Aug. 2013.
- [14] M. E. Khodayar, M. Shahidehpour, and L. Wu, "Enhancing the dispatchability of variable wind generation by coordination with pumped-storage hydro units in stochastic power systems," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2808–2818, Aug. 2013.
- [15] G. Pritchard, A. B. Philpott, and P. J. Neame, "Hydroelectric reservoir optimization in a pool market," *Math. Program.*, vol. 103, no. 3, pp. 445–461, 2005.
- [16] L. Wu, M. Shahidehpour, and Z. Li, "Genco's risk-constrained hydrothermal scheduling," *IEEE Trans. Power Syst.*, vol. 23, no. 4, pp. 1847–1858, Nov. 2008.
- [17] F. J. Díaz, J. Contreras, J. I. Muñoz, and D. Pozo, "Optimal scheduling of a price-taker cascaded reservoir system in a pool-based electricity market," *IEEE Trans. Power Syst.*, vol. 26, no. 2, pp. 604–615, May 2011.
- [18] S. de la Torre, J. M. Arroyo, A. J. Conejo, and J. Contreras, "Price maker self-scheduling in a pool-based electricity market: A mixed-integer LP approach," *IEEE Trans. Power Syst.*, vol. 17, no. 4, pp. 1037–1042, Nov. 2002.
- [19] C. G. Baslis and A. G. Bakirtzis, "Mid-term stochastic scheduling of a price-maker hydro producer with pumped storage," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 1856–1865, Nov. 2011.
- [20] H. M. I. Pousinho, J. Contreras, and J. P. S. Catalao, "Short-term optimal scheduling of a price-maker hydro producer in a pool-based day-ahead market," *IET Gener., Transm., Distrib.*, vol. 6, no. 12, pp. 1243–1251, 2012.
- [21] A. Thiele, T. Terry, and M. Epelman, "Robust Linear Optimization with Recourse," *Tech. Rep.*. Bethlehem, PA, USA: Lehigh Univ., 2009.
- [22] R. Jiang, M. Zhang, G. Li, and Y. Guan, "Two-stage network constrained robust unit commitment problem," *Eur. J. Oper. Res.*, vol. 234, no. 3, pp. 751–762, 2014.
- [23] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 52–63, Feb. 2013.
- [24] Q. Wang, J.-P. Watson, and Y. Guan, "Two-stage robust optimization for N-k contingency-constrained unit commitment," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2366–2375, Aug. 2013.
- [25] A. H. Hajimiragha, C. A. Canizares, M. W. Fowler, S. Moazeni, and A. Elkamel, "A robust optimization approach for planning the transition to plug-in hybrid electric vehicles," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2264–2274, Nov. 2011.
- [26] L. Baringo and A. J. Conejo, "Offering strategy via robust optimization," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1418–1425, Aug. 2011.
- [27] L. Fan, J. Wang, R. Jiang, and Y. Guan, "Min-max regret bidding strategy for thermal generator considering price uncertainty," *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2169–2179, Sep. 2014.
- [28] G. Gross and F. D. Galiana, "Short-term load forecasting," *Proc. IEEE*, vol. 75, no. 12, pp. 1558–1573, Dec. 1987.
- [29] M. T. Hagan and S. M. Behr, "The time series approach to short term load forecasting," *IEEE Trans. Power Syst.*, vol. 2, no. 3, pp. 785–791, Aug. 1987.
- [30] F. J. Nogales, J. Contreras, A. J. Conejo, and R. Espinola, "Forecasting next-day electricity prices by time series models," *IEEE Trans. Power Syst.*, vol. 17, no. 2, pp. 342–348, May 2002.
- [31] J. Contreras, R. Espinola, F. J. Nogales, and A. J. Conejo, "ARIMA models to predict next-day electricity prices," *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1014–1020, Aug. 2003.
- [32] L. Söder, "Simulation of wind speed forecast errors for operation planning of multi-area power systems," in *Proc. 2004 IEEE Int. Conf. Probabilistic Methods Appl. to Power Systems*, 2004, pp. 723–728.
- [33] J. Matevosyan and L. Söder, "Optimal daily planning for hydro power system coordinated with wind power in areas with limited export capability," in *Proc. 2006 IEEE Int. Conf. Probabilistic Methods Appl. to Power Systems*, 2006, pp. 1–8.
- [34] W. Wangdee and R. Billinton, "Probing the intermittent energy resource contributions from generation adequacy and security perspectives," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2306–2313, Nov. 2012.
- [35] C. Zhao, J. Wang, J.-P. Watson, and Y. Guan, "Multi-stage robust unit commitment considering wind and demand response uncertainties," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2708–2717, Aug. 2013.
- [36] A. R. Bergen and V. Vittal, *Power Systems Analysis*, 2nd ed. Englewood Cliffs, NJ, USA: Prentice Hall, 1999.
- [37] A. Gómez-Expósito, A. J. Conejo, and C. Cañizares, *Electric Energy Systems: Analysis and Operation*. Boca Raton, FL, USA: CRC Press, 2008.
- [38] D. Bertsimas and M. Sim, "The price of robustness," *Oper. Res.*, vol. 52, no. 1, pp. 35–53, 2004.
- [39] R: A Language and Environment for Statistical Computing. Vienna, Austria, R Core Team, R Foundation for Statistical Computing, 2014 [Online]. Available: <http://www.R-project.org>
- [40] Á. Lorca and X. A. Sun, "Adaptive robust optimization with dynamic uncertainty sets for multi-period economic dispatch under significant wind," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1702–1713, Jul. 2015.
- [41] D. G. Luenberger and Y. Ye, *Linear and Nonlinear Programming*. Springer, 2008.

- [42] H. Konno, "A cutting plane algorithm for solving bilinear programs," *Math. Program.*, vol. 11, no. 1, pp. 14–27, 1976.
- [43] The Reliability Test System Task Force, "The IEEE reliability test system – 1996," *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 1010–1020, Aug. 1999.



Hossein Dashti (S'15) received the B.S. degree in Industrial Engineering from the University of Tehran, Tehran, Iran, in 2013. He is currently pursuing the Ph.D. degree in the Department of Systems and Industrial Engineering at the University of Arizona, Tucson, AZ, USA.

His research interests include robust optimization of power system operations and renewable energy management.



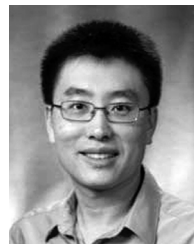
Antonio J. Conejo (F'04) received the M.S. degree from MIT, Cambridge, MA, USA, in 1987, and the Ph.D. degree from the Royal Institute of Technology, Stockholm, Sweden, in 1990.

He is currently a full professor at the Integrated System Engineering and the Electrical & Computer Engineering Departments, The Ohio State University, Columbus, OH, USA. His research interests include control, operations, planning, economics and regulation of electric energy systems, as well as statistics and optimization theory and its applications.



Ruiwei Jiang (M'14) received the B.S. degree in Industrial Engineering from the Tsinghua University, Beijing, China, in 2009, and the Ph.D. degree in Industrial and Systems Engineering from the University of Florida, Gainesville, FL, USA, in 2013.

Presently, he is an Assistant Professor with the Department of Industrial and Operations Engineering at the University of Michigan, Ann Arbor, MI, USA. His research interests include power system planning and operations, renewable energy management, and water distribution operations and system analysis.



Jianhui Wang (M'07–SM'12) received the Ph.D. degree in electrical engineering from Illinois Institute of Technology, Chicago, IL, USA, in 2007.

Presently, he is the Section Lead for Advanced Power Grid Modeling at the Energy Systems Division at Argonne National Laboratory, Argonne, IL, USA.

Dr. Wang is the secretary of the IEEE Power & Energy Society (PES) Power System Operations Committee. He is an associate editor of the *Journal of Energy Engineering* and an editorial board member of *Applied Energy*. He is also an affiliate professor at Auburn University and an adjunct professor at University of Notre Dame. He has held visiting positions in Europe, Australia and Hong Kong including a VELUX Visiting Professorship at the Technical University of Denmark (DTU). He is the Editor-in-Chief of the IEEE TRANSACTIONS ON SMART GRID and an IEEE PES Distinguished Lecturer. He is also the recipient of the IEEE PES Power System Operation Committee Prize Paper Award in 2015.