Stochastic Optimization Approach to Water Management in Cooling-Constrained Power Plants*

Juan M. Salazar and Urmila Diwekar
Vishwamitra Research Institute
Center for Uncertain Systems: Tools for Optimization and Management
Clarendon Hills, IL 60439, USA
Emil Constantinescu and Victor M. Zavala
Mathematics and Computer Science Division
Argonne National Laboratory, 9700 South Cass Avenue, Argonne, IL 60439, USA

Abstract

We propose a stochastic optimization framework to perform water management in cooling-constrained power plants. The approach determines optimal set-points to maximize power output in the presence of uncertain weather conditions and water intake constraints. Weather uncertainty is quantified in the form of ensembles using the state-of-the-art numerical weather prediction model WRF. The framework enables us to handle first-principles black-box simulation models and to construct empirical distributions from limited samples obtained from WRF. Using these capabilities, we investigate the effects of cooling capacity constraints and weather conditions on generation capacity. In a pulverized coal power plant study we have found that weather fluctuations make the maximum plant output vary in the range of 5-10% of the nominal capacity in intraday operations. In addition, we have found that stochastic optimization can lead to *daily* capacity gains of as much as 245 MWh over current practice and enables more robust bidding procedures. We demonstrate that the framework is computationally feasible.

Keywords: stochastic optimization, power plants, weather forecasting, uncertainty, water constraints.

1 Introduction

The U.S. geological survey estimates that thermoelectric generation accounts for approximately 136,000 million gallons per day of *fresh* water withdrawals, ranking only slightly behind agricultural irrigation as the largest source of freshwater withdrawals in the United States [10]. In base-load power plants (i.e., coal and nuclear) huge amounts of water are lost to the environment in the cooling towers as a result of evaporation. A 500 MW coal-fired power plant that employs once-through cooling can use more than 10 million gallons per hour of water for cooling and other process requirements.

*Preprint ANL/MCS-P2026-0212.

The latest forecast estimates of the Energy Information Administration indicate that U.S. coal-fired generating capacity will grow from approximately 305 GW in 2004 to 453 GW in 2030. This growth is a concern because of decreasing fresh water levels and the increasing frequency and severity of drought conditions [2, 1].

Water consumption in power plants is intimately related to weather conditions such as ambient temperature and humidity because these conditions affect the capacity of the cooling towers. For instance, Independent System Operators (ISOs) normally represent plant capacity as a function of ambient temperature in clearing procedures. Unfavorable weather conditions can force power plants to drop their power outputs or even to fully shut down, thus limiting their participation in electricity markets. Consequently, weather conditions can have a strong effect on market prices and grid reliability particularly if the plant is located in an area with significant transmission congestion [8]. Reduced water availability and increasing environmental regulation suggest that a more restricted water intake will be enforced in power plants in order to ensure long-term availability. This will force cooling systems to operate closer to their maximum capacity and will make them highly sensitive to weather conditions. Consequently, quantifying the uncertainty of the ambient conditions and modifying operations in real time will become increasingly important.

Stochastic optimization has been widely used to coordinate generation at central grid levels [24, 25, 5, 7]. Studies at the power plant level have been focused mostly on determining optimal bidding strategies [17]. These studies, however, neglect detailed physical limitations of the power plant. Recently, a stochastic optimization approach has been proposed to minimize water consumption in power plants [22, 21]. Here, the authors determine optimal operating conditions (set-points for the control system) for a pulverized coal power plant that minimize water consumption under fluctuations of load and seasonal ambient conditions. Empirical distributions for ambient temperature and relative humidity were constructed by using historical weather data.

In this work, we use the numerical weather prediction system WRF to obtain forecast and uncertainty information [23, 7]. This enables us to obtain physical representations of uncertainty and to investigate their size, shape, and accuracy. We use a specialized resolution targeting scheme in WRF to enable fine resolutions at the location of the power plant. This approach has been used in diverse studies to obtain weather information affecting energy system performance [27, 6]. We present a detailed case study integrating uncertainty quantification with WRF, stochastic optimization, and a detailed power plant first-principles model to investigate computational feasibility, effects of weather, and uncertainty levels on plant performance and to evaluate potential over the state of the art. To the best of our knowledge, this is the first study that integrates physical weather and power plant models. Our study demonstrates that stochastic optimization can yield savings in daily plant capacity of more than 245 MWh over current practice. In addition, this approach enables more robust bidding procedures.

We use the BONUS solver, which implements a reweighting-based algorithm, to solve the associated nonlinear stochastic optimization problem [19]. We argue that reweighting is a powerful paradigm to handle black-box simulation models because it enables the construction of empirical distributions of the cost function *off-line* that can be updated *on-line* along the search. This avoids the need to perform simulations on-line, hence, it dramatically decreases computational times and

ultimately enables feasible industrial implementations.

The paper is structured as follows. In Section 2 we present a high-level description of power plant operations and effects of weather and water intake constraints. In Section 3 we present a stochastic optimization formulation and the proposed solution strategy. In Section 4 we describe the weather forecasting system WRF and the procedure used to compute physical uncertainty representations in the form of ensembles. In Section 5 we present the integrative case study. In Section 6 we provide concluding remarks and briefly discuss future work.

2 Power Plant Operations

In this section we present a high-level description of cooling-constrained power plants and explain the effect of weather conditions and water constraints on generation capacity.

A typical interface between a typical power plant equipped with a wet recirculating cooling system is sketched in Figure 1. For a detailed description of power plant operations and other types of cooling systems we refer the reader to [3]. The power plant burns fuel to generate steam that is used to generate power in a turbine train. The steam coming out of the turbines must be condensed before being recirculated to the power plant. This steam has a thermal load at time t denoted by $Q_p(t)$.

The thermal load depends on the total amount of power produced by the plant $P(t)$, which implicitly sets the amount of steam to be condensed. We will denote as $x_p(t)$ the plant output variables of interest in this case given by $P(t)$ and $Q_p(t)$. The operational degrees of freedom of the power plant will be denoted as $u_p(t)$. This can be the set-points, fuel flow rate, stoichiometric ratios, and so on. We will assume that the plant model does not exhibit coupling in time (i.e., it is at steady-state). Consequently, we will remove the explicit dependence in time from the nomenclature. We have a plant model in the following abstract form:

$$0 = f_p(x_p, u_p). \tag{2.1}$$

The cooling system capacity is given by Q_c. This is part of the cooling system states denoted by x_c. The cooling system capacity depends on the weather conditions (relative humidity and dry bulb temperature) denoted as w. The operational degrees of freedom of the cooling system are denoted as u_c. The cooling system model has the following form:

$$0 = f_c(x_c, u_c, w). \tag{2.2}$$

At steady-state, the cooling system capacity should match that required by the power plant. This condition is given by the abstract condenser model of the form

$$0 = g(x_p, x_c). \tag{2.3}$$

The capacity of the cooling system can be denoted in implicit form as $Q_c(u_c, w)$. The maximum capacity by design under *ideal* weather conditions will be defined as Q_c^{max}. This maximum capacity is constrained by physical limitations of the system such as diameter, height, flooding conditions, and pumping capacity, by environmental regulations, and by drought. These limitations typically

manifest as *water constraints*, which limit the amount of fresh water (make-up water) that the cooling system can use to provide cooling capacity. For any other *nonideal* weather conditions we have $Q_c(u_c, w) \leq Q_c^{max}$. The feasible operating region for the cooling system is given by the set of inequality constraints,

$$0 \leq h_c(x_c, u_c, w), \tag{2.4}$$

which depend on the weather conditions and design specifications. This set of constraints can also include bounds on states and decision variables for the cooling system. We denote the feasible operating region for the power plant as an inequality system of the form

$$0 \leq h_p(x_p, u_p, w). \tag{2.5}$$

This set of constraints includes bounds on states and decision variables for the plant. The entire

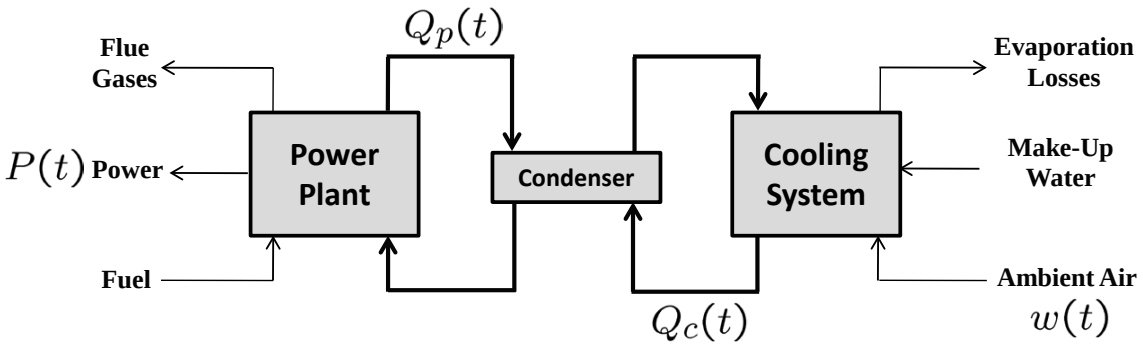

Figure 1: Schematic representation of the interface between the generation and cooling systems.

model coupling the power plant and the cooling system can be written as follows:

$$0 = f(x, u, w) \tag{2.6a}$$

$$0 \leq h(x, u, w). \tag{2.6b}$$

For the power plant, the maximum power output under design specifications (nominal capacity) is given by P^{max} with corresponding cooling demand Q_p^{max}. If the available cooling capacity satisfies $Q_c(u_c, w) \geq Q_p^{max}$, then the power output is constrained only by the plant-side design specifications and is given by P^{max}. If the cooling capacity is constrained $Q_c(u_c, w) \leq Q_p^{max}$, then $P^{max} = P^{max}(Q_c(u_c, w)$ and, in order to have feasible operation, we need $P \leq P^{max}(Q_c(u_c, w))$ and $Q \leq Q_p^{max}(Q_c(u_c, w))$. In other words, the maximum capacity of the power plant is an implicit function of the available cooling capacity. In this case, the power plant is operationally constrained by the prevailing weather conditions because the cooling capacity is constrained in turn by weather and design/regulation specifications. Therefore, there exists a feasibility threshold value Q_p^{max} where the system moves from being design (physically) constrained to cooling constrained. We note that the relative influence of the weather conditions depends on the design of the plant. In particular, it depends on the distance of the optimal cooling capacity $Q_c^*(w)$ from the feasibility threshold Q_p^{max}.

The thresholds for two different plant designs $Q_p^{max}, \bar{Q}_p^{max}$ with maximum power outputs $P^{max}, \bar{P}^{max}$ are sketched in Figure 2. The sensitivity of the power plant output to weather below the threshold is related to the design conditions of the cooling and condenser systems. For instance, the cooling system for $\bar{P}^{max}$ is less sensitive to the weather conditions than is the cooling system for $\bar{Q}_p^{max}$. This fact is of importance because it implies that a lack of sensitivity of the power plant output to weather conditions can occur if the plant is design constrained instead of cooling constrained. An interesting question is how to design or retrofit, in a cost-optimal manner, the power plant and cooling system in order to make the threshold Q_p^{max} as high as possible and to reduce sensitivity to weather conditions. Doing so would require robust optimization and flexibility analysis techniques. In this study, we will consider fixed design conditions for the power plant and cooling system and consider constraints only on operating conditions. We also note that weather also affects the boiler efficiency, since ambient air is used in the combustion chamber. However, we do not consider this case here.

In this study, we will assume that the main operational objective in the power plant is to maximize plant output power at prevailing weather conditions. To achieve this objective, one can adjust the set-points represented by u. For known weather conditions, the optimal set-points can be determined by solving a nonlinear optimization problem of the form

$$\max_{u} \quad P(x, u, w) \tag{2.7a}$$

$$\text{s.t.} \quad 0 = f(x, u, w) \tag{2.7b}$$

$$0 \leq h(x, u, w). \tag{2.7c}$$

We note that the set-points on the plant side can be manipulated to minimize the cooling demand of the plant in order to avoid the combined effect of weather and water constraints. In other words, if the set-points are not adapted to the prevailing weather conditions, the power plant output will be constrained by the water intake constraints, and the power output will decrease.

The steady-state modeling assumption implies that the distance between time steps is sufficiently long such that there exists a low-level control system capable of making the dynamic transition. This assumption is justified for slow base-load plants with time constants on the order of hours.

3 Stochastic Optimization

In this section we present a stochastic optimization formulation and solution strategy to mitigate effects of weather on power plant capacity.

3.1 Formulation

The objective of the stochastic formulation is to find set-points for the current time t, denoted by u_{HN}, that maximize the expected value of the plant capacity. We assume that the weather conditions w are

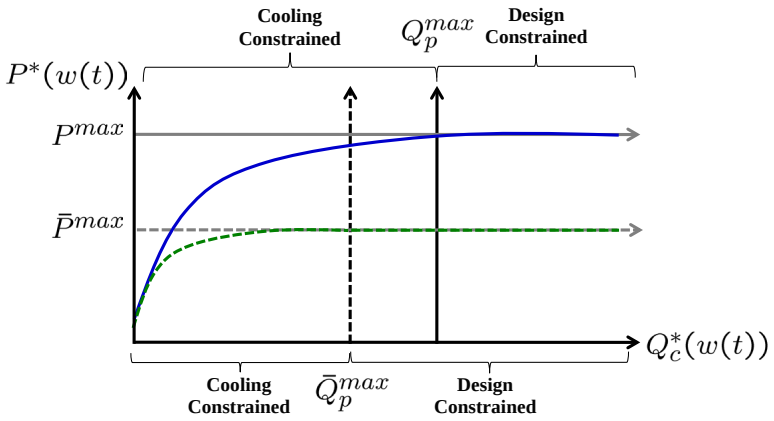

Figure 2: Effect of weather on maximum power plant capacity.

random variables.

$$\max_{u} \quad \mathbb{E}_{w}[P(x(w), u, w)] \tag{3.8a}$$

$$\text{s.t.} \quad 0 = f(x(w), u, w), \ w \in \Omega \tag{3.8b}$$

$$0 \leq h(x(w), u, w), \ w \in \Omega. \tag{3.8c}$$

Here, $\mathbb{E}_w[\cdot]$ denotes the expected value with respect to w. In this formulation the constraints are assumed to hold almost surely. We denote the optimal value of this problem as $\varphi(u_{HN})$, where the subscript HN indicates the *here-and-now* nature of the solution u_{HN}. In other words, a single (scenario independent) set of implementable set-points is sought. In practical terms, this set-point ensures the feasibility of the model and constraints for each realization of the weather conditions $w \in \Omega$. We note that the state variables are scenario-dependent and thus represented as $x(w)$. We also note that since the model is assumed to be in steady-state, the problems at different time instants are decoupled. Consequently, we can solve the above problem over an horizon $t, ..., t+T$ to determine the cost functions $\varphi(u_{HN}(\tau)), \quad \tau = t, ..., t+T$

Related to the here-and-now optimization problem (3.8) is the deterministic problem:

$$\max_{u} \quad P(x, u, \mathbb{E}[w]) \tag{3.9a}$$

$$\text{s.t.} \quad 0 = f(x, u, \mathbb{E}[w]) \tag{3.9b}$$

$$0 \leq h(x, u, \mathbb{E}[w]). \tag{3.9c}$$

This problem uses the expected value of the weather conditions $\mathbb{E}[w]$ to determine the optimal set-points, which we denote as u_D. If we fix the set-points u_D and evaluate the cost function for each scenario $w \in \Omega$, we can obtain $\varphi_D(u_D, w)$ and compute the deterministic cost function

$$\varphi(u_D) = \mathbb{E}_{w}[\varphi_D(u_D, w)]. \tag{3.10}$$

From [24] we have that

$$\varphi(u_{HN}) \geq \varphi(u_D). \tag{3.11}$$

Consequently, the stochastic optimization solution will always give better performance than will the deterministic counterpart (as long as the global maximum can be computed). The improvement (normally called the *value of the stochastic solution*), however, is problem dependent and thus can be determined only empirically. In the power plant context this value is determined by the design specifications and the ambient conditions at the particular climate and season. In this work, we seek to determine this value for a realistic setting.

We also seek to determine the effect of using weather forecasting to optimize operating conditions in real time (intra day). To this end, we consider the solution of the deterministic problem using *day-average* weather conditions, which we represent by $\bar{w}$. This gives a single fixed set-point for the entire time horizon $t, ..., t+T$ denoted by u_S and associated cost functions $\varphi(u_S(\tau)), \quad \tau = t, ..., t+T$. The cost improvement with respect to the deterministic cost $\varphi(u_D)$ gives a measure of the impact of using weather forecast information to perform intra day optimization of operating conditions. This is important because, in a typical market setting, it is necessary to determine the policy of set-points over the entire horizon $t, ..., t+T$ in advance (normally within 24 hours) in order to determine a plant capacity profile to be bid into the market. In other words, inaccurate estimation of weather conditions can ultimately affect plant profit as well as grid reliability and market prices because it can lead to under- and overestimation of capacity.

3.2 Solution

We can write the stochastic optimization problem (3.8) in the following abstract form:

$$\min_{u \in \mathcal{U}} \; V(u) := \mathop{\mathbb{E}}_{w} \left[\varphi(u, w) \right], \tag{3.12}$$

Here, $\mathcal{U}$ is the box set $\mathcal{U} := \{ u \,|\, \underline{u} \leq u \leq \overline{u} \}$, where $\underline{u}$ and $\overline{u}$ are lower and upper bounds, respectively. We can use this representation if we assume inequality constraints only in u and if we assume that the implicit solution of the black-box model $f(x, u, w) = 0$ can be represented as a *smooth* function of u and w, $x(u, w)$.

The random vector w has a *prior* distribution $\underline{\mathcal{P}}(w)$. We note that the cost $\varphi(u, w)$ also has an associated posterior distribution $\overline{\mathcal{P}}(\varphi(u, w))$ from which we seek to minimize its expected value. The KKT conditions of this problem are

$$0 = \nabla_u V(u) + \overline{\nu} - \underline{\nu} \tag{3.13a}$$

$$0 \leq \overline{\nu} \perp (\overline{u} - u) \geq 0 \tag{3.13b}$$

$$0 \leq \underline{\nu} \perp (u - \underline{u}) \geq 0. \tag{3.13c}$$

Here, $\overline{\nu}$ and $\underline{nu}$ are multipliers for upper and lower bounds, respectively. To solve the stochastic optimization problem, one can follow exterior or interior sampling approaches. In exterior sampling approaches (also known as sample-average approximation approaches) N_{sam} samples are drawn from the *prior* distribution $\underline{\mathcal{P}}(w)$ to convert the problem (3.14) into a deterministic counterpart [18] of the form

$$\min_{u \in \mathcal{U}} \; \tilde{V}(u) := \frac{1}{N_{sam}} \sum_{j=1}^{N_{sam}} \left[\varphi(u, w_j) \right]. \tag{3.14}$$

Inference is performed a posteriori to check the quality of the solution because the cost $\tilde{V}(u)$ is an empirical approximation of $V(u)$ [15]. One of the issues with the exterior sampling approach is that the number of samples normally must be overestimated in order to guarantee an appropriate solution accuracy. This is an issue of concern since, in many applications, only a few samples are available and because the complexity of the deterministic counterpart increases with the number of samples. For instance, sampling of the prior distribution might require the solution of a computationally intensive model (e.g., a weather forecasting model); or, if the prior is a high-dimensional normal, sampling will require the factorization of a large and dense covariance matrix (cubic complexity). In [7], the authors proposed a reweighting scheme to estimate additional samples from the limited number of samples obtained from a weather forecasting system.

Interior sampling approaches estimate the cost and derivative information internally *along the search*. One of the key advantages is that flexibility can be used to minimize the number of samples needed by identifying promising solutions early in the search. This concept was implemented by Sahin and Diwekar in the BONUS solver [19]. Here, the authors propose a reweighting scheme to estimate *directly* the *posterior* distribution of the cost $\overline{\mathcal{P}}(\varphi(u,w))$ and associated expected value and derivative information. This is different from the reweighting scheme presented in [7], which is applied to the *prior* distribution of the random variables.

Reweighting can also be interpreted as an importance sampling approach [11]. Using this observation, one can estimate the statistical properties of the cost posterior indirectly by updating *design* distributions of the random and input variables and the cost distribution generated *off-line* (before running the optimization search). This is particularly useful in a simulation-based approach where the evaluation of the cost function requires the simulation of a black-box model. The approach contrasts with traditional sampling algorithms that simulate the model along the search to estimate the expected cost and gradient information. Consequently, the internal reweighting approach can save large amounts of *on-line* computational time. In addition, reweighting enables the use of highly complex models (e.g., using internal convergence loops) and black-box legacy codes. Further, the use of reweighting enable prior distributions to be created using samples from another distribution that is more expensive to sample. Because of these reasons, we use the internal reweighting approach implemented in the BONUS solver [19].

A detailed explanation of the procedure can be found in [19]. The procedure involves the following general steps:

1. *Off-Line Computations.* Draw independently distributed samples $j = 1, ..., N_{sam}$ for random variables $\hat{w}_j$ and inputs $\hat{u}_j$ covering the convex set $\mathcal{U}$. Use these samples to generate the *design* prior density $\underline{\mathcal{P}}^d(u,w)$ using kernel density estimation (KDE). Evaluate the cost function for each sample $j = 1, ..., N_{sam}$ to obtain the cost function response $\varphi(\hat{u}_j, \hat{w}_j)$, and estimate the *design* posterior distribution $\overline{\mathcal{P}}^d(\varphi(u,w))$ using KDE.

2. *On-Line Computations*. Given the prior distribution for the random variables $\underline{\mathcal{P}}(w)$ and the design distributions $\underline{\mathcal{P}}^d(u,w)$, $\overline{\mathcal{P}}^d(\varphi(u,w))$ and cost values $\varphi(\hat{u}_j, \hat{w}_j),\ \ j = 1, ..., N_{sam}$, initialize decision variables u^k.

- 2.1.*Optimality Check.* At iteration k, and using the current iterate of decision variables u^k, define a narrow normal distribution around this point, and draw samples u_j^k from it. Draw samples w_j^k from the available distribution. Use samples to generate $\underline{\mathcal{P}}^k(u,w)$ using KDE. Estimate the cost function using the reweighting formula:

$$\tilde{V}(u^k) \approx \mathbb{E}_w[\varphi(u^k,w)] \approx \sum_{j=1}^{N_{sam}} \omega_j^k \varphi(\hat{u}_j,\hat{w}_j), \tag{3.15}$$

where the weights ω_j^k are obtained from

$$\omega_j^k = \frac{\frac{\underline{\mathcal{P}}^k(u_j^k,w_j^k)}{\underline{\mathcal{P}}^d(\hat{u}_j,\hat{w}_j)}}{\sum_{i=1}^{N_{sam}} \frac{\underline{\mathcal{P}}^k(u_i^k,w_i^k)}{\underline{\mathcal{P}}^d(\hat{u}_i,\hat{w}_i)}} \tag{3.16}$$

and satisfy $\sum_{j=1}^k \omega_j^k = 1$. Perturb the decision variables $u^k + \delta u^k$ to estimate the perturbed cost $\tilde{V}(u^k+\delta u^k)$ using reweighting to estimate the gradient $\nabla_u \tilde{V}(u^k)$. If the KKT residual $\|\nabla_u \tilde{V}(u) + \overline{\nu} - \underline{\nu}\|$ and complementarity products are sufficiently small, terminate. Otherwise, go to step 2.2.

- 2.2. *SQP Step Computation.* Use the gradient to compute the Hessian approximation H_k using the BFGS formula, and compute step Δu for decision variables by solving the quadratic program (QP):

$$\min_{\Delta u} \quad \nabla_u \tilde{V}(u^k)^T \Delta u + \Delta u^T H^k \Delta u, \quad \text{s.t.} \quad u^k + \Delta u \in \mathcal{U}. \tag{3.17}$$

Cut the step if necessary to obtain a new iterate $u^{k+1} = u^k + \alpha \Delta u$ with $\alpha \in (0,1]$. Go to step 2.1.

As can be seen, the computation of the cost at each iteration requires only the off-line values $\varphi(\hat{u}_j,\hat{w}_j)$; the black-box model is never run along the search. This is a key feature in real-time applications where the stochastic optimization problem must be solved at different instances in time. If higher moments are needed to estimate, for instance, variance of the cost then the design posterior distribution can be used. We also note that reweighting can be used to compute confidence intervals for the gradient and thus have a more robust termination criterion. This comes, however, at a much higher computational cost because *full* problems must be solved with different realizations of uncertainty in order to compute the variance of the cost, at solutions satisfying the KKT conditions [15, 12]. In our approach, we terminate when the norm of the KKT residual is sufficiently small. This procedure will converge as long as the QP step obtained with the estimate $\nabla_u \tilde{V}(u^k)$ provides a direction of sufficient decrease on the actual cost $V(u^k)$. This is normally observed in practice [4], and the chances are enhanced if sampling is performed along the search, as we do here [16]. There is no general guarantee, however, that the solution is obtained unless an explicit inference task is performed. In our case, we also perform a cheap a posteriori validation by computing the value of the stochastic solution to determine whether the procedure is improving the cost over the deterministic solution.

One of the limitations of this approach is that complexity increases with the number of degrees of freedom u and random variables w because the number of samples needed to cover the set $\mathcal{U}$ and

the uncertain space increases. Convergence is normally exponential in the number of samples. Note, however, that this complexity is reflected in off-line computations only and can be reduced in practice by using variance reduction techniques such as Hammersley sequence sampling or Latin hypercube sampling [9]. The number of samples needed to estimate the distributions along the search by using KDE is problem dependent. We also note that the complexity of the QP increases cubically with the number of degrees of freedom since the QP it is nearly dense.

4 Weather Forecasting and Uncertainty Quantification

We describe the procedures used to compute real-time forecasts and uncertainty information for ambient temperature and humidity using the numerical weather prediction (NWP) model WRF. We present details of the ensemble initialization and restarting procedures capable of providing frequent forecast updates as needed in real-time power plant operations.

The WRF model [23] is a state-of-the-art numerical weather prediction system designed to serve both operational forecasting and atmospheric research needs. WRF is the result of a multi-agency and university effort to build a highly parallelizable code that can run across scales ranging from large-eddy to global simulations. WRF has a comprehensive description of the atmospheric physics that includes cloud parameterization, land-surface models, atmosphere-ocean coupling, and broad radiation models. The terrain resolution can go up to 30 seconds of a degree (less than 1 km^2). The discretization of the underlying nonlinear partial-differential equation system over a 3-D field typically leads to a state dimensionality of $O(10^8 - 10^{10})$. Consequently, the WRF model is computationally intensive, and special domain targeting techniques and computational resources are needed to enable computational feasibility [7]. We currently run an implementation of the WRF system at Argonne National Laboratory. This implementation can be used to generate validated weather data and uncertainty information for energy optimization studies [27, 26, 7].

To initialize the NWP simulations, we use reanalyzed fields, that is, simulated atmospheric states reconciled with observations (the entire atmospheric state space is required by the model) as initial conditions, since only a small subset of the state space is available through measurement at any given time [13]. The reanalyzed states are generated from an internal data assimilation procedure. In particular, we use the North American Regional Reanalysis (NARR) data set that covers the North American continent (160W-20W; 10N-80N) with a resolution of 10 minutes of a degree, 29 pressure levels (1000-100 hPa, excluding the surface), every three hours from 1979 until present. We use an ensemble of realizations to represent uncertainty in the initial (random) wind field and propagate it through the WRF nonlinear model. The initial ensemble is obtained by sampling from an empirical distribution, a procedure similar to the National Centers for Environmental Prediction (NCEP) method [14]. In the following subsections we describe in more detail the procedures needed for generating the forecast and its uncertainty.

4.1 Ensemble Initialization

In a normal operational mode, the NWP system evolves a given state from an initial time t_0 (current time) to a final time t_F. The initial state is produced from past simulations and reanalysis fields. Because of observation sparseness in the atmospheric field and the incomplete numerical representation of its dynamics, the initial states are not known exactly and can be represented only statistically. Therefore, we use a distribution of the initial conditions to describe the confidence in the knowledge of the initial state of the atmosphere. We assume a normal distribution of the uncertainty field of the initial state, a typical assumption in weather forecasting. The distribution is centered on the NARR field at the initial time, the most accurate information available. In other words, the expectation is exactly the NARR solution. The second statistical moment of the distribution described by the covariance matrix $\mathbf{V}$ is approximated by the sample variance or pointwise uncertainty and its correlation, $\mathbf{C}$. The initial N_S-member ensemble field $x_s^{t_0} := x_s(t_0)$, $i \in \{1 \dots N_S\}$, is sampled from $\mathcal{N}\left(x_{\text{NARR}}, \mathbf{V}\right)$:

$$x_s^{t_0} = x_{\text{NARR}} + \mathbf{V}^{\frac{1}{2}} \xi_s \,,\ \xi_s \sim \mathcal{N}(0, I) \,,\ s \in \{1 \dots N_S\} \,, \tag{4.18}$$

where $\mathbf{C} = \mathbf{V}_{ij} / \sqrt{\mathbf{V}_{ii} \mathbf{V}_{jj}}$ and $\mathbf{V}_{ii}$ is the variance of variable i. This is equivalent to perturbing the NARR field with $\mathcal{N}\left(0, \mathbf{V}\right)$. That is, $x_s = x_{\text{NARR}} + \mathcal{N}\left(0, \mathbf{V}\right)$. We note that the correlation matrix is huge (square of the discretized number of states in a 3-D field) and thus cannot be formed and stored. In what follows, we describe the procedure used to estimate the state correlation matrix.

4.2 Estimation of the Correlation Matrix

In weather models the correlation structure typically is localized in space. Therefore, in creating the initial ensemble one needs to estimate the spatial scales associated with each variable. To obtain these spatial scales, we build correlation matrices of the forecast errors using the WRF model. These forecast errors are estimated by using the NCEP method [14], which is based on starting several simulations staggered in time in such a way that, at any time, two forecasts are available. In particular, we run a month of day-long simulations started every twelve hours so that every twelve hours we have two forecasts, one started one day before and one started half-a-day before. The differences between two staggered simulations is denoted as $d_{ij} \in \mathbb{R}^{N \times (2 \times 30\text{days})}$, that is, the difference at the $i\text{th}$ point in space between the $j\text{th}$ pair of forecasts, where N is the number of points in space multiplied by the number of variables of interest. We can then define ϵ_s as the $i\text{th}$ row, each of which correspond to the deviations for a single point in space. Therefore, the covariance matrix can be approximated by $\mathbf{V} \approx \mathbf{d}\mathbf{d}^{\mathbf{T}}$. Calculating and storing the entire covariance matrix are computationally intractable. Consequently, we describe the correlation distance at each vertical level and for each variable by two parameters representing the East-West and North-South directions. This approach captures the Coriolis effect and the Earth rotation, as well as faster and larger-scale winds in the upper atmosphere. We assume that correlations and winds are roughly similar in nature across the continental U.S. This process is repeated in the vertical direction. To create the perturbations from these length scales, we take a normally distributed noisy field and apply Gaussian filters in each direction with appropriate length scales to obtain the same effect as in (4.18).

4.3 Ensemble Propagation through the WRF Model

The initial state distribution is evolved through the NWP model dynamics. The resulting trajectories can then be assembled to obtain an approximation of the forecast covariance matrix:

$$x_s^{t_F} = \mathcal{M}_{t_0 \to t_F}\left(x_s^{t_0}\right) + \eta_s(t)\,,\ s \in \{1 \dots N_S\}\,, \tag{4.19}$$

where $x_s^{t_0} \sim \mathcal{N}\left(x_{\text{NARR}}, \mathbf{V}^{t_0}\right)$, $\eta_s \sim \mathcal{N}\left(0, \mathbf{Q}\right)$, and $\mathcal{M}_{t_0 \to t_F}(\bullet)$ represents the evolution of the initial condition through the WRF model from time t_0 to time t_F. The initial condition is perturbed by the additive noise η that accounts for the various error sources during the model evolution. An analysis of the covariance propagation through the model is given in [7].

In this study, we assume that the numerical model (WRF) is perfect, that is, $\eta \equiv 0$, and thus given the exact real initial conditions, the model produces error-free forecasts. For long prediction windows, this is a strong assumption. In this study, however, we restrict the forecast windows to no longer than one day ahead, thus making this assumption reasonable.

Relative humidity (RH) is typically a derived quantity or a diagnostic variable in the NWP model, which means it is not modeled directly. RH is computed based on air density, temperature, and vapor pressure [23]. In particular, the water mixing ratio is available as a standard WRF output; then the RH is obtained as the ratio between the water mixing ratio and the saturation mixing ratios, which are based on local temperature and pressure.

5 Integration Study

We present a numerical study integrating stochastic optimization and weather uncertainty quantification using detailed physics models for weather and power plant simulation. The framework is sketched in Figure 3. We first demonstrate that stochastic optimization can yield significant improvements over deterministic operating policies. In addition, we discuss computational feasibility of the proposed framework.

5.1 Plant Description

A general scheme of a pulverized coal power plant is shown in Figure 4. A more detailed description can be found in [22, 21, 20] and the references therein. Powdered coal is blown into the boiler combustion chamber, and the hot gases and heat produced by the combustion of coal are used to produce steam within the tubes inside the boiler. High pressure steam is fed into a train of high pressure (HP), intermediate pressure (IP) and low pressure (LP) turbines where it does mechanical work causing the turbine shaft to rotate at high speed. The steam returns to the boiler after the first pass through HP turbines, where it is reheated before being fed to the IP turbines. After passing through the turbine train, the steam is cooled to condensation and returned to the boiler while being preheated with steam extracted from the turbines in eight stages. A flue gas desulfurization (FGD) unit is attached to the boiler effluent to remove the SO_2 before it is released to the atmosphere. This FGD unit employs limestone and oxygen in a reactive absorber to react with SO_2 and produce calcium

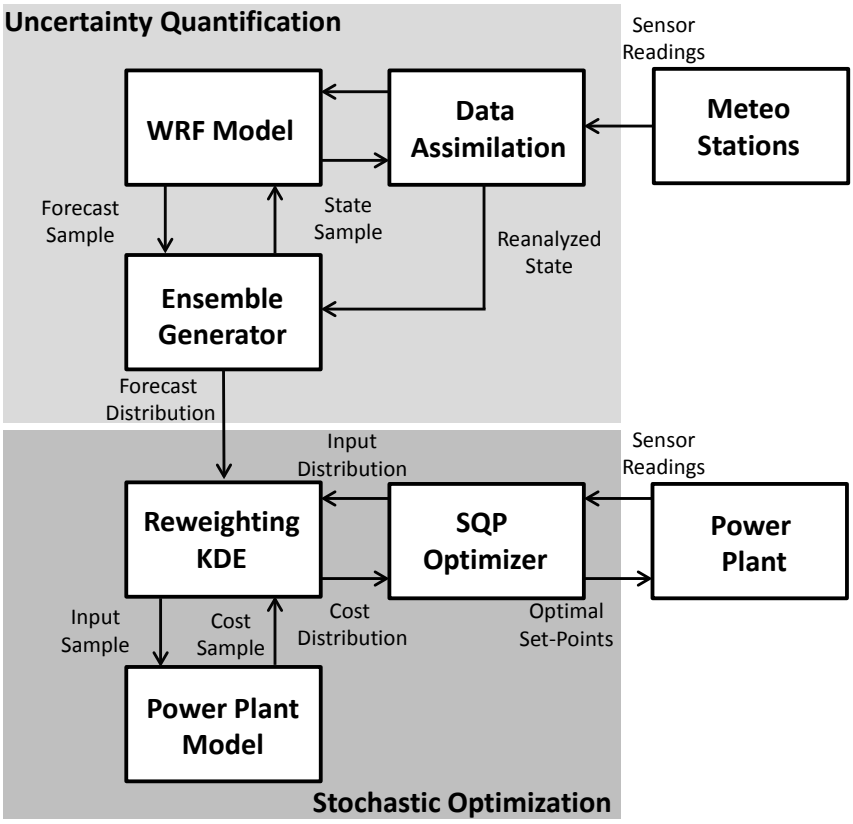

Figure 3: Integration framework of weather uncertainty quantification.

sulfate (gypsum). Limestone is fed to the unit as slurry, and the resulting sulfate is removed from the absorber as suspension water. In order to compensate for losses by evaporative cooling and water employed in the flue gas treatment units (ash removal and flue gas desulfurization) large makeup streams are required in the cooling cycle. As we have discussed, the amount of make-up water can become constrained under different scenarios, thus limiting capacity.

The model is a supercritical steady-state flow sheet without carbon capture designed to generate 700 MW of electricity. The process performance parameters (generation, efficiency, emission and water consumption) are calculated with the process simulator model Aspen Plus which allows the use of available unit operation models, property data bases and simulation algorithms. The specifications of the Aspen model are summarized in Table 1. A sequential modular approach was used to simulate the model.

Accurate estimation of the water intake strongly depends on the cooling tower model used. In this work, an equilibrium-based model for the cooling tower was used to estimate the evaporation rates [22]. The model comprises two flash units and one heat exchanger. One of the flashes is set to be adiabatic and is employed to calculate the wet bulb temperature from dry bulb temperature and relative humidity. The second flash is coupled with the heat exchanger and has no liquid outlet. A design specification ensures that heat provided to the flash equals the condenser cooling requirement. Additionally, a set of design specification and calculators determine the cold water temperature,

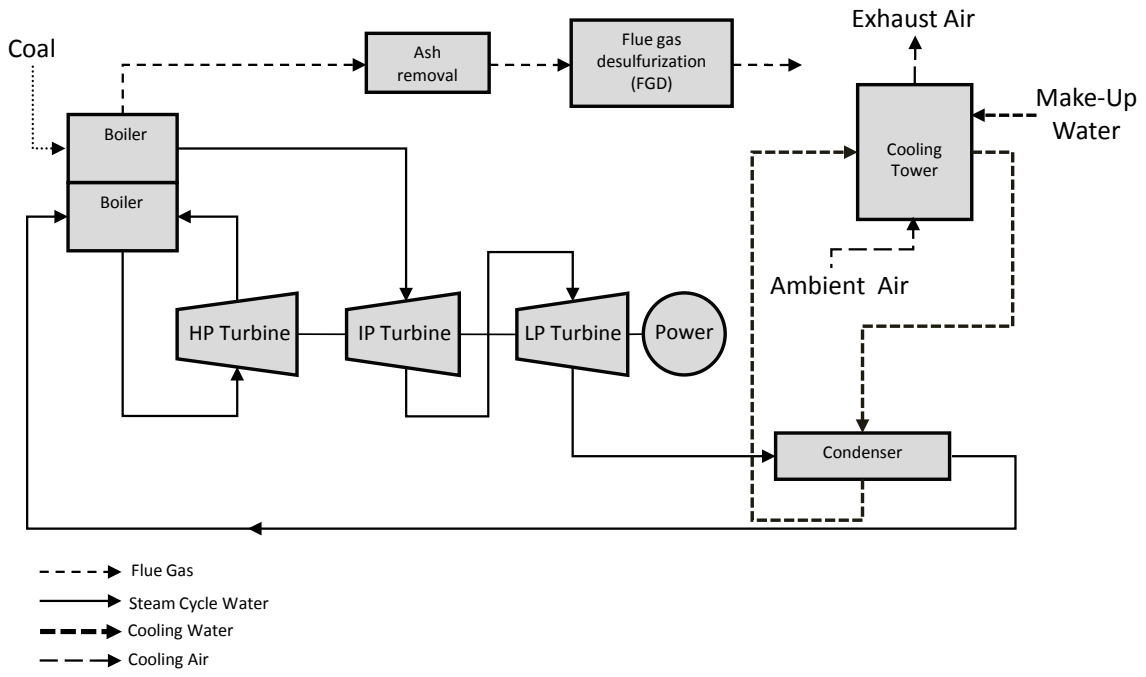

Figure 4: Schematic representation of coal-fired plant.

circulating water flow rate, and air flow rate for a constant volume, forced-drift cooling tower. We enforced a constraint on the maximum water intake of 2×10^6 lb/hr. This represents the nominal consumption under average atmospheric conditions: of 59 ^{o}F and 60% relative humidity.

To determine the plant operational and design parameters that influence water consumption the most and therefore the maximum generation, we performed a partial rank correlation coefficients (PRCC) analysis. Re-heater temperature, air excess to the boiler, water contents to for the FGD slurry and pressure drops at HP turbines 1 and 2 were chosen as decision variables.

5.2 Validation of Weather Forecast

In Figure 5 we present average and 30 ensemble profiles for the dry bulb temperature and relative humidity for a day, respectively, in a random location in the Midwest region of the U.S. for June 1, 2006. The day-long profile is forecast at the location of a meteorological tower located near Chicago, IL ($41°42'04''$, $87°59'42''$). The crossed dots show the actual temperatures and humidity observed at the closest meteorological station as measured by the instruments mounted on the weather tower. The envelope surrounding the ensembles is the 95% confidence region. Time zero represents 6 a.m. central time.

One can see that the projection for a day ahead is reasonably accurate from the WRF model, capturing well the trend and sensor readings. Temperature varies significantly throughout the day in the range of 15-26^{o}C, while the relative humidity varies in the range of 20-90%. Humidity is much higher at night, whereas temperature is much higher during the day, the variables are anticorrelated. The temperature and RH uncertainties are the result of different NWP model forecasts for the same

Table 1: Aspen model specifications.

Units	Number of Units
Sections	3
Boiler	1
Steam Cycle	1
Cooling Water	1
Total Streams	**178**
Mass Streams	120
Energy Streams	37
Work Streams	21
Total ASPEN Blocks	**128**
Unit Operations	115
Heat Mixers/Splitters	10
Work Mixers/Splitters	3
Calculator Blocks	25
Design Specifications	31
Convergence Blocks	31

prediction window. While the temperature is relatively well modeled, the RH is the result of the moisture transport in the atmosphere and complex physical interactions that lead to phase changes. Therefore, as expected, we observed a relatively wider uncertainty estimate for RH and a fanning effect toward the end of the forecast window. We also note that while uncertainty for temperature projections has been discussed in other studies, we were unable to find projected uncertainty estimates for RH in the literature. These estimates are a contribution of our work.

The uncertainty envelopes for the ambient temperature are narrow. The uncertainty envelope for the relative humidity, on the other hand, is wide and can be as large as 30-40% toward the end of the day. Both envelopes follow complex shapes due to the nonlinearity of the weather prediction model and current meteorological conditions. Toward the end of the day, the mean forecast does not predict well the shape of the humidity profile, but the uncertainty envelope covers this region reasonably well. This indicates the additional complexity in forecasting humidity. This also reflects the fact that the uncertainty envelope is unable to fully cover the sensor readings at all times.

We fitted the ambient temperature and relative humidity WRF ensembles to a log-normal and uniform distributions, respectively. These are illustrated in Figure 6 and 7. These were used to generate the design distribution $\underline{\mathcal{P}}^d(u, w)$ and corresponding cost posterior $\overline{\mathcal{P}}^d(\varphi(u, w))$.

5.3 Maximum Power Performance

To demonstrate the economic benefits of stochastic optimization over current practice, we performed extensive simulations using three settings. In the first setting, we solved the here-and-now stochastic

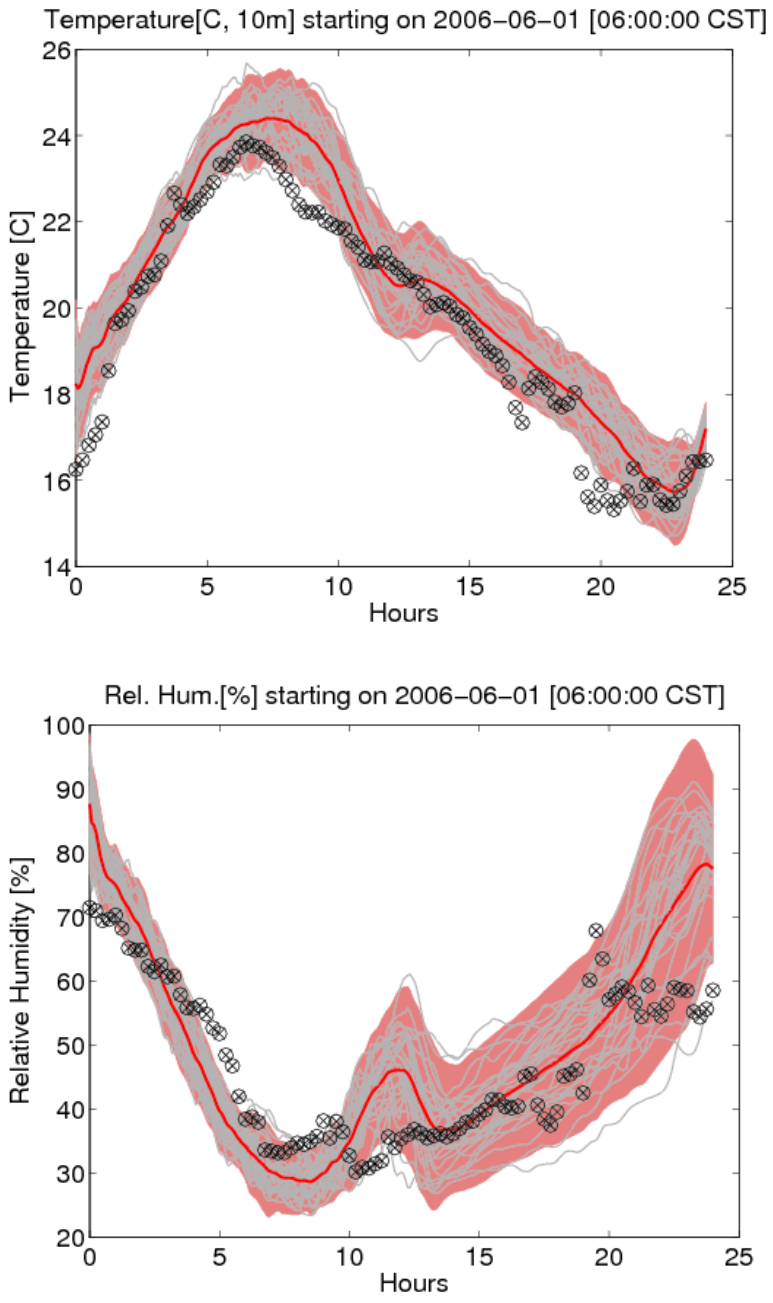

Figure 5: Forecast and ensemble profiles for dry bulb temperature and relative humidity in Midwest U.S. for June 1, 2006.

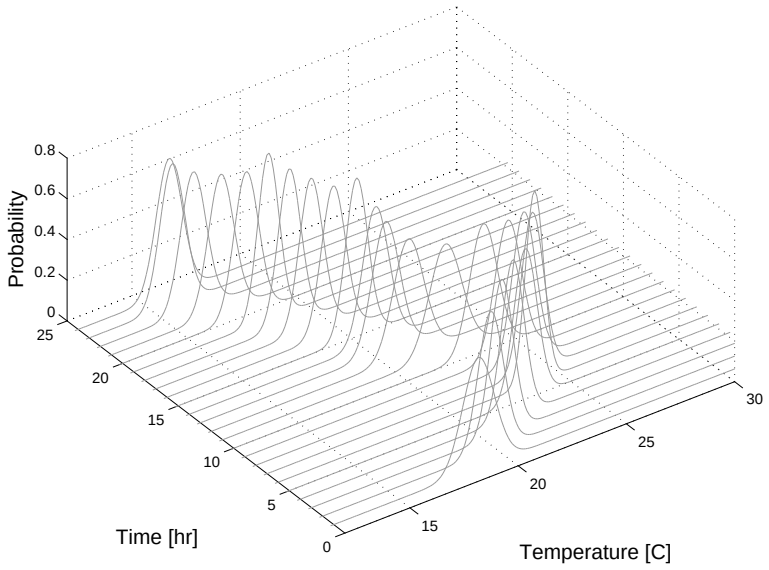

Figure 6: Fit of ambient temperature WRF ensembles to log normal distribution.

optimization problem to maximize the expected value of power during a timeframe of 24 hours with time steps of one hour. The WRF ensemble information was used to solve the problem. This profile is labeled as *"Stochastic"*. In the second setting, we solved the deterministic problem using the mean forecast obtained from WRF and we computed the expected value of the maximum power using the ensembles from WRF. The maximum power profile is labeled *"Deterministic b)"*. In the third setting, we compute day-average conditions for temperature and humidity and leave the operating conditions fixed throughout the day. The expected value of the maximum power was again computed from the WRF ensembles. The maximum power profile is labeled *"Deterministic a)"*. Figure 8 shows the result of optimization. We have the following findings:

- From Deterministic a) we can observe that the power profile does vary with the weather conditions throughout the day. Specifically, the power levels vary from 715 to 665 MW, which represents 7% of the base load of 700 MW. Consequently, the set-points must be adjusted throughout the day to mitigate the effects of weather and water constraints.

- High power is obtained at night when ambient temperature is lower, even if humidity is high. Low power is obtained around noon when temperature is high, even if humidity is low. This indicates that temperature dominates the cooling capacity. In addition, it indicates that the plant becomes constrained by cooling capacity at high temperatures since we have imposed a constraint on water intake.

- From Deterministic a) and b) we can see that adjusting operating conditions according to the mean forecast can increase the maximum power output. The total power output for a) is 16,675 MWh while for b) is 16,823 MWh, a performance gain of 146 MWh.

- Stochastic optimization gives a total power output of 16,922 MWh. This gives a performance

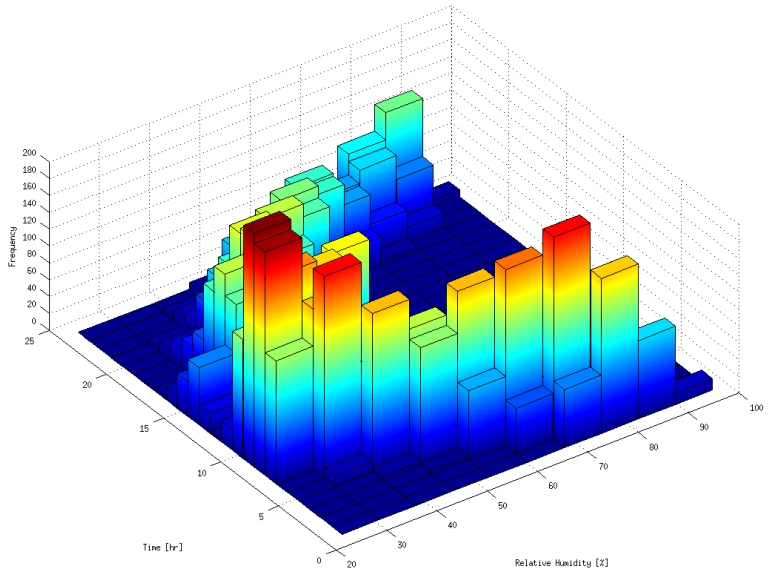

Figure 7: Fit of relative humidity WRF ensembles to uniform distribution.

gain of 99 MWh over Deterministic b) and 245 MWh over a).

- Stochastic optimization displaces the power output curve but does not fully mitigate the variability of power throughout the day. In particular, the variability for the stochastic profile is 45 MW. This indicates the strong impact of weather and corroborates the fact that the plant is constrained by cooling capacity.

- At hour 15 and 23 the stochastic profile is lower than that of Deterministic b). This is because of the estimation error introduced from the small number of samples. The profiles, however, indicate that the stochastic optimization solution is consistently better than the deterministic one, indicating that a small number of samples seem sufficient to obtain accurate solutions. This, in turn, indicates that the weather uncertainty information is accurate and the reweighting scheme is efficient in identifying good-quality solutions.

These results indicate that, for a given water intake constraint, power can be maximized by using stochastic optimization. This also implies that, as water constraints are tightened, it is possible to maintain high output levels with stochastic optimization, thus reducing water consumption. We also note that tightening the water intake constraint will increase the sensitivity of the power profile to weather conditions.

If we project the gain of 245 MWh over a total of 100 days in a year (summer conditions), this translates into 24.5 GWh that, if paid at a price of \$100/MWh, translate into \$2,450,000. Of course, a more rigorous analysis capturing seasonal effects and different constraint scenarios is needed. The presented analysis, however, provides an order of magnitude of the expected savings through stochastic optimization. We also note that the results confirm that both the weather uncertainty information and the proposed solution approach are sufficiently accurate to generate perceivable improvements over existing practice.

We highlight the potential of stochastic optimization in aiding *bidding* procedures. In particular, the proposed approach enables a more robust prediction of the expected performance of the plant. This is important in day-ahead bidding procedures where maximum capacity is an important parameter needed in market participation. The computed maximum capacity profiles obtained with stochastic optimization can implicitly minimize economic penalties arising in real-time markets. In particular, if capacity is overestimated, then the plant owner will have to pay for the unmet generation at a real-time price which tends to be higher since more more expensive peaking plants are used in real-time grid operations. On the other hand, if capacity is underestimated, there will be an opportunity cost if the available surplus cannot be sold in the real-time market.

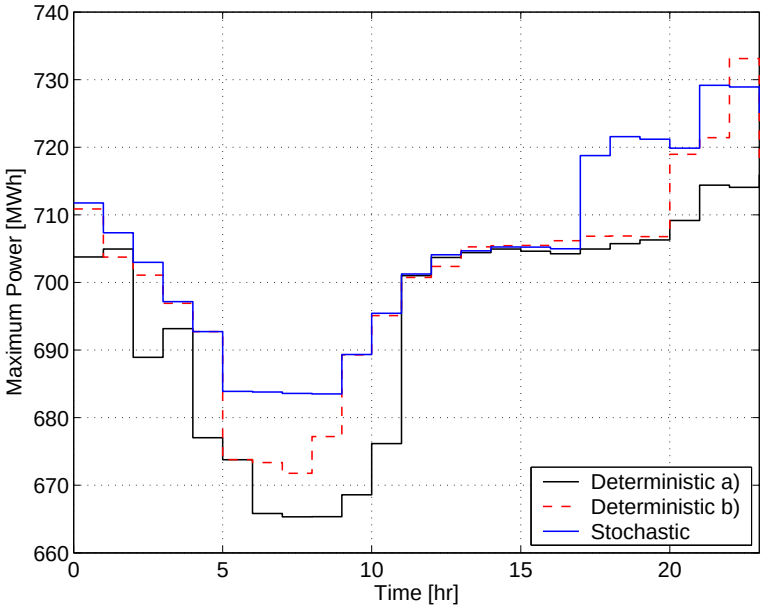

Figure 8: Maximum power profiles for stochastic and deterministic approaches.

5.4 Numerical Performance

A single simulation of the Aspen power plant model takes, on average, 10 minutes on a Intel Core 2 Duo processor running at 2 GHz. This time is required for all the loops to converge and satisfy the design specifications within the simulator. To set-up the initial distributions of the objective function, we ran 600 scenarios, each requiring a simulation of the Aspen model (10 min). This required as total time of about 100 hours. We emphasize that this is *off-line* time. To generate these scenarios, we sampled the uncertain variable domain (from uniform distributions) considering weather variations of one week. We also emphasize that we can reuse the design distributions to solve stochastic optimization problems over several days while we can update the design distributions for the next week as updated weather data is obtained.

The average number of SQP iterations was 7, with a minimum of 2 and a maximum of 16. Each SQP iteration takes less than a minute. We used 50 *different* samples at each SQP step. To verify that

global optimality was achieved, we ran the optimizer from 7 different starting points at each time step. A total of 1,184 SQP iterations was required for the 24-hours timeframe. Starting from multiple starting points also enhances the chances of identifying higher quality solutions [16].

A traditional stochastic optimization approach (with no reweighting) would have used 355,200 runs of the ASPEN model to solve the problems over the 24 hour timeframe. That in turn would have required 59,200 hours of computation (2,400 days). This total number of runs comprises 6 Aspen model simulations per SQP iteration (one for function evaluation and one for each degree of freedom to evaluate derivatives). The proposed solution approach with reweighting does not requires *on-line* model runs. The total on-line time for 1,184 SQP iterations is around 2 hours. This translates into 99.8% savings in computational load and enables feasibility.

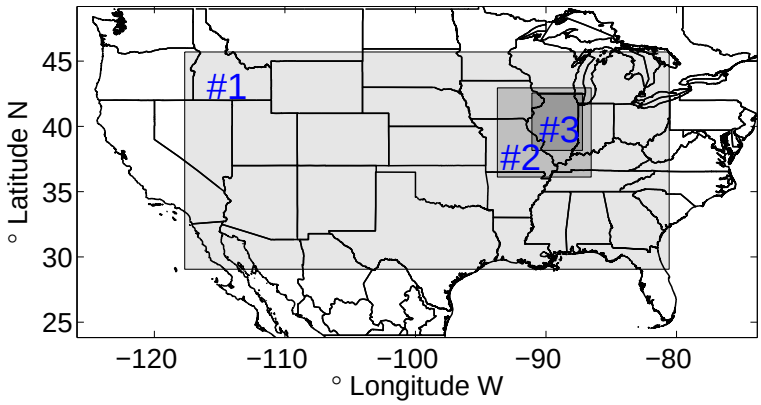

Figure 9: Simulation domain and resolution for WRF.

The strategy for estimating the uncertainty in weather forecasts is highly parallelizable because each scenario evolves independently through the numerical weather model WRF, once the initial scenario ensemble has been generated. The most expensive computational element is the evolution of each member through the WRF system. We consider a two-level parallel implementation scheme. The first level is a coarse-grain task division represented by each individual member. A secondary fine-grained level consists in the parallelization of each individual member. This approach is readily available because WRF is already parallelized, and this strategy yields a highly scalable solution.

The WRF simulations were carried out on Argonne National Laboratory's Fusion cluster. This cluster has 320 nodes each with two Nehalem 2.6 GHz Pentium Xeon and 36 GB of RAM. We considered three nested domains, the coarsest having a horizontal resolution of 32 by 32 km (#1), and the finest 2 by 2 km (#3), as illustrated in Figure 9. This targeted domain partition provides an accurate representation of the target region, which was selected to be the state of Illinois and dramatically decreases computational complexity. In particular, it avoids the use of fine *uniform* resolutions over the entire country, which is the approach used by most providers of forecast information.

For this study, we carried our NWP simulations asynchronously and on a small number of cores. We estimate, however, that we can compute members for a 24-hour simulation in *less than an hour* on

about 400 cores. Consequently, it is feasible to perform rigorous uncertainty quantification tasks with WRF to aid power plant operations.

6 Conclusions and Future Work

We propose a stochastic optimization framework to perform integrated water and energy management in cooling-constrained power plants. The approach integrates detailed physical weather and power plant simulation models to assess effects of ambient conditions on plant performance. We have found that the maximum plant capacity can vary in the range of 5-10% of the nominal capacity in intraday operations because of weather variations. In addition, we have found that stochastic optimization can lead to *daily* capacity gains of as much as 245 MWh over current practice and enables more robust bidding procedures. Moreover, we have demonstrated that a reweighting scheme enables computationally feasibility of the approach.

The results open interesting directions for future work. A more extensive study is needed of the effects of seasonal weather conditions on plant performance and on the accuracy of the weather forecasts throughout the year. This will require, however, extensive computational resources to perform integrated weather forecasting and stochastic optimization. It is also necessary to revisit the internal reweighting scheme to assess, on the fly, the quality of the solutions using inference techniques. It is also of interest to study the effects of weather on other plant configurations such as combined cycle power plants where ambient conditions affect combustion processes. Formulations using dynamic effects such as ramp constraints must also be developed. This will increase dramatically the complexity of the problem as time decoupling is no longer possible. Finally, it is of interest to study the effect of risk functions and long-term water reservoir dynamics (e.g., seasonal) on plant performance.

Acknowledgments

This work was supported by the U.S. Department of Energy, under Contract No. DE-AC02-06CH11357. We also thank Argonne National Laboratory's Laboratory Computing Resource Center for the use of the Fusion cluster.

References

[1] An analysis of the effects of drought conditions on electric power generation in the western united states. *Technical Report DOE/NETL-2009/1365, National Technology Energy Laboratory*, 2009.

[2] Impact of drought on U.S. steam electric power plant cooling water intakes and related water resource management issues. *Technical Report DOE/NETL-2009/1364, National Technology Energy Laboratory*, 2009.

[3] Cost and performance baseline for fossil energy plants. Volume 1: Bituminous coal and natural gas to electricity. *Technical Report DOE/NETL-2010/1397, National Technology Energy Laboratory*, 2010.

[4] R. H. Byrd, G. M. Chin, W. Neveitt, and J. Nocedal. On the use of stochastic hessian information in optimization methods for machine learning. *SIAM Journal on Optimization*, 21(3):977–995, 2011.

[5] C. C. Caroe and R. Schultz. A two-stage stochastic program for unit commitment under uncertainty in a hydro-thermal power system. Technical Report SC 98-11, ZIB, 1998.

[6] A. Cioaca, V. M. Zaval, and E. M. Constantinescu. Adjoint sensitivity analysis for numerical weather prediction: Applications to power grid optimization. *In Proceedings of International Workshop on High Performance Computing, Networking and Analytics for the Power Grid*, 2011.

[7] E. M. Constantinescu, V. M. Zavala, M. Rocklin, S. Lee, and M. Anitescu. A computational framework for uncertainty quantification and stochastic optimization in unit commitment with wind power generation. *IEEE Transactions on Power Systems*, 26:431–441, 2011.

[8] G. Conzelmann, V. Koritarov, L. Poch, P. Thimmapuram, and T. Veselka. Power system simulations of western united states region:. *Technical Report ANL/DIS-10-05, Argonne National Laboratory*, 2010.

[9] Urmila M. Diwekar and Jayant R. Kalagnanam. Efficient sampling technique for optimization under uncertainty. *AIChE Journal*, 43(2):440–447, 1997.

[10] T. J. Feeley, L. Green, J. T. Murphy, J. Hoffmann, and B. A. Carney. department of Energy/Office of fossil energy power plant water management R&D program. *Technical Report, National Energy Technology Laboratory*, 2005.

[11] Tim Hesterberg. Weighted average importance sampling and defensive mixture distributions. *Technometrics*, 37(2):185–194, 1995.

[12] J. Higle and S. Sen. Duality and statistical tests of optimality for two stage stochastic programs. *Mathematical Programming*, 75:257–275, 1996.

[13] E. Kalnay. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 2003.

[14] E. Kalnay, M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, et al. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, 77(3):437–471, 1996.

[15] J. Linderoth, A. Shapiro, and S. Wright. The empirical behavior of sampling methods for stochastic programming. *Annals of Operations Research*, 142(1):215–241, 2006.

[16] Yu. Nesterov and J.-Ph. Vial. Confidence level solutions for stochastic programming. *Automatica*, 44(6):1559 – 1568, 2008.

[17] C.P. Rodriguez and G.J. Anders. Bidding strategy design for different types of electric power market participants. *IEEE Transactions on Power Systems*, 19(2):964 – 971, 2004.

[18] A. Ruszczynski and A. Shapiro. *Stochastic Programming (Handbooks in Operations Research and Management Series)*. Elsevier Science BV, Amsterdam, 2003.

[19] K. Sahin and U. Diwekar. Better optimization of nonlinear uncertain systems (BONUS): A new algorithm for stochastic programming using reweighting through kernel density estimation. *Annals of Operations Research*, 132:47–68, 2004.

[20] Juan Salazar and Urmila Diwekar. An efficient stochastic optimization framework for studying the impact of seasonal variation on the minimum water consumption of pulverized coal (pc) power plants. *Energy Systems*, 2:263–279, 2011. 10.1007/s12667-011-0035-8.

[21] Juan M. Salazar, Urmila M. Diwekar, and Stephen E. Zitney. Stochastic simulation of pulverized coal (PC) processes. *Energy and Fuels*, 24(9):4961–4970, 2010.

[22] Juan M. Salazar, Stephen E. Zitney, and Urmila M. Diwekar. Minimization of water consumption under uncertainty for a pulverized coal power plant. *Environmental Science and Technology*, 45(10):4645–4651, 2011.

[23] W.C. Skamarock, J.B. Klemp, J. Dudhia, D.O. Gill, D.M. Barker, M.G. Duda, X.-Y. Huang, W. Wang, and J.G. Powers. A description of the Advanced Research WRF version 3. Technical Report Tech Notes-475+ STR, NCAR, 2008.

[24] S. Takriti, J. R. Birge, and E. Long. A stochastic model for the unit commitment problem. *IEEE Transactions on Power Systems*, 11:1497–1508, 1996.

[25] J. Wang, M. Shahidehpour, and Z. Li. Security-constrained unit commitment with volatile wind power generation. *IEEE Transactions on Power Systems*, 23:1319–1327, 2008.

[26] V. M. Zavala, E. M. Constantinescu, and M. Anitescu. Economic impacts of advanced weather forecasting on energy system operations. In *IEEE PES Conference on Innovative Smart Grid Technologies*, 2010.

[27] V. M. Zavala, E. M. Constantinescu, M. Anitescu, and T. Krause. On-line economic optimization of energy systems using weather forecast information. *Journal of Process Control*, 19:1725–1736, 2009.