Argonne NATIONAL LABORATORY

ANL/MCS-TM-309

Unit Commitment with Wind Power Generation: Integrating Wind Forecast Uncertainty and Stochastic Programming

Mathematics and Computer Science Division

About Argonne National Laboratory
Argonne is a U.S. Department of Energy laboratory managed by UChicago Argonne, LLC under contract DE-AC02-06CH11357. The Laboratory's main facility is outside Chicago, at 9700 South Cass Avenue, Argonne, Illinois 60439. For information about Argonne and its pioneering science and technology programs, see www.anl.gov.

Availability of This Report
This report is available, at no cost, at http://www.osti.gov/bridge. It is also available on paper to the U.S. Department of Energy and its contractors, for a processing fee, from:

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
phone (865) 576-8401
fax (865) 576-5728
reports@adonis.osti.gov

Disclaimer
This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor UChicago Argonne, LLC, nor any of their employees or officers, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of document authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, Argonne National Laboratory, or UChicago Argonne, LLC.

ANL/MCS-TM-309

Unit Commitment with Wind Power Generation: Integrating Wind Forecast Uncertainty and Stochastic Programming

by
E.M. Constantinescu[1], V.M. Zavala[1], M. Rocklin[2], S. Lee[3], and M. Anitescu[1]
[1]Mathematics and Computer Science Division, Argonne National Laboratory
[2]Department of Computer Science, University of Chicago
[3]Courant Institute of Mathematical Sciences, New York University

September 2009

Contents

Abstract **1**

1 Introduction **2**

2 Wind Forecast and Uncertainty Estimation Using WRF **3**
2.1 Numerical Weather Prediction 3
2.2 Uncertainty Estimation in Wind Prediction 4
2.3 Sensitivity Analysis 7

3 Unit Commitment and Energy Dispatch **8**
3.1 Deterministic Formulation 8
3.2 Stochastic Programming Formulation 11
3.3 Closed-Loop Implementation 12
3.4 Inference Analysis 12

4 Integrative Study **15**
4.1 Wind Forecast and Uncertainty Quantification 15
4.2 Economic Study Unit Commitment/Energy Dispatch 22

5 Conclusions and Future Work **25**

Acknowledgments **26**

References **27**

Unit Commitment with Wind Power Generation: Integrating Wind Forecast Uncertainty and Stochastic Programming

Emil M. Constantinescu[†], **Victor M. Zavala**[†],
Matthew Rocklin[‡], **Sangmin Lee**[§], **and Mihai Anitescu**[†,*]

Abstract

We present a computational framework for integrating of the state-of-the-art Weather Research and Forecasting (WRF) model in stochastic unit commitment/energy dispatch formulations that account for wind power uncertainty. We first enhance the WRF model with adjoint sensitivity analysis capabilities and a sampling technique implemented in a distributed-memory parallel computing architecture. We use these capabilities through an ensemble approach to model the uncertainty of the forecast errors. The wind power realizations are exploited through a closed-loop stochastic unit commitment/energy dispatch formulation. We discuss computational issues arising in the implementation of the framework. In addition, we validate the framework using real wind speed data obtained from a set of meteorological stations. We also build a simulated power system to demonstrate the developments.

[†] Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL, 60439 USA, Email:{emconsta,vzavala,anitescu}@mcs.anl.gov

[‡] Department of Computer Science, University of Chicago, 1100 East 58th Street, Chicago, IL 60637, USA

[§] Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, New York 10012, USA

[*] Managing Author, Email:anitescu@mcs.anl.gov

1 Introduction

Wind power is becoming worldwide a significant component of the power generation portfolio. In Europe, several countries already exhibit adoption levels in the range of 5-20% of the total annual demand. In the U.S. an adoption level of 20% is expected by the year 2030 [1]. Such a large-scale adoption resents many challenges to the operation of the electrical power grid because wind power is highly intermittent and difficult to predict. In particular, unit commitment (UC) and energy dispatch (ED) operations are of great importance because of their strong economic impact (on the order of billions of dollars per year) and increasing emissions concerns.

Several UC studies analyzing the impact of increasing adoption levels of wind power have been performed recently. In [26], a security-constrained stochastic UC formulation that accounts for wind power volatility is presented together with an efficient Benders decomposition solution technique. However, the issue of constructing probability distributions for the wind power is not addressed. In [24], a detailed closed-loop stochastic UC formulation is reported. The authors analyze the impact of the frequency of recommitment on the production, startup, and shutdown costs. They find that increasing the recommitment frequency can reduce costs and increase the reliability of the system. However, the authors do not present details on the wind power forecast model and uncertainty information used to support their conclusions. In [15, 19], artificial neural network (ANN) models are used to compute forecasts and confidence intervals for the total aggregated power for a set of distributed wind generators. The authors observed that forecasting the aggregated power tends to reduce the overall forecast error because it smoothes out local individual variations. A problem with empirical modeling approaches, however, is that their predictive capabilities rely strongly on the presence of persistent trends. In addition, they neglect the presence of spatio-temporal physical phenomena that can lead to time-varying correlations of the wind speeds at neighboring locations. Such approaches can thus result in inaccurate medium and long-term forecasts and over- or under-estimated uncertainty levels [18, 11], which in turn affect the expected cost and robustness of the UC solution.

In this work, we seek to exploit recent advances in numerical weather prediction (NWP) models to perform UC/EP studies with wind power adoption. The use of physical models is desirable because consistent and accurate uncertainty information can be obtained [16]. As an example, consider the missing effects of turbulence during night time, which would allow one to obtain much tighter uncertainty intervals and lower operating costs. These physical effects cannot be captured adequately through empirical modeling techniques [8]. On the other hand, the practical capabilities of NWP models are also limited. One of the major limiting factors is their computational complexity. For instance, performing data assimilation every hour at a high spatial resolution is currently not practical. In addition, extracting uncertainty information from NWP models becomes quickly intractable both from the point of view of simulation time and memory requirements. To give a reference, a single forecast run for a day ahead with a resolution of about 2 km^2 for an area that covers most of the US state of Illinois takes about 50 hours and produces around 50 GB of data. The question is: *From an operational point of view, how suitable are the forecasting capabilities of state-of-the-art NWP models?* This is an important question because NWP

models are expected to be used to make real-time operational decisions with important economic implications. To analyze this issue, we present a computational framework that integrates the Weather Research and Forecast (WRF) model with a closed-loop stochastic UC/EP formulation. In particular, we are interested in analyzing computational issues and to analyze the effects of wind uncertainty on UC/ED operations.

We focus our attention on wind speed forecasting with WRF. Arguably, more sophisticated hybrid methods that combine both NWP wind speed forecasts and empirical models are needed to map the resolution of NWP forecasts down to a specific domain and to account for system-specific characteristics (e.g., power curves, orography) [16]. Unfortunately, most real-time power data of operational wind farms is confidential, so it is complicated to validate wind power models. We model the uncertainty of the wind speed forecasts using a sampling technique that generates an ensemble of the future realizations in the targeted geographical region. The ensembles are obtained by using a scalable implementation on a distributed-memory parallel computing. In addition, we perform adjoint sensitivity analyzes that are useful in determining the domain size and control variables in the weather system. The generated ensembles are sent to a stochastic UC/EP optimization problem. We validate the forecast information using real wind speed data obtained from a set of meteorological stations located in Illinois. We use this forecast information to perform an economic analysis of the impact of increasing adoption levels of wind power on a simulated power generation system.

The paper is structured as follows. Section 2 presents details about the WRF model and uncertainty quantification. Section 3 describes the stochastic unit commitment formulation and a resampling technique used to perform inference analysis. Section 4 presents numerical validation results for WRF and the closed-loop UC simulations results. We close with final remarks and directions for future work.

2 Wind Forecast and Uncertainty Estimation Using WRF

In this section, we describe the procedures used to forecast the wind speed using WRF. The uncertainty in the wind speed is estimated by using an ensemble approach. We describe in detail the ensemble initialization and restarting procedures required in an operational framework. Furthermore, we illustrate a targeted sensitivity analysis approach that can be used for estimating the simulation domain size and placement, as well as identifying the variables in the weather system that have a strong impact on the wind forecast.

2.1 Numerical Weather Prediction

The Weather Research and Forecasting 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, which corresponds to less than 1 km^2.

To initialize the numerical weather prediction simulations, we use reanalyzed fields. 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.

2.2 Uncertainty Estimation in Wind Prediction

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 NCEP method introduced by Parrish and Derber [20, 12]. In the following sections we describe in more detail the procedures needed for generating the forecast and its uncertainty. We introduced a similar approach in [28].

Ensemble Initialization

In a normal operational mode, the NWP system evolves a given state from an initial time t_0, to a final time t_F. The initial state is produced from past simulations and reanalysis fields, that is, simulated atmospheric states reconciled with real observations. Because of the observation sparseness in the atmospheric field and the incomplete numerical representation of the atmosphere dynamics, the initial states are not known exactly and can be correctly represented only in a statistical sense. Therefore, we use an ensemble of initial conditions to describe the confidence in the knowledge of the initial state of the atmosphere.

The ensemble of the initial states is centered on the NARR field at the initial time, the most accurate information available of the atmospheric state. In other words, the ensemble expectation is exactly the NARR solution. The second statistical moment of the ensemble described by the covariance matrix $\mathbf{V}$ is approximated by the sample variance or pointwise uncertainty and its correlation. We assume a normal distribution of the uncertainty field of the initial state, a typical assumption in weather forecasting, and thus the first two statistical moments give a complete description of the uncertainty.

The initial N_{ens}-member ensemble field $x_i^{t_0} := x_i(t_0)$, $i \in \{1 \dots N_{\text{ens}}\}$, is sampled from $\mathcal{N}(x_{\text{NARR}}, \mathbf{V})$:

$$x_i^{t_0} = x_{\text{NARR}} + \mathbf{L}\xi_i\,, \quad \xi_i \sim \mathcal{N}(0, I)\,, \quad i \in \{1 \dots N_{\text{ens}}\}\,, \; \mathbf{L}\mathbf{L}^T = \mathbf{V}\,, \quad \mathbf{C}_{ij} = \frac{\mathbf{V}_{ij}}{\sqrt{\mathbf{V}_{ii}\mathbf{V}_{jj}}}\,, \quad (1)$$

where $\mathbf{C}$ is the correlation matrix and $\mathbf{V}_{ii}$ is the variance of variable i. This is equivalent to perturbing the NARR field with $\mathcal{N}(0, \mathbf{V})$. That is, $x_i = x_{\text{NARR}} + \mathcal{N}(0, \mathbf{V})$. In what follows, we describe the procedure used to estimate the correlation matrix.

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 method formally known as the National Meteorological Center (NMC) method, now NCEP [20, 12, 6], 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 ith point in space between the jth 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_i as the ith row, each of which correspond to the deviations for a single point in space. Therefore, the covariance and correlation matrices can be approximated by

$$\mathbf{V}_{ik} \approx \mathbf{d}\mathbf{d}^{\mathbf{T}} = \sum_j d_{ij} d_{kj}^T = \epsilon_i \cdot \epsilon_k = \left[\begin{array}{c|c|c|c} \epsilon_0 \cdot \epsilon_0 & \epsilon_1 \cdot \epsilon_0 & \cdots & \epsilon_n \cdot \epsilon_0 \\ \epsilon_0 \cdot \epsilon_1 & \epsilon_1 \cdot \epsilon_1 & \cdots & \epsilon_n \cdot \epsilon_1 \\ \cdots & \cdots & \cdots & \cdots \\ \epsilon_0 \cdot \epsilon_n & \epsilon_1 \cdot \epsilon_n & \cdots & \epsilon_n \cdot \epsilon_n \end{array}\right], \quad \mathbf{C}_{ik} = \frac{\epsilon_i \cdot \epsilon_k}{|\epsilon_i||\epsilon_k|}.$$

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 can be efficiently estimated by computing several times the correlation along a ray cast into this vertical level in the East-West direction, fitting a Gaussian to the central peak, and then averaging among several rays. This approach captures the difference between East-West and North-South winds due to the Coriolis effect and the earth rotation, and also faster and larger scale winds in the upper atmosphere. However, we assume that correlations and winds are roughly similar in nature across the continental U.S. This process is repeated in the vertical direction such that our results are also correlated level to level.

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 (1).

Ensemble Propagation through the WRF Model

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

$$x_i^{t_F} = \mathcal{M}_{t_0 \to t_F}\left(x_i^{t_0}\right) + \eta_i(t),\ x_i^{t_0} \sim \mathcal{N}\left(x_{\text{NARR}}, \mathbf{V}^{t_0}\right),\ \eta_i \sim \mathcal{N}\left(0, \mathbf{Q}\right),\ i \in \{1 \ldots N_{\text{ens}}\}, \quad (2)$$

where $\mathcal{M}_{t_0 \to t_F}(\bullet)$ 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 [28].

In this study, we assume that the numerical model (WRF) is perfect, that is, $\eta \equiv 0$, and 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.

Accounting for Error Underestimation

In an operational setting, observations become available periodically and can be assimilated in the atmospheric state. In order to account for the new information, the ensemble needs to be recentered on the new reanalyzed field. In our example, we consider 12-hour windows between restarts. This simple adjustment corresponds to a correction in the ensemble expectation. However, the pointwise error estimates given by the ensemble variance may be over- or underestimated because of simulation errors accumulated along the model trajectory due to uncertain data and incomplete NWP physics. In other words, the ensemble statistics may diverge from the true statistics. Therefore, the error levels need to be re-estimated before each initialization.

One sensible approach used to correct the ensemble statistics is as follows. Consider the reanalyzed field x_{NARR} as the true state, for computing corrections purposes only, and require that this reanalyzed solution be on average within one standard deviation as given by the ensemble spread. This approach corresponds to finding a factor γ that inflates the ensemble spread about its expectation. Let us consider again the ensemble $x_{i=1\ldots N_{\text{ens}}} \in \mathbb{R}^M$ and the reanalyzed solution x_{NARR}. Denote by $\overline{x} = \frac{1}{N_{\text{ens}}} \sum_{i=1}^{N_{\text{ens}}} x_i$ the sample expectation and by $\sigma_j = \sqrt{\mathbf{S}^2{}_{jj}}$, $j \in \{1 \ldots M\}$, the standard deviation, where $\mathbf{S}^2 = \frac{1}{N_{\text{ens}}-1} \sum_{i=1}^{N_{\text{ens}}} (x_i - \overline{x})(x_i - \overline{x})^T$ is the sample covariance estimation. Then, we have

$$\gamma = max\left(1, min\left(\gamma_\sigma, 4\right)\right) \ , \quad \gamma_\sigma = \text{mean}_{U,V,T|k=1\ldots5} \left(\frac{|x_{\text{NARR}} - \overline{x}|}{\sigma} \right) \ , \tag{3}$$

where U, V, and T are the the wind field components and the temperature, that is, the ensemble variables under consideration. For this comparison, we consider only the first five layers, which include grid points located below 300 m. The new ensemble is then inflated by

$$x_i \leftarrow \overline{x} + \gamma \left(x_i - \overline{x}\right) \ , \ i \in \{1 \ldots N_{\text{ens}}\} \, . \tag{4}$$

The factor is bounded between one, because the model error is underestimated in our case, and four to avoid large jumps in the solution and destabilize the NWP model. Experimentally, however, we noticed that $\gamma \approx 2$, which confirms that this approach tends to underestimate the uncertainty. This fact is not unexpected because the model error is not considered.

2.3 Sensitivity Analysis

Sensitivity analysis is the study of how the model solution is affected by small perturbations in model variables and parameters [2, 3]. We write the sensitivity of the solution x with respect to parameter p_i as $S_i(t) = \frac{\partial x(t)}{\partial p_i}$ or scaled to be unitless, $S_i(t) = \frac{\partial x(t)}{\partial p_i}\frac{p_i}{x(t)}$ [29]. In this study, we are interested in the effect that the initial condition at location i, $p_i \equiv x_i^{t_0} := x_i(t_0)$, has at some targeted locations in the final system state $x^{t_F} := x(t_F)$. Therefore, the sensitivity takes the form

$$S_i = \frac{\partial x^{t_F}}{\partial x_i^{t_0}} \frac{x_i^{t_0}}{x^{t_F}} \,. \tag{5}$$

Just as the model state x^{t_0} is evolved through $\mathcal{M}$, the sensitivity S_i is evolved by the gradient of the model, also known as the *tangent linear model*,

$$S_i^k = \frac{\partial \mathcal{M}}{\partial x}(x^{t_{k-1}}, p) S^{k-1} + \frac{\partial \mathcal{M}}{\partial p_i}(x^{t_{k-1}}, p) \,, \qquad S_i^0 = \frac{\partial x^{t_0}}{\partial p_i} \,, \qquad t_k \in [t_0, t_F] \,,$$

where for brevity we do not include the implicit dependence of $\mathcal{M}$ on t. This is useful if one is interested in what effect a small perturbation at a single source location would have on the future states, x^{t_k}. Alternatively, one could consider the inverse or *adjoint process* of observing some target point in the state space at future times and inferring what points in the initial conditions had a strong influence on the target point.

In this study, we are interested in finding the regions in the initial state to which target points at later times are most sensitive. Therefore, the sensitivities are computed in terms of a cost function, which is a function of the state at the final time,

$$\Psi(x^{t_F}(x^{t_0})) \in \mathbb{R} \,, \quad \left(\frac{\partial \Psi}{\partial x^{t_0}}\right)^T = \left[\frac{\partial \Psi}{\partial x_1^{t_0}} \; \cdots \; \frac{\partial \Psi}{\partial x_M^{t_0}}\right]^T \in \mathbb{R}^M, \tag{6}$$

where M is the dimension of the initial state vector. Using the chain rule, one obtains

$$\frac{\partial \Psi(x^{t_F})}{\partial x_i^{t_0}} = \frac{\partial \Psi(x^{t_F})}{\partial x^{t_F}} \frac{\partial x^{t_F}}{\partial x^{t_0}} = \frac{\partial \Psi(x^{t_F})}{\partial x^{t_F}} S_i^{t_F} \,.$$

Following [29], the direct sensitivity can be extended for all time indexes

$$\frac{\partial \Psi(x^{t_F})}{\partial x^{t_0}} = \frac{\partial \Psi(x^{t_F})}{\partial x^{t_F}} \frac{\partial x^{t_F}}{\partial x^{t_{F-1}}} \cdots \frac{\partial x^{t_1}}{\partial x^{t_0}} \frac{\partial x^{t_0}}{\partial p_i} \,, \tag{7a}$$

$$\frac{\partial x^{t_k}}{\partial x^{t_{k-1}}} = \frac{\partial \mathcal{M}}{\partial x}(x^{t_{k-1}}) \in \mathbb{R}^{M \times M} \,, \tag{7b}$$

$$\frac{\partial x^{t_0}}{\partial p_i} = e_i \in \mathbb{R}^M \,; \; i \in \{1 \ldots M\} \,. \tag{7c}$$

Alternatively, by transposing (7), this can be solved by using the adjoint process to evolve the sensitivity in reverse order:

$$\left(\frac{\partial \Psi(x^{t_F})}{\partial x^{t_0}}\right)^T = \left(\frac{\partial x^{t_1}}{\partial x^{t_0}}\right)^T \cdots \left(\frac{\partial x^{t_F}}{\partial x^{t_{F-1}}}\right)^T \left(\frac{\partial \Psi(x^{t_F})}{\partial x^{t_F}}\right)^T \,.$$

Because we start directly from the final state, we need only to evolve forward the system once and not for each parameter. We further consider the adjoint variables λ^{t_k}, which represent the gradients of the cost function with respect to perturbations in the state at earlier times $\lambda^{t_k} = \left(\frac{\partial \Psi(x^{t_F})}{\partial x^{t_k}}\right)^T$. The following equations need to be satisfied [29]:

$$\lambda^{t_{k-1}} = \left(\frac{\partial x^{t_k}}{\partial x^{t_{k-1}}}\right)^T \lambda^{t_k} = \left(\frac{\partial \mathcal{M}}{\partial x}(x^{t_{k-1}})\right)^T \lambda^{t_k} = \mathbf{M}^T{}_{k-1}\lambda^{t_k}, \quad k \in \{1 \dots F\},$$
$$\lambda^{t_F} = \left(\frac{\partial \Psi(x^{t_F})}{\partial x^{t_F}}\right)^T.$$

Note that we evolve the adjoint variable λ^{t_k} backwards in time, starting at the final time and taking steps with the *adjoint model* $\mathbf{M}^T = \left(\frac{\partial \mathcal{M}}{\partial x}\right)^T$ back to the initial time. As we did in equation (5), we can also consider the scaled adjoint sensitivity, which can be physically interpreted as the percentage change in the cost function when the variable $x_i^{t_k}$ is changed 1%:

$$\lambda_i^{t_k} = \frac{\partial \Psi(x^{t_F})}{\partial x_i^{t_k}} \frac{x_i^{t_k}}{\Psi(x^{t_F})}. \tag{8}$$

Some practical difficulties arise in obtaining the adjoint model, as discussed in [14, 7]. The WRF model contains simplified adjoint (WRFPlus-AD) and tangent linear (WRFPlus-TL) models, which are encapsulated in WRFPLUS [27], and based on a simplified WRF model that includes only some simple physical processes such as vertical diffusion and large-scale condensation. This simplified WRF code has been run through a source-to-source program called Transformation of Algorithm in FORTRAN (TAF) [9] to automatically produce both gradients and adjoints of the gradients.

3 Unit Commitment and Energy Dispatch

The unit commitment problem consists in finding a cost-optimal plan of on/off states and power levels for a set of distributed power generators. This problem normally is solved by the system operator each time new price information arrives from the generator owners (market biding). Depending on the hierarchical structure used (centralized or decentralized), several variants of the UC problem can be established. Here, we adopt the traditional centralized approach, in which the system operator has control over the whole set of generators. The UC problem uses a forecast of uncertain factors such as the load or, in our case, of wind power generation. Since the forecast is not perfect or disturbances might arise, an energy dispatch problem is solved in closed-loop to correct the power levels and satisfy the demands.

3.1 Deterministic Formulation

Many different mathematical formulations of the UC problem exist. They differ mostly in the constraints used to capture the dynamic performance of the generators (e.g., ramp

limits) or in the assumed cost models [22]. The UC formulation considered here is based on the mixed-integer linear programming (MILP) formulation of Carrion and Arroyo [5]. This has the general form

$$\min_{p_{j,k},\overline{p}_{j,k}\nu_{j,k}} \sum_{j\in\mathcal{N}}\sum_{k\in\mathcal{T}} c^p_{j,k} + c^u_{j,k} + c^d_{j,k} \tag{9a}$$

$$\text{s.t.} \sum_{j\in\mathcal{N}} p_{j,k} + \sum_{j\in\mathcal{N}_{wind}} \mathbb{E}[p^{wind}_{j,k}] = D_k,\ k\in\mathcal{T} \tag{9b}$$

$$\sum_{j\in\mathcal{N}} \overline{p}_{j,k} + \sum_{j\in\mathcal{N}_{wind}} \mathbb{E}[p^{wind}_{j,k}] \geq D_k + R_k,\ k\in\mathcal{T} \tag{9c}$$

$$(10) - (17).$$

The sets $\mathcal{T} := \{1..T\}$, $\mathcal{N} := \{1..N\}$, and $\mathcal{N}_{wind} := \{1..N_{wind}\}$ represent the time periods, thermal units, and wind generators, respectively. The demand at each time period k is denoted by D_k while the reserve requirement is R_k. The power output of the thermal unit j at time k is given by the continuous variable $p_{j,k}$ while the expected value of the power output of the wind unit j at time k is given by the fixed parameter $\mathbb{E}[p^{wind}_{j,k}]$. The continuous variable $\overline{p}_{j,k}$ represents the maximum power output of the thermal unit j at time k. This variable is introduced in order to model the spinning reserve requirements which are given by the differences $\overline{p}_{j,k} - p_{j,k}$. The units of all the power outputs are MW. The on/off status of thermal unit j at time k is given by the binary variable $\nu_{j,k}$.

The production cost for each thermal unit is approximated by using the linear model [4]

$$c^p_{j,k} = a_j\nu_{j,k} + b_j p_{j,k},\ j\in\mathcal{N},\ k\in\mathcal{T}\,, \tag{10}$$

where a_j and b_j are cost coefficients. To model the startup cost $c^u_{j,k}$ we use a staircase cost of the form

$$K^t_j = \begin{cases} cc_j & \text{if } t > t^{\text{cold}}_j + DT_j \\ hc_j & \text{otherwise} \end{cases},\ j\in\mathcal{N},\ t = 1,...,ND_j.$$

where K^t_j is the cost of interval t of the staircase cost, ND_j is the number of intervals of the staircase cost, and cc_j and hc_j are cost coefficients. This leads to the following set of inequality constraints:

$$c^u_{j,k} \geq K^t_j\left(\nu_{j,k} - \sum_{n=1}^{t}\nu_{j,k-n}\right),\ j\in\mathcal{N},\ k\in\mathcal{T},\ t = 1,...,ND_j\,, \tag{11a}$$

$$c^u_{j,k} \geq 0,\ j\in\mathcal{N},\ k\in\mathcal{T}\,. \tag{11b}$$

The formulation of the shutdown cost is given by

$$c^d_{j,k} \geq C_j\big(\nu_{j,k-1} - \nu_{j,k}\big),\ j\in\mathcal{N},\ k\in\mathcal{T}\,, \tag{12a}$$

$$c^d_{j,k} \geq 0,\ j\in\mathcal{N},\ k\in\mathcal{T}\,, \tag{12b}$$

where C_j is the shutdown cost of unit j. The power output of each unit at each period must satisfy the bounds

$$\underline{P}_j \nu_{j,k} \leq p_{j,k} \leq \overline{p}_{j,k},\ j \in \mathcal{N},\ k \in \mathcal{T}\,, \tag{13a}$$

$$0 \leq \overline{p}_{j,k} \leq \overline{P}_j \nu_{j,k},\ j \in \mathcal{N},\ k \in \mathcal{T}\,, \tag{13b}$$

where $\overline{P}_j$ and $\underline{P}_j$ are the maximum and minimum capacities of unit j, respectively. The thermal power outputs must also satisfy the ramp-up limits

$$\overline{p}_{j,k} \leq p_{j,k-1} + RU_j \nu_{j,k-1} + SU_j\big(\nu_{j,k} - \nu_{j,k-1}\big) + \overline{P}_j(1-\nu_{j,k}),\ j \in \mathcal{N},\ k \in \mathcal{T}. \tag{14}$$

The shutdown and ramp-down limits are

$$\overline{p}_{j,k-1} \leq p_{j,k} + RD_j \nu_{j,k} + SD_j\big(\nu_{j,k-1} - \nu_{j,k}\big) + \overline{P}_j(1-\nu_{j,k-1}),\ j \in \mathcal{N},\ k \in \mathcal{T}. \tag{15}$$

Here, RD_j, RU_j, SD_j, and SU_j are the ramp-down, ramp-up, shutdown ramp, and startup ramp limits of unit j, respectively. The minimum up time constraints are

$$\sum_{k=1}^{G_j}(1-\nu_{j,k}) = 0,\ j \in \mathcal{N} \tag{16a}$$

$$\sum_{n=k}^{k+UT_j-1} \nu_{j,n} \geq UT_j\big(\nu_{j,k} - \nu_{j,k-1}\big),\ j \in \mathcal{N},\ k = G_j+1,\ldots,T-UT_j+1 \tag{16b}$$

$$\sum_{n=k}^{T}\big(\nu_{j,n} - (\nu_{j,k} - \nu_{j,k-1})\big) \geq 0,\ j \in \mathcal{N},\ k = T-UT_j+2,\ldots,T\,, \tag{16c}$$

where UT_j are the minimum up time limits and $G_j = \min(T, (UT_j - U_j^0)\nu_{j,0})$ is the number of periods unit j must be initially ON. The initial state of unit j is denoted by $\nu_{j,0}$ and is a fixed parameter. The minimum down time constraints are formulated as

$$\sum_{k=1}^{L_j}\nu_{j,k} = 0,\ j \in \mathcal{N} \tag{17a}$$

$$\sum_{n=k}^{k+DT_j-1}(1-\nu_{j,n}) \geq DT_j\big(\nu_{j,k-1} - \nu_{j,k}\big),\ j \in \mathcal{N},\ k = L_j+1,\ldots,T-DT_j+1 \tag{17b}$$

$$\sum_{n=k}^{T}\big(1-\nu_{j,n} - (\nu_{j,k-1} - \nu_{j,k})\big) \geq 0,\ j \in \mathcal{N},\ k = T-DT_j+2,\ldots,T\,, \tag{17c}$$

where DT_j denote the minimum down-time limits and $L_j = \min(T, (DT_j - S_j^0)(1-\nu_{j,0}))$ is the number of periods unit j must be initially OFF.

As noted in [5], the above UC formulation requires a significantly smaller number of binary variables compared to traditional formulations. For instance, in all the formulations reviewed in [10], an extra set of variables is required to identify the units that are *started* at the beginning of each period. The proposed formulation also yields an accurate representation of ramping limits and individual contributions to the spinning reserve requirements. Modeling the ramp limits in a consistent manner is particularly critical as this simulates the responsiveness of the thermal units in the presence of short-term fluctuations of the load and wind power [25]. Note that it is possible to use this model to simulate the performance of the energy dispatch problem by fixing the commitment variables $\nu_{j,k}$.

3.2 Stochastic Programming Formulation

We extend the previous deterministic formulation by considering corrective actions on the power outputs of the thermal generators to account for the uncertainty in the wind power outputs. The problem can be cast as a two-stage stochastic programming problem similar to the ones proposed in [4, 26, 24]. The first stage decision variables are the current thermal power outputs $p_{j,1}$, $\overline{p}_{j,1}$ and the commitment profiles over the entire planning horizon $\nu_{j,k}$. The power outputs are nonanticipatory (here and now) because it is assumed that the current wind power outputs $p_{j,1}^{wind}$ are known and given by $p_{j,1}^{wind,true}$. To formulate the second stage, we consider multiple realizations of the wind power outputs $p_{s,j,k}^{wind}$,and we define scenario-dependent thermal power outputs $p_{s,j,k}$ and $\overline{p}_{s,j,k}$ with $k > 1$ (wait and see). Note that we do not define second-stage scenario-dependent commitment variables because we wish to keep the problem computationally tractable. The second-stage power outputs implicitly simulate the closed-loop response of the energy dispatch problem. The formulation of the stochastic optimization problem is given by

$$\min_{p_{s,j,k},\overline{p}_{s,j,k},\nu_{j,k}} \frac{1}{N_S}\sum_{s\in\mathcal{S}}\left(\sum_{j\in\mathcal{N}}\sum_{k\in\mathcal{T}} c^p_{s,j,k} + c^u_{j,k} + c^d_{j,k}\right) \tag{18a}$$

$$\text{s.t.} \sum_{j\in\mathcal{N}} p_{s,j,k} + \sum_{j\in\mathcal{N}_{wind}} p^{wind}_{s,j,k} = D_k,\ s\in\mathcal{S}, k\in\mathcal{T} \tag{18b}$$

$$\sum_{j\in\mathcal{N}} \overline{p}_{s,j,k} + \sum_{j\in\mathcal{N}_{wind}} p^{wind}_{s,j,k} \geq D_k + R_k,\ s\in\mathcal{S}, k\in\mathcal{T} \tag{18c}$$

$$p_{s,j,1} = p_{1,j,1}\ s\in\mathcal{S},\ j\in\mathcal{N} \tag{18d}$$

$$\overline{p}_{s,j,1} = \overline{p}_{1,j,1}\ s\in\mathcal{S},\ j\in\mathcal{N} \tag{18e}$$

$$(10)-(15), \quad s\in\mathcal{S}$$

$$(16)-(17). \tag{18f}$$

The ramp and power limit constraints are defined over each scenario, $s \in \mathcal{S}$ where we substitute $p_{j,k} \leftarrow p_{s,j,k}$ and $\overline{p}_{j,k} \leftarrow \overline{p}_{s,j,k}$. The nonanticipativity constraints for the power outputs in the first time step are given by equations (18d) and (18e). For the known wind power outputs we set $p^{wind}_{s,j,1} \leftarrow p^{wind,true}_{j,1}$. Note that if the stochastic formulation is able

to capture the uncertainty of the wind power accurately, the reserve requirements can be reduced to less conservative levels or even be removed. In other words, the reserves already incorporate some robustness into the UC problem. We have decided to include the reserves in order to analyze the interplay between the explicit robustness introduced by the reserves and the implicit robustness introduced by the stochastic formulation. This approach has been proposed in [21] and will be used in Section 4 to analyze the effect of increasing levels of penetration that can be achieved through the stochastic UC formulation. Load uncertainty has not been considered in this study in order to isolate the effects of wind power uncertainty. As in the deterministic formulation, one can solve a closed-loop *stochastic* energy dispatch problem by fixing the commitment actions.

3.3 Closed-Loop Implementation

To simulate the closed-loop performance of the power system, we consider a *rolling-shrinking* horizon approach. The starting rolling time is reset to one each time new price information is obtained. This period is assumed to be $T = 24$ hours. At the starting rolling time, we assume that the wind power forecast becomes available from WRF for the next 24 hours. At this point, the stochastic unit commitment problem is solved by using the current wind power outputs $p^{wind}_{s,j,\ell} \leftarrow p^{wind,true}_{j,\ell}$ and the future forecasts $p^{wind}_{s,j,k},\ k = \ell + 1, ..., T$ where ℓ is the current time step. The solution of this problem gives the commitment profiles $\nu_{j,k}$ over the 24-hour rolling horizon. At each step inside the rolling horizon $\ell = 2, ..., T$, the horizon is shrunk by one time step $T \leftarrow T - \ell$ and the stochastic energy dispatch is solved over the remaining horizon with the new true wind power but the same forecast information. Each of these shrinking horizon problems gives the current power outputs $p_{j,\ell}$ and $\overline{p}_{j,\ell}$ at current time ℓ. Note that this requires shifting the nonanticipativity constraints to the corresponding initial times of the shrinking problems. In addition, it is necessary to shift the initial state of the thermal units and the corresponding minimum up- and down-time parameters G_j and L_j.

3.4 Inference Analysis

In the above stochastic formulation, the wind power outputs are assumed to have a probability distribution $\mathbb{P}$. In most stochastic optimization studies this distribution is assumed to be known. As seen in Section 2, obtaining this distribution is part of the modeling task. Since many different forecast models (autoregressive, ANN, physics based) can be used to construct the error distribution, there is not a unique distribution. From a practical point of view, what we expect from a distribution is that it is able to encapsulate the actual realizations of wind power and it has tight confidence intervals. The encapsulation property ensures robustness of the solution (it satisfies the load in each possible scenario) while tightness ensures that the cost penalty incurred (with respect to the perfect information cost) is not too strong. As explained in Section 2, we model the wind power distribution by propagating an assumed Gaussian distribution of the initial state conditions through the WRF model. Because of the complexity of the model, we are limited to a single batch of a few (less than a hundred) weather samples. From a stochastic optimization point of

view, this is an issue because we are not solving the problem with the full distribution. Nevertheless, we want to at least get an idea of how sensitive the costs are to changes in the random information. In addition, we might want to compute the cost variance or confidence intervals. Performing inference analysis through sampling techniques [13] would require resampling many times the initial state distribution and propagating each sample through the WRF model. This approach is not practical from a computational point of view. To avoid this limitation, we next present a heuristic resampling technique.

Weighted Average Sampling

To create new wind speed time series from the existing WRF realizations, we express a new realization as a weighted average of the available ones. Suppose the WRF model is $x(t) = \mathcal{M}(t, x(0))$, where $x(t)$ is the state vector at time t. If we are given N_S samples x_j and we can write $x(0) = \sum_{j\in\mathcal{S}} w_j x_j(0)$, the propagation of $x(0)$ is $x(t) = \mathcal{M}(t, x(0)) = \mathcal{M}(t, \sum_{j\in\mathcal{S}} w_j x_j(0))$. Assuming the variance of the samples is small, we can write $x_j(0) = \bar{x}(0) + \epsilon_j(0)$. We can justify the computation of weighted averages of the time series because

$$\begin{aligned}
x(t) &= \mathcal{M}\left(t, \bar{x}(0) + \sum_j w_j \epsilon_j(0)\right) \\
&\approx \mathcal{M}(t, \bar{x}(0)) + \sum_j w_j \frac{\partial \mathcal{M}}{\partial x}(t, \bar{x}(0))\epsilon_j(0) \\
&= \sum_j w_j \Big(\mathcal{M}(t, \bar{x}(0)) + \frac{\partial \mathcal{M}}{\partial x}(t, \bar{x}(0))\epsilon_j(0)\Big) \\
&\approx \sum_j w_j \mathcal{M}\Big(t, \bar{x}(0) + \epsilon_j(0)\Big) \\
&= \sum_j w_j \mathcal{M}(t, x_j(0)) \\
&= \sum_j w_j x_j(t).
\end{aligned}$$

In other words, the weighted average of the time series approximates, to first order, the nonlinear propagation of weighted samples of the initial conditions. Inspired by kernel density estimation, the weights are chosen to be Gaussian near the unit vectors in the standard basis on a hyperplane $\sum_{j\in\mathcal{S}} w_j = 1$ in the w space.

Computation of Confidence Intervals

The two-stage stochastic UC problem with fixed binary variables can be expressed in the following abstract form :

$$\min_{x\geq 0} \; c^T x + \mathcal{Q}(x), \text{ s.t. } Ax = b. \tag{19}$$

Here, x are the first-stage decision variables, and $\mathcal{Q}(x) = \mathbb{E}\left[Q(x,\xi)\right]$ is the second-stage cost. We assume that the probability distribution $\mathbb{P}$ of ξ has finite support; that is, ξ has a finite number of scenarios $\{\xi_1, ..., \xi_K\}$ with probabilities $p_k \in (0,1)$. Consequently, we can write $\mathcal{Q}(x) = \frac{1}{K}\sum_{k=1}^{K} Q(x,\xi_k)$, where

$$Q(x,\xi_k) = \min_{y_k \geq 0} \; q^T y_k, \text{ s.t. } Tx + Wy_k = \xi_k, \; k = 1, ..., K. \tag{20}$$

Here, y_k are the second-stage decision variables, and ξ_k are the realizations of the wind power outputs. K is a very large number so it is impractical to solve the stochastic problem exactly. Therefore, given a fixed number of realizations $N_S << K$, we solve the approximate problem,

$$\min_{x \geq 0} \; c^T x + \frac{1}{N_S}\sum_{k=1}^{N_S} Q(x,\xi_k), \text{ s.t. } Ax = b. \tag{21}$$

We seek to estimate lower and upper bounds of the true optimal solution v^* (using the entire set of K realizations) and their corresponding confidence intervals. A lower bound can be estimated generating $j = 1, ..., M$ batches, each of size N_S and solve (21) for each batch. If we denote as $\hat{v}^j_{N_S}$ the optimal cost of each SAA problem, we can estimate the lower bound as

$$L_{N_S,M} = \frac{1}{M}\sum_{j=1}^{M} \hat{v}^j_{N_S}. \tag{22}$$

The sample variance estimator is given by

$$s_L^2(M) = \frac{1}{M-1}\sum_{j=1}^{M} \left(\hat{v}^j_{N_S} - L_{N_S,M}\right)^2, \tag{23}$$

and the $(1-\alpha)$ confidence intervals of the lower bound are

$$\left[L_{N_S,M} - \frac{z_\alpha s_L(M)}{\sqrt{M}} \; , \; L_{N_S,M} + \frac{z_\alpha s_L(M)}{\sqrt{M}}\right], \tag{24}$$

where z_α satisfies $\mathbb{P}\left\{\mathcal{N}(0,1) \leq z_\alpha\right\} = 1-\alpha$. To estimate the upper bound, we pick a given value for the first-stage variables $\hat{x}$ and generate a new set of $j = 1, ..., M$ batches of data. We then evaluate (19), leading to $\hat{f}^j_N(\hat{x})$. Note that each evaluation involves the solution of the second-stage problem (20). As before, the upper bound confidence intervals are given by

$$\left[U_{N_S,M} - \frac{z_\alpha s_U(M)}{\sqrt{M}} \; , \; U_{N_S,M} + \frac{z_\alpha s_U(M)}{\sqrt{M}}\right], \tag{25}$$

with $U_{N_S,M} = \frac{1}{M}\sum_{j=1}^{M} \hat{f}^j_{N_S}$ and $s_U^2(M) = \frac{1}{M-1}\sum_{j=1}^{M}\left(\hat{f}^j_{N_S} - U_{N_S,M}\right)^2$.

4 Integrative Study

In this section we integrate the wind speed forecasts produced by WRF by following the procedure described in Section 2 with the stochastic unit commitment/energy dispatch formulations described in Section 3.

4.1 Wind Forecast and Uncertainty Quantification

We begin with a description of the weather model setup and illustrate its performance in a parallel environment that mimics a real operational setup. We use the WRF model to forecast the wind speed in a specific region that covers the state of Illinois. We set up a computational nested domain structure including a high-resolution sector that covers the target area and two additional domains of larger coverage but lower resolution. The parent domains supply the boundary conditions for the nested ones, and the largest domain has prescribed boundary conditions from coarser simulations. This setup is illustrated in Figure 1. A similar setup with one coarse domain is described in [28].

ID	Size	Grid
#1	130×60	$32\,\mathrm{km}^2$
#2	126×121	$6\,\mathrm{km}^2$
#3	202×232	$2\,\mathrm{km}^2$

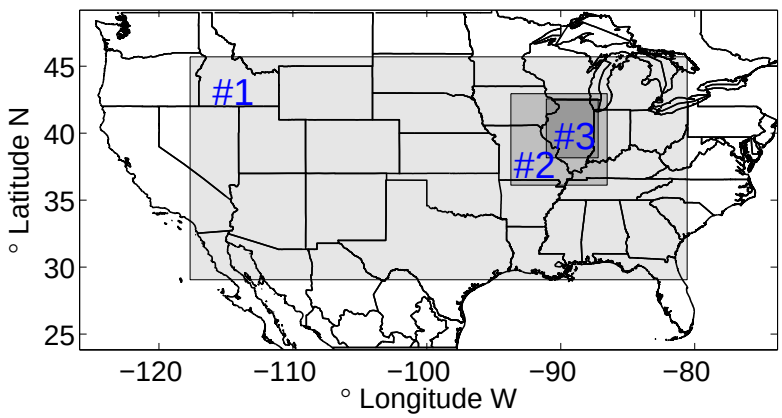


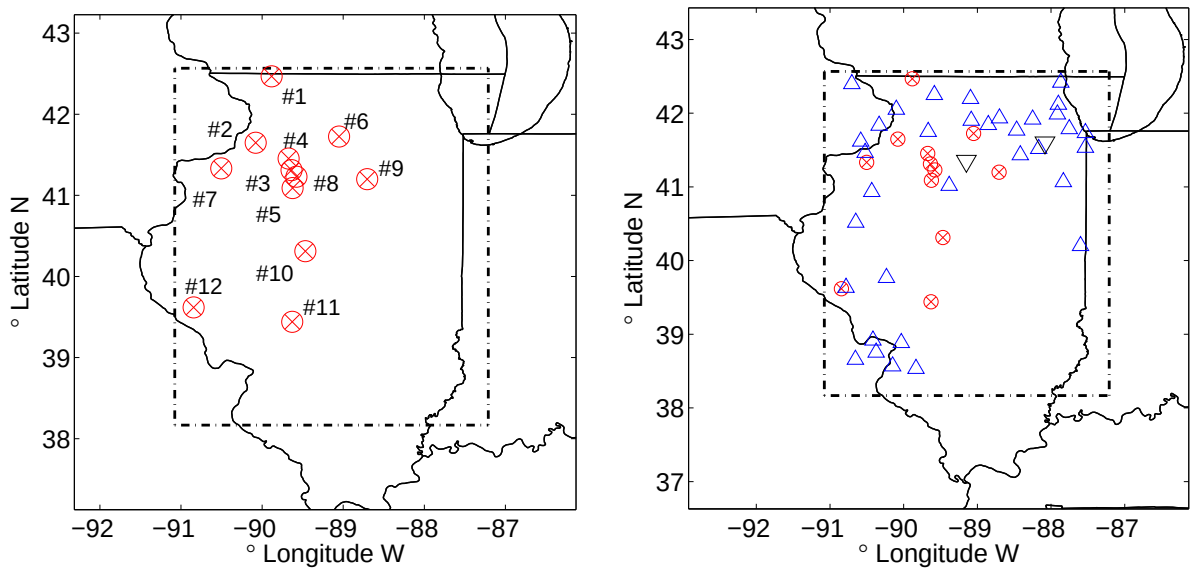
Figure 1: Size and spatial resolution of the computational domain.

To simulate realistic conditions, we generate six ensemble data sets, each containing the predicted wind speed for Illinois region corresponding to domain # 3 (the innermost) in Figure 1. Each ensemble has $N_{\text{ens}} = 30$ members. The data is sampled every 10 minutes, and each ensemble is evolved one day ahead. The starting time of the experiment t_0 corresponds to June 1st, 2006, 6:00 PM CT (local time), with each data set restarted from the reanalyzed solution at time $t_0 + (k-1) \times (12\,\text{hours})$ with $k = 1, \ldots, 6$. In other words, each data set is started at the revalidation time with 12-hour increments.

Validation Using Wind and Temperature Data Measurements

We validate the WRF model against observations at several weather stations throughout Illinois. The weather station observations were obtained from the National Climatic Data Center (NCDC), and their locations are illustrated in Figure 2.b.

In Figure 3 we show the wind speed ($\pm 2\sigma$) and temperature ($\pm 3\sigma$) predictions and measurements for Peru and Chicago, IL, identified as rotated triangles in Figure 2.b. Each ensemble evolves for 24 hours and new ones are started every 12 hours; therefore, the last half of each forecast overlaps with a more recent one. In our illustrations we show the most recent results on top.



a) Wind turbine location and ID

b) Location of weather stations

Figure 2: Windmill farms and weather station (triangles) locations in Illinois.

We remark that the wind speed measurements obtained from NCDC are given in miles per hour rounded to the nearest integer. Doing so has the unfortunate effect of diminishing the wind variability and yielding more pessimistic than real validation results. The temperature measurements are also rounded, but this effect is less visible. In this test case, however, the temperature is relatively easier to predict whereas the wind typically has a wider variability and is much more difficult to predict accurately. This point is enforced by the correlogram for the temperature and wind speed at Peru, IL shown in Figure 6. The uncertainty intervals as given by WRF capture the trends very well, with few exceptions. Similar results are obtained at the other weather stations. These results show that WRF can be used to make accurate wind predictions and provide confidence intervals.

Validation at Wind Farm Locations in Illinois

We present validation results at 12 active wind-farms in the state of Illinois. The locations were obtained from `http://www.windforillinois.org`; the coordinates are summarized in Table 1 and are displayed in Figure 2. Currently, the power produced by wind turbines depends on the wind speed at elevations of about 40-120 meters. The wind speed fields at

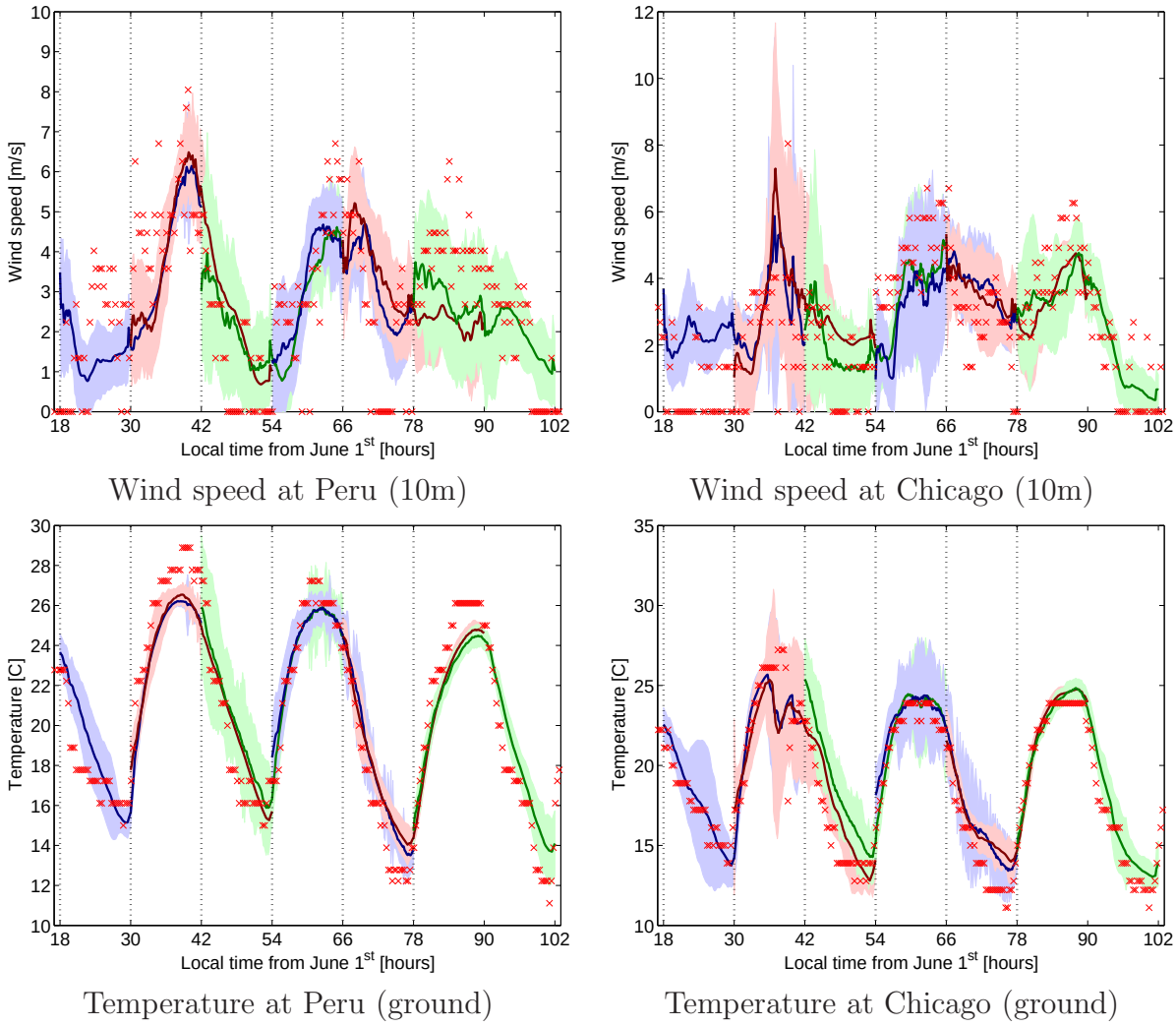


Figure 3: Wind speed ($\pm 2\sigma$) and temperature ($\pm 3\sigma$) predictions and measurements (x) for Peru and Chicago, IL. The vertical dashed lines denote the beginning of a new 24-hour prediction window; different colors are used to indicate ensembles started at different times.

these heights can be extracted from WRF. Unfortunately, the NCDC data available for validation is reported only at 10 meters. Obtaining wind speed data at higher altitudes requires access to proprietary databases of operational wind farms. Consequently, we provide validation results only at 10 meters.

The wind speed fields at 10 and 100 meters above the ground for three consecutive days of June 2006 are presented in Figures 4 and 5, respectively. The order of the windows goes from left to right and coincide with the wind-farm location numbering shown in Figure 2. From Figure 4, we note that the WRF realizations are able to capture the general trends of the actual observations. In addition, they are able to encapsulate the observations. Note, however, that the wind speed is relatively low at this height. The

Table 1: Location of wind farms in the state of Illinois. Their location is also displayed on the state map in Figure 2.

ID	Latitude °N	Longitude °W	ID	Latitude °N	Longitude °W
1	42.467032	89.88327	7	41.333256	90.503311
2	41.648256	90.078578	8	41.453728	89.673843
3	41.089702	89.623718	9	41.199323	88.704987
4	41.310953	89.637623	10	40.311996	89.467163
5	41.230315	89.577026	11	39.44017	89.629029
6	41.725462	89.051399	12	39.617987	90.845089

maximum average is around 6-7 meters per second. From Figure 5 we see that the wind speeds increase significantly at 100 meters, reaching a maximum average of around 10 meters per second. Note also that the uncertainty levels increase at this height as a result of the larger range and variability. This increase is also expected because most of the wind speed data assimilated in WRF is near ground level.

In Figure 7 we show the spatial correlations of the wind speed for wind farm #8 on June 5, 1:50 AM, as inferred from the 30-member WRF ensemble simulation. In this analysis, we have observed that the wind speed is highly correlated over a wide geographical region and that it has a nontrivial spatial structure. This observation is confirmed by comparing Figures 10 and 2. Here, we can see that the wind speed realizations for wind farms #2, 3, 4, and 8 are strongly correlated, as predicted by the correlation mapping. The correlation information is highly valuable, because it can augment the temporal uncertainty analysis by identifying the locations that are likely to experience a similar behavior. This can be used, for instance, to identify faulty sensors at certain locations or to aggregate the performance of multiple turbines in a consistent manner.

Adjoint Sensitivity Analysis Results

In this section we describe the procedure to estimate the sensitivity of the wind speed $\Psi = \sqrt{U^2 + V^2}$, where U and V are the West-East and South-North wind components, respectively. The initial adjoint values at the final time is therefore given by

$$\lambda^{t_F} = \frac{\partial \Psi}{\partial \bullet} = \frac{\bullet}{\sqrt{U^2 + V^2}}\,, \qquad \bullet = \{U, V\}\,, \tag{26}$$

which is the derivative of the wind speed (or cost function) at the final time, as discussed in Section 2.3. The initialization with the scaled cost function is then

$$\lambda^{t_F} = \frac{\partial \Psi}{\partial \bullet}\frac{\bullet}{\Psi} = \frac{\bullet}{\sqrt{U^2 + V^2}}\frac{\bullet}{\sqrt{U^2 + V^2}} = \frac{(\bullet)^2}{U^2 + V^2}\,, \qquad \bullet = \{U, V\}\,. \tag{27}$$

The scaled version is useful because it allows us to compare sensitivities in model states with different units of measure; for instance, wind speed and temperature.

To illustrate the results that can be obtained from such an analysis, in Figure 8 we show the wind speed sensitivity in a region in northern Texas with respect to the wind speed, 6

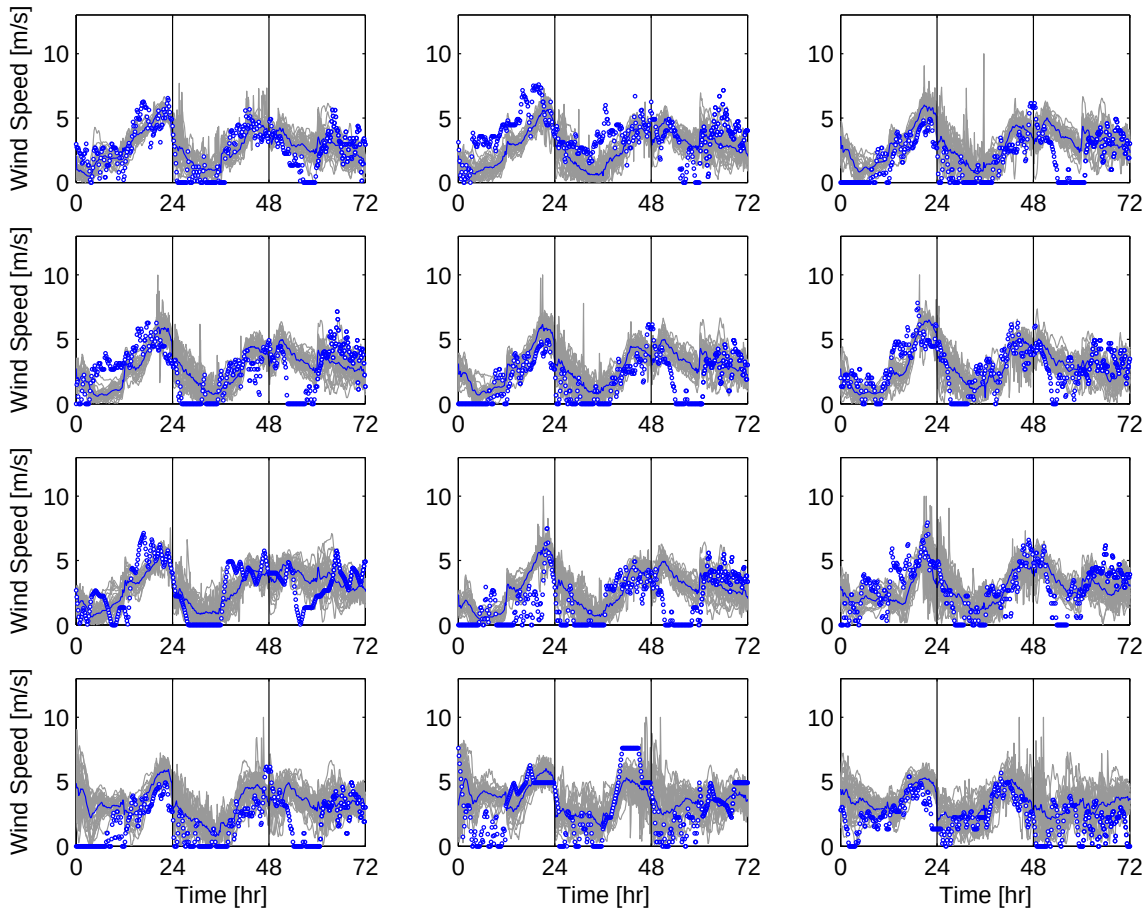


Figure 4: Wind speed realizations for 12 wind farm locations in Illinois at 10 m and observations (dots) at nearest meteorological stations. Vertical lines represent beginning of day (12:00 AM).

and 12 hours before the final (target) time. The higher the value, the more sensitive is the final-time target solution to the corresponding (past) state. In other words, the sensitivity at the final time (27) is propagated backwards 6 and 12 hours, and this approach gives a measure of influence of the initial condition on the final target state 6 and 12 hours ahead. A similar procedure was used to determine the size and shape of the finest domain (#3 in Figure 1).

The adjoint consistency can be tested by using finite-difference approximations or the linear algebra properties of the transpose operator, as described below.

Finite Differences. To perform this test, we perturb the initial state at some point $\hat{x}_i^{t_0} \leftarrow x_i^{t_0} + \varepsilon$ and look at the effect caused in the cost function Ψ at some nearby point $\Psi(\hat{x}_j^{t_F}) - \Psi(x_j^{t_F})$. The finite difference $\frac{\Psi(\hat{x}_j^{t_F}) - \Psi(x_j^{t_F})}{\varepsilon}$ can be compared directly to the adjoint sensitivity of the same cost function. We found that our results converge with decreasing

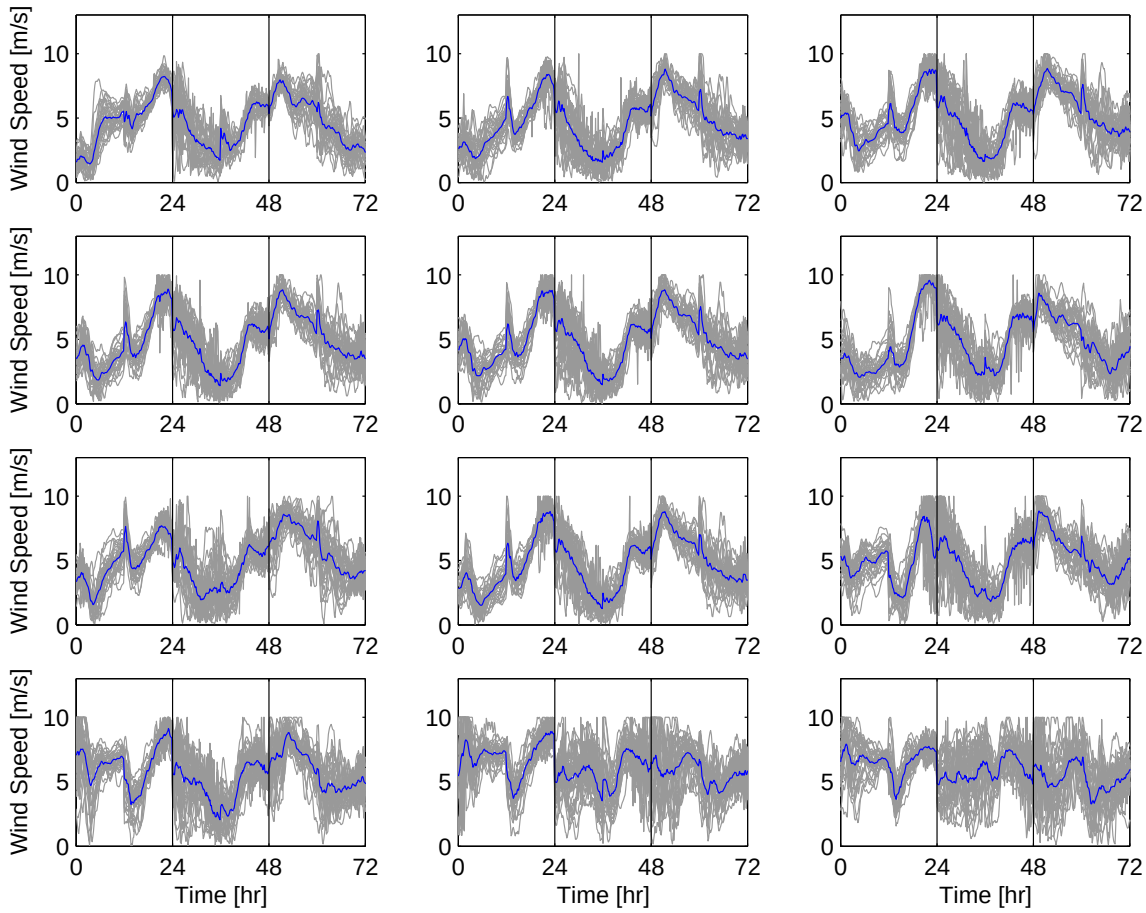


Figure 5: Wind speed realizations for 12 wind farm locations in Illinois at 100 m. Vertical lines represent beginning of day (12:00 AM).

ε. The results are shown in Table 2.

Table 2: Convergence of the adjoint solution to the finite differences approximation.

ε [%]	15	3	1
Error [%]	21	7	4

TL-AD Test. The second test was proposed by [17] and is based on the following observation:

$$\left(\mathbf{M}x^{t_0}\right)^T\left(\mathbf{M}x^{t_0}\right) = \left(x^{t_0}\right)^T\mathbf{M}^T\mathbf{M}x^{t_0} = \left(x^{t_0}\right)^T\mathbf{M}^T x^{t_F} = \left(x^{t_0}\right)^T\lambda^{t_0}\,, \tag{28a}$$

$$= \left(x^{t_F}\right)^T x^{t_F}\,, \tag{28b}$$

where $\mathbf{M} = \frac{\partial\mathcal{M}}{\partial x}$. Thus, from $(\mathbf{M}x^{t_0})^T(\mathbf{M}x^{t_0})$ we can evolve the initial state x^{t_0} forward to x^{t_F} twice using the tangent linear model ($\mathbf{M}$) and then take the inner product (28a), or we

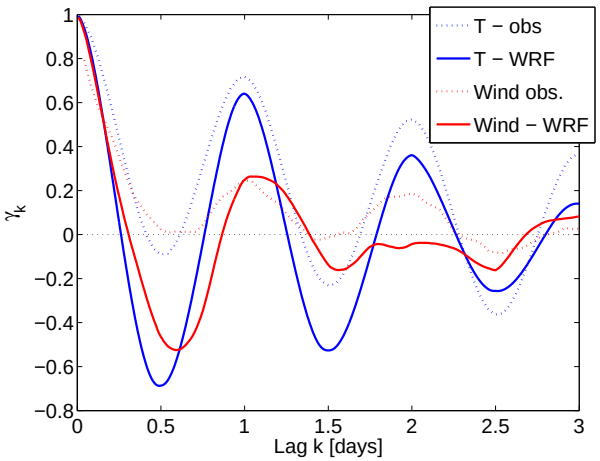


Figure 6: Correlogram for the wind and temperature measurements at Peru, IL.

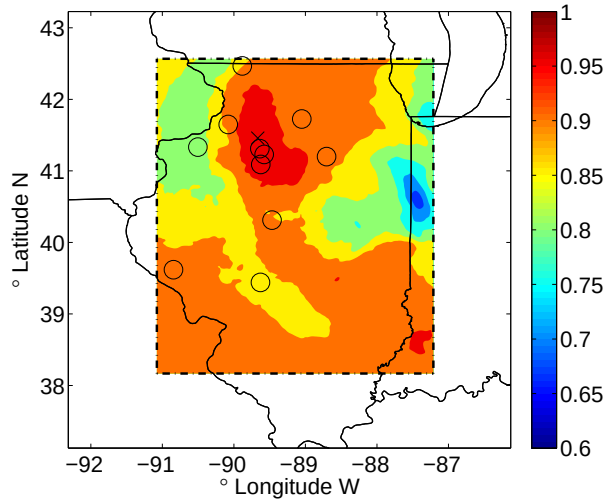


Figure 7: The spatial correlation for the wind field for wind farm #8 on June 5, 1:50 AM, denoted by "X." The circle markers denote the other wind farms in Illinois.

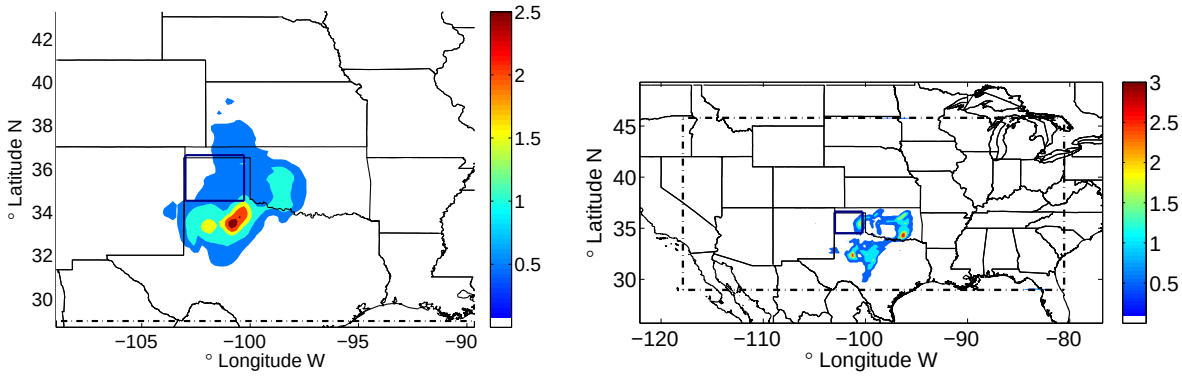


Figure 8: The wind speed sensitivity in a northern Texas region (solid line caret) with respect to the wind speed with 6 hours zoomed-in (left) and 12 hours (right) before the final time accumulated on all height levels.

can evolve it forward once and then use the adjoint operator ($\mathbf{M}^T$) to evolve it backward and take its inner product with x^{t_0} (28b). These two expressions should evaluate to the same value within roundoff. As a test, we ran WRFPlus in TL and AD mode on an initial state for a two-hour simulation forth and back. We obtained the same value to within a factor of 10^{-5}, a deviation that can be accounted for by the slightly different paths that the tangent linear and adjoint operators take as they integrate because of chaining in multiple operators and the differences between the full nonlinear model and the simplified model.

Implementation Considerations

WRF is an open-source weather and climate model available to the research community. In this study we used the latest version available at the time of preparing this manuscript; that is, version 3.1 [23]. In what follows we provide some implementation considerations arising during the execution of WRF in an operational setting.

The ensemble approach taken for estimating the uncertainty in the weather system is highly parallelizable because each scenario evolves independently through WRF, once the initial ensemble has been generated. The most expensive computational element is the evolution of each member through the WRF system. We therefore consider a two-level parallel implementation scheme. The first level is a coarse-grained task decomposition represented by each individual member. A secondary finer-grain level consists in the parallelization of each individual member. This approach yields a highly scalable solution.

The simulations were performed on the Jazz Linux cluster at Argonne National laboratory. Jazz has 350 compute nodes, each with a 2.4 GHz Pentium Xeon with 1.5 GB of RAM. The cluster uses Myrinet 2000 and Ethernet for interconnect and has 20 TB of on-line storage in PVFS and GFS file systems. Our actual running times for a single 24-hour job on Jazz are illustrated in Figure 9. We note that solution times can vary by a factor of 2 throughout the day. Consequently, to make this run a one hour of real time, one would need on the order of 500 CPUs or cores of a similar power. Additional computational power may be required to compensate for increasing cost of disk access.

CPUs	Wall-time [hr]
4	50
8	28
16	17
32	10

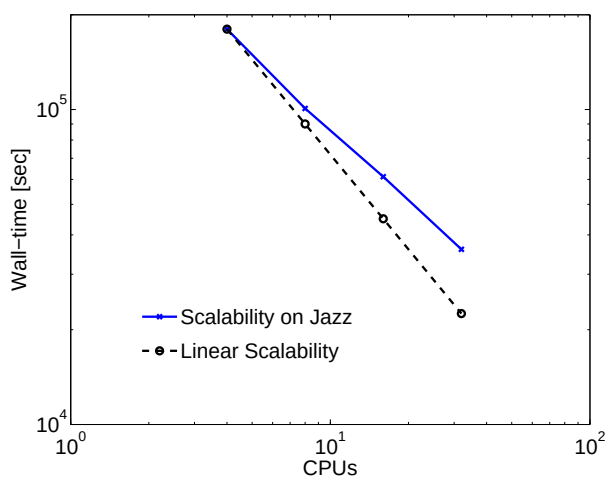


Figure 9: Scalability of WRF on the compute cluster Jazz for 24 hours (extrapolation based on 2-hour runs).

4.2 Economic Study Unit Commitment/Energy Dispatch

Because of the lack of detailed design data of thermal and wind power units in the open literature, we have constructed an artificial simulation study. We first describe the thermal and wind power assumptions used and then discuss our results from the simulation.

Power System Description

The thermal power system specifications used in this work are based on those reported in [5]. The system contains a total of 10 thermal generators with a total installed capacity of 1662 MW. The peak demand is 1326 MW. The ramp limits of the units are not reported, so we have assumed them to be 50% of the corresponding maximum capacity. The reserve requirements are assumed to be 10% of the demand. To simulate increasing level of wind power adoption, we increase the number of wind turbines at each of the 12 wind farm locations in Illinois.

Results

To generate wind power forecasts, we propagate the real wind speed observations and the realizations from WRF at a height of *10 meters* through a typical wind power curve with a maximum capacity of 1.5 MW. The nominal curve has a cut-in speed of 3 meters per second and reaches the rated capacity at 12 meters per second. The wind speed observations, forecast and ensembles used are summarized in Figure 4. As previously mentioned, we used the height of 10 meters because the NCDC data used for validation is only reported at this level. In addition, most wind power data is proprietary and difficult to obtain. As expected, the wind speeds are relatively low at this level, thus leading to small power outputs. Instead of using the wind speed WRF forecasts at 100 meters, we have kept the 10 meters WRF forecasts and observations and mapped these using a shifted power curve obtained by displacing the nominal cut-in speed from 3 to 2 meters per second. With this, the rated capacity is reached at around 11 meters per second. This strategy allowed more realistic validation results for wind power compared to linear interpolation of the wind speed observations. The resulting artificial wind power realizations and observations are presented in Figure 10. The order of the windows goes from left to right and coincides with the wind farm location numbering shown in Figure 2. The wind power distribution is clearly affected by the nonlinear structure of the power curve. The WRF realizations are able to encapsulate the actual power observations. The largest differences are observed at the beginning of the third day. Unfortunately, the rounding of the NCDC wind speed data significantly increases the errors when mapped through he power curve. Nevertheless, in general, the overall trends are captured fairly accurately. We emphasize that the uncertainty structure of the wind power forecasts strongly depends on the particular characteristics of the installed system.

We have run the closed-loop UC/ED system assuming a rolling horizon and a forecast frequency of 24 hours. The energy dispatch problem runs every hour. A total of 30 WRF realizations are used to solve the stochastic problem. The resulting MILP problems are implemented in AMPL and are solved with the CBC solver from the COIN-OR repository. The MILP contains 38,651 variables from which 240 are binary, 783 equality constraints, and 40,747 inequality constraints. CBC is used with its default algorithm settings but in multi-thread mode with four threads. The average solution time for the full UC problem in a quad-core Intel processor running Linux is about 9 minutes in cold-start mode. The solution time of the energy dispatch problem is less than 10 seconds.

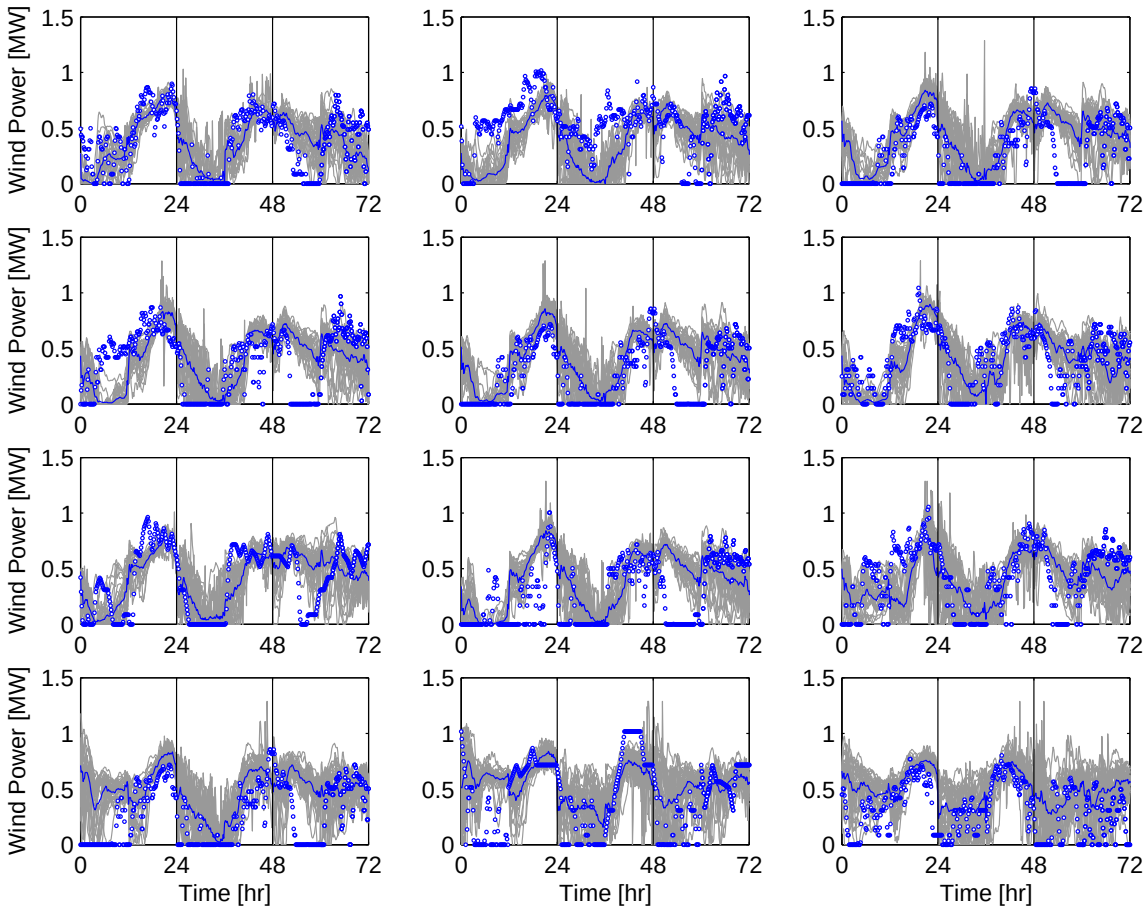


Figure 10: Wind power realizations for 12 different wind farm locations in Illinois at 10 m and observations (dots) at nearest meteorological stations. Vertical lines represent beginning of day (12:00 am).

The results for the 20% penetration study are presented in Figures 11-13. In Figure 11, we present the optimal commitment profiles for the closed-loop optimizer and for the perfect information problem. As can be seen, the profiles are similar but the generators tend to be ON more with the stochastic policies. We note that the optimal cost of the stochastic strategy over three days of operation is only about 1% larger than with the perfect information strategy. In Figure 12, we present the policies for the thermal power levels. We notice that the sensitivity of the power levels of some units to the uncertainty of the wind power is very small. Generators #2 and #5 are the most sensitive, while generators #3 and #4 exhibit no sensitivity. We have found that the sensitivity levels depend strongly on the design characteristics and prices of the generators.

We also performed an inference analysis using the weighting sampling strategy of Section 3.4 for the first day of operation using $M = 30$ different batches. The upper bound mean was found to be $U_{N_S,M} = \$474,064$ with variance $s_U^2(M) = 1,082\$^2$. The lower

bound mean was found to be $L_{N_S,M} = \$474,317$ with variance $s_L^2(M) = 1,656\$^2$. Both variances are less than 0.25% of the mean cost. This value indicates that 30 WRF realizations are sufficient to estimate the optimal cost. We have also found that updating the WRF forecast every 12 hours instead of every 24 hours does not bring important economic benefits. The explanation resets with the forecast trends presented Figure 3. Note that the forecasts are not improved significantly at the middle of the day, perhaps because measurements assimilated during the day are not as informative as those assimilated during night, where the wind currents tend to be stronger.

In Figure 13 we present the profiles of total aggregated demand, thermal power, and wind power. The thin gray lines represent the scenarios foreseen by the stochastic optimizer at the beginning of each day. As can be seen, the realizations are able to encapsulate the actual closed-loop profiles (solid lines) during the first two days. As a result, the optimizer is always able to satisfy the load, even for an adoption level of 20%. In the third day, however, we see a significant mismatch between the forecasted wind power and the realized one in the first 12 hours of operation. This directly translates to a mismatch in the planned closed-loop behavior of the thermal power levels. In this case, the reserves are sufficient to satisfy the load. Nevertheless, this result is important to illustrate that modeling the probability distribution in an adequate manner is critical. This effect could potentially be ameliorated by inflating the initial conditions of the WRF ensembles. However, this effect cannot be predicted *a priori*, and computational limitations preclude running the WRF model more frequently to capture this mismatch. This situation suggests that a higher frequency forecaster (e.g., autoregressive model) or an artificial variance corrector should be added to the system. How to connect this short-term forecaster with WRF is still not clear.

We found that a purely deterministic optimizer (using only the WRF forecast mean) is not able to sustain adoption levels of more than 10% even with the allocated reserves. We also observed that increasing the adoption levels increases the start-up and shutdown costs, but these are negligible (on the order of $10,000) with respect to the total production costs.

5 Conclusions and Future Work

We presented a computational framework for the integration of the state-of-the-art Weather Research and Forecasting (WRF) model in stochastic unit commitment/energy dispatch formulations that account for wind power uncertainty. We extended the WRF model with a sampling technique implemented in a distributed-memory parallel computing architecture and perform an adjoint sensitivity analysis on the wind speed. We have used the uncertainty information in a stochastic unit commitment formulation to analyze the impacts of wind power uncertainty on the system's economic performance. In addition, we have developed a weighting-average resampling strategy that avoids expensive WRF simulations to perform inference analysis. Our numerical experiments indicate that it is relatively costly to generate forecast and uncertainty information from WRF at a lower frequency than 12 hours. We have not accounted for the cost of assimilating observations

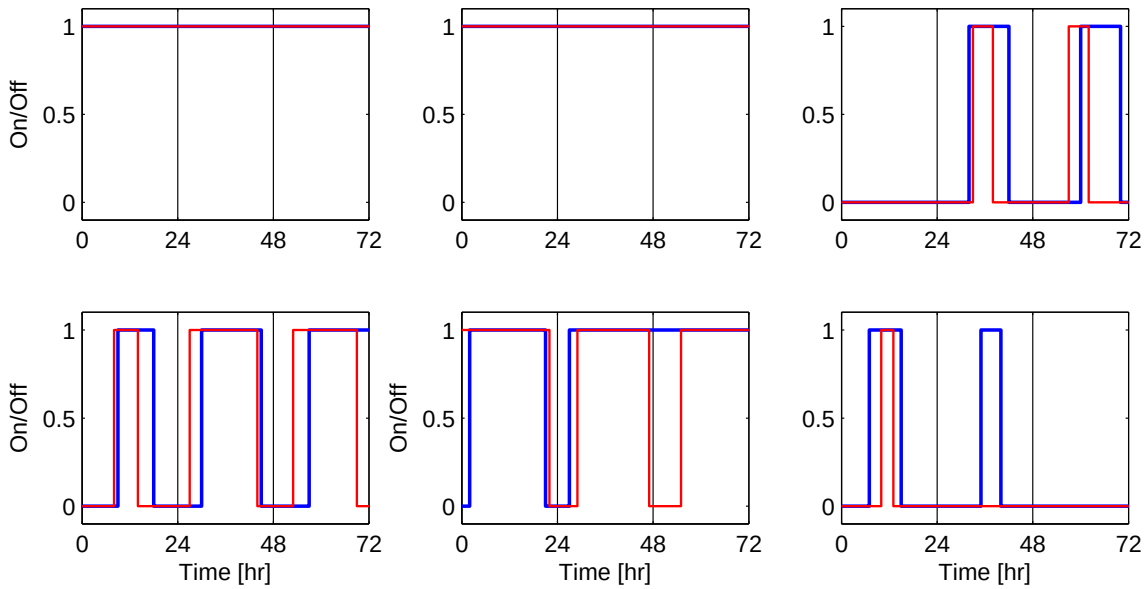


Figure 11: Closed-loop commitment profiles for thermal units. Solid thin line is optimal profile (with perfect information); solid thick line is stochastic UC solution.

in the model. However, given the power of current compute clusters, this does not seem to be a limiting factor.

Our simulated stochastic commitment study indicates that using WRF forecasts and uncertainty information is critical to achieve high adoption levels with minimum reserves. In this study, however, we have not found significant benefit of updating the WRF forecasts every 12 hours, as opposed to 24 hours. The benefits of updating the forecast more frequently is an issue of ongoing research. One must keep in mind, however, though the high computational cost attached to such frequent updates. Our study illustrates a real operational setting, and points to several issues and limitations that are not present in idealized experiments using artificial uncertainty information. Our inference strategy shows that a moderate (N=30) number of WRF realizations is sufficient to produce a good estimate of the optimal cost.

As future work, we are interested in developing techniques to generate forecasts at a higher resolution. In addition, we are interesting in generating wind-power forecast models by fusing WRF wind speed forecasts and operational wind-power data. In addition, we are interested in exploring real-time optimization strategies to reduce the on-line computational times. We are also interested in exploring the interplay between wind power uncertainty and the power system design and price characteristics. We will extend our formulation to consider detailed network constraints where the effects of wind power are more pronounced.

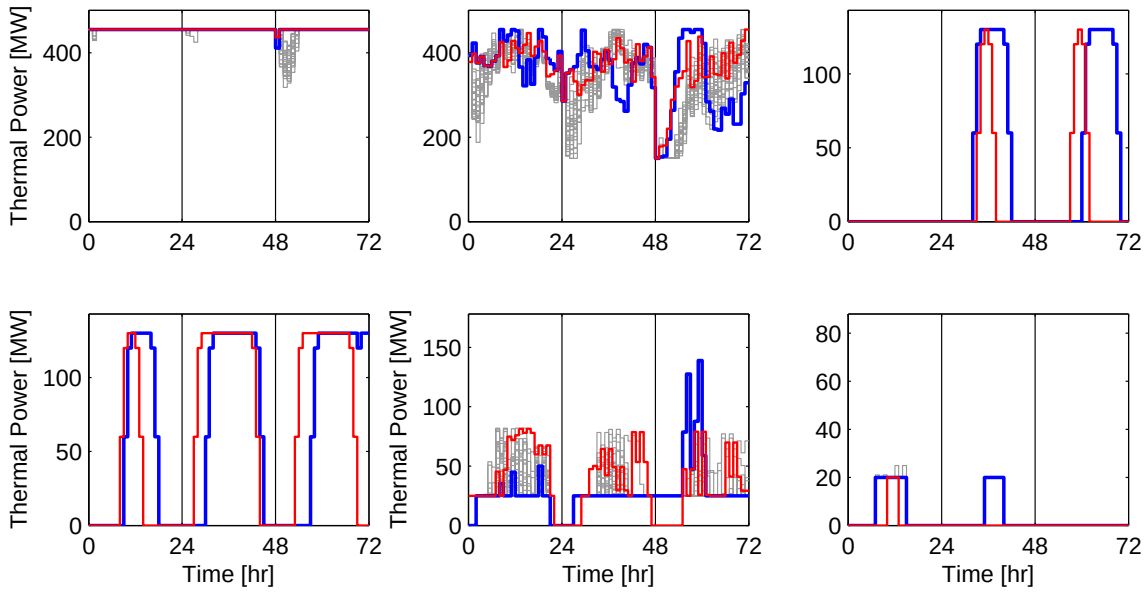


Figure 12: Closed-loop thermal profiles for thermal units. Solid thin line is optimal profile (with perfect information), solid thick line is stochastic UC solution, and thick gray lines are planned scenarios at the beginning of each day.

Acknowledgments

This work was supported by the Department of Energy, through Contract No. DE-AC02-06CH11357. We thank Argonne National Laboratory's Laboratory Computing Resource Center (LCRC) and John Valdes for helping us to set up and run WRF on the Jazz compute cluster.

References

[1] *20% wind energy by 2030: Increasing wind energy's contribution to U.S. electricity supply*, Tech. Rep. DOE/GO-102008-2567, Office of Energy Efficiency and Renewable Energy, 2009.

[2] D. Cacuci, *Sensitivity theory for nonlinear systems I. Nonlinear functional analysis approach*, Journal of Mathematical Physics, 22 (1981), pp. 2794–2802.

[3] ——, *Sensitivity theory for nonlinear systems. II. Extensions to additional classes of responses*, Journal of Mathematical Physics, 22 (1981), pp. 2803–2812.

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

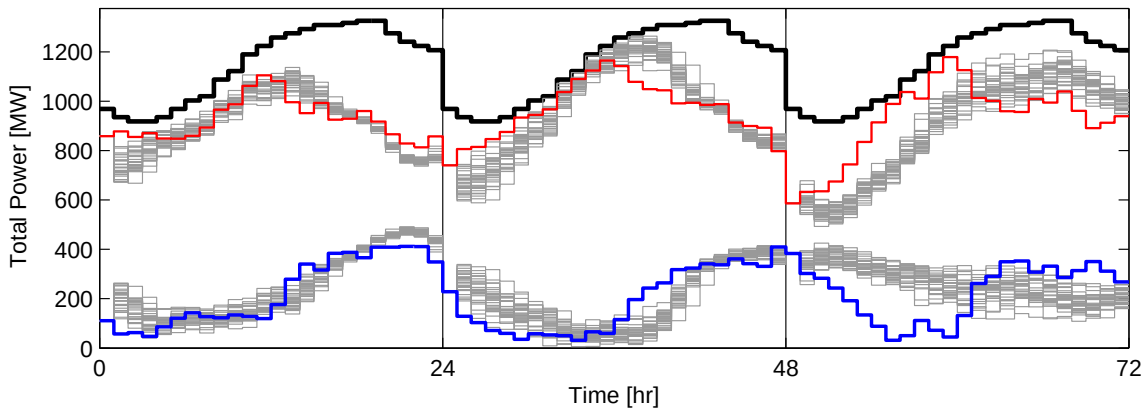


Figure 13: Closed-loop total power profiles obtained with stochastic UC formulation. Top thick line is demand profile, medium thick line is the implemented thermal profile, gray lines are planned realizations at beginning of each day, bottom thick line is actual total wind power, and the adjacent gray lines are forecasted profiles.

[5] M. Carrion and J. M. Arroyo, *A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem*, IEEE Transactions on Power Systems, 21 (2006), pp. 1371–1378.

[6] T. Chai, G. Carmichael, Y. Tang, A. Sandu, M. Hardesty, P. Pilewskie, S. Whitlow, E. Browell, M. Avery, P. Nedelec, J. Merrill, A. Thompson, and E. Williams, *Four-dimensional data assimilation experiments with International Consortium for Atmospheric Research on Transport and Transformation ozone measurements*, Journal of Geophysical Research, 112 (2007), p. D12S15.

[7] R. Errico, *What is an adjoint model?*, Bulletin of the American Meteorological Society, 78 (1997), pp. 2577–2591.

[8] G. Giebel, R. Brownsword, and G. Kariniotakis, *The state-of-the-art in short-term prediction of wind power – A literature overview*, tech. rep., ANEMOS Deliverable Report D1.1, 2003.

[9] R. Giering and T. Kaminski, *Applying TAF to generate efficient derivative code of FORTRAN 77-95 programs*, PAMM, 2 (2003), pp. 54–57.

[10] J. Goez, J. Luedtke, D. Rajan, and J. Kalagnanam, *Stochastic unit commitment problem*, Tech. Rep. RC24713, IBM, 2008.

[11] J. Juban, L. Fugon, and G. Kariniotakis, *Uncertainty estimation of wind power forecasts*, in Proceedings of the European Wind Energy Conference EWEC08, Brussels, Belgium, 2008.

[12] 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 (1996), pp. 437–471.

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

[14] J. Marotzke, R. Giering, K. Q. Zhang, D. Stammer, C. Hill, and T. Lee, *Construction of the adjoint MIT ocean general circulation model and application to Atlantic heat transport sensitivity*, J. Geophys. Res., 104 (1999), pp. 29,52929,547.

[15] K. Methaprayoon, W. J. Lee, C. Yingvivatanapong, and J. Liao, *An integration of ANN wind power estimation into uc considering the forecasting uncertainty*, IEEE Trans. Ind. Appl., 43 (2007), pp. 116–124.

[16] C. Monteiro, R. Bessa, V. Miranda, A. Boterrud, J. Wang, and G. Conzelmann, *Wind power forecasting: State-of-the-art 2009*, tech. rep., INESC Porto and Argonne National Laboratory, 2009.

[17] I. Navon, *Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography*, Dynamics of Atmospheres and Oceans. Special issue in honor of Richard Pfeffer, 27 (1998), pp. 55–79.

[18] T. Palmer, G. Shutts, R. Hagedorn, F. Doblas-Reyes, T. Jung, and M. Leutbecher, *Representing model uncertainty in weather and climate prediction*, Annual Review of Earth and Planetary Sciences, 33 (2005), pp. 163–193.

[19] V. S. Pappala, I. Erlich, K. Rohrig, and J. Dobschinski, *A stochastic model for the optimal operation of a wind-thermal power system*, IEEE Transactions on Power Systems, 24 (2009), pp. 940–950.

[20] D. Parrish and J. Derber, *The National Meteorological Center's spectral statistical-interpolation analysis system*, Monthly Weather Review, 120 (1992), pp. 1747–1763.

[21] P. A. Ruiz, R. Philbrick, E. Zak, K. W. Cheung, and P. Sauer, *Uncertainty management in the unit commitment problem*, IEEE Transactions on Power Systems, 24 (2009), pp. 642–651.

[22] G. Sheble and G. N. Fahd, *Unit commitment literature synopsis*, IEEE Transactions on Power Systems, 9 (1994), pp. 128–135.

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

[24] A. Tuohy, P. Meibom, E. Denny, and M. O'Malley, *Unit commitment for systems with significant wind penetration*, IEEE Transactions on Power Systems, 24 (2009), pp. 592–601.

[25] C. Wang and S. M. Shahidehpour, *Effects of ramp-rate limits on unit commitment and economic dispatch*, IEEE Transactions on Power Systems, 8 (1993), pp. 1341–1350.

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

[27] Y. Xiao, M. Xue, W. Martin, and J. Gao, *Development of an adjoint for a complex atmospheric model, the ARPS, using TAF*, Lecture Notes in Computational Science and Engineering, 50 (2006), p. 263.

[28] V. Zavala, E. Constantinescu, T. Krause, and M. Anitescu, *On-line economic optimization of energy systems using weather forecast information*, To appear in the Journal of Process Control, (2009).

[29] L. Zhang, E. Constantinescu, A. Sandu, Y. Tang, T. Chai, G. Carmichael, D. Byun, and E. Olaguer, *An adjoint sensitivity analysis and 4D-Var data assimilation study of Texas air quality*, Atmospheric Environment (ACM Issue), 42 (2008), pp. 5787–5804.

Mathematics and Computer Science Division
Argonne National Laboratory
9700 South Cass Avenue, Bldg. 240
Argonne, IL 60439-4847

www.anl.gov

Argonne National Laboratory is a U.S. Department of Energy
laboratory managed by UChicago Argonne, LLC