Adjoint Sensitivity Analysis for Numerical Weather Prediction: Applications to Power Grid Optimization

Alexandru G. Cioaca, Victor M. Zavala *Member, IEEE*, and Emil M. Constantinescu

***Abstract*—We present an approach to estimate adjoint sensitivities of economic metrics of relevance in the power grid with respect to physical weather variables using numerical weather prediction models. We demonstrate that this capability can significantly enhance planning and operations. We illustrate the method using a large-scale computational study where we compute sensitivities of the regional generation cost in the state of Illinois with respect to wind speed and temperature fields inside and outside the state.**

***Index Terms*—adjoint sensitivity analysis, numerical weather prediction, WRF, power grid planning, sensor siting.**

I. Introduction

Weather forecasts are essential in estimating electricity demand and renewable power generation [1], [2]. Weather forecasts can be obtained by using data-based models (e.g., autoregressive (AR), artificial neural networks (ANN), Gaussian process (GP) models), physics-based numerical weather prediction (NWP) models, or a combination of both. In previous work [2] we showed that NWP models are superior in producing weather forecasts and uncertainty information since they can capture complex physical spatio-temporal phenomena over wide geographical regions that cannot be captured by using data-based models in isolation. In addition, we proposed computational strategies to make NWP models practical for power grid planning and operations.

In this work, we demonstrate that NWP models can also provide valuable information about the effect of physical weather variables on power grid economic performance. In particular, we present an adjoint sensitivity analysis (ASA) method. ASA is used to determine the sensitivity of a model state or parameter (e.g., future 3D wind speed field) with respect to input states (e.g., current fields of ambient conditions). In the context of wind power generation, we show how ASA can be used to determine simulation domain size and resolution, to identify power grid variables and locations that should be monitored more closely, and to determine suitable locations for sensor and wind farm placement. Furthermore, we discuss how adjoint analysis can provide information to guide the development of low-complexity, data-based AR/GP/ANN models.

A. Cioaca is with the Department of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061. e-mail: alexgc@vt.edu.

E. M. Constantinescu and V. M. Zavala are with the Mathematics and Computer Science Division at Argonne National Laboratory, Argonne, IL 60439. e-mail: emconsta@mcs.anl.gov, zavala@mcs.anl.gov.

As an example of the power grid operation, consider the following economic cost function:

$$\Psi(t) = c(t) + \lambda(t)^T \omega(w(t)), \tag{1}$$

where $c(t)$ is the regional cost of the power grid at time t, $\omega(\cdot)$ is a wind power function at different grid buses, $w(t)$ is the wind speed field, and $\lambda(t)$ are cost coefficients representing power grid economic performance as a function of wind power supply.

Using this economic cost function, we can calculate the following *sensitivity*:

$$\mathbb{S}(\Psi) = \frac{\partial \Psi(t)}{\partial \mathbf{W}(\mathbf{t})}, \tag{2}$$

where $\mathbf{W}(t)$ are the 3-D weather spatial fields at time t, which include temperature, wind directions, solar radiation, and so on. As can be seen, different weather variables such as temperature can influence wind speed; hence, the sensitivity structure is not obvious. Computing sensitivity information is challenging from a computational point of view because of the complexity of NWP models.

The paper is structured as follows. In Section II we describe the basic capabilities of the NWP model WRF. In Section III we provide a brief mathematical presentation of ASA. In Section IV we discuss how to construct economic cost functions from complex optimization problems such as economic dispatch and optimal power flow using optimization sensitivity capabilities. In Section V we illustrate the developments using a large-scale study. The paper closes with conclusions and directions for future research.

II. Numerical Weather Prediction

In this section, we describe the procedures used to forecast ambient conditions using the Weather Research and Forecasting (WRF) model. The WRF model [3] is a state-of-the-art numerical weather prediction system designed to serve for both operational forecasting and atmospheric research. WRF is the result of a multiagency 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 be as fine as 30 seconds of a degree (less than 1 km^2).

To initialize the NWP simulations, we use reanalyzed fields, that is, simulated atmospheric states reconciled with observations (i.e., using data assimilation), because the entire atmospheric state space is required by the model as initial conditions

whereas only a small subset of the state space is available through measurement at any given time [4]. 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. For more details, please refer to [5].

III. Adjoint Sensitivity Analysis

Consider a numerical model $\mathcal{M}$ that evolves an initial state x_{t_0} to a given final time t_N (e.g., 24 hours ahead):

$$x^k = \mathcal{M}\left(t_{k-1}, x^{k-1}, p\right)\ ,\ x^0 = x_i(t_0, p)\ ,\ k \in \{1 \dots N\}\,, \tag{3}$$

where p are model parameters. For instance, $\mathcal{M}$ may represent the discretization operator of a partial differential equation. Sensitivity analysis reveals how a model solution is affected by small perturbations in the model variables and parameters [6]. 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, $S_i(t) = \frac{\partial x(t)}{\partial p_i}\frac{p_i}{x(t)}$ [7]. Just as the model state x^{t_0} is evolved through $\mathcal{M}$, the sensitivity S_i is evolved by the gradient (also known as tangent linear) model

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

where $t_k \in [t_0, t_N]$. 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_N}. Therefore, the sensitivity takes the form

$$S_i = \frac{\partial x^{t_N}}{\partial x_i^{t_0}}\frac{x_i^{t_0}}{x^{t_N}}\,. \tag{4}$$

and its evolution is described by

$$S_i^k = \frac{\partial \mathcal{M}}{\partial x}(x^{t_{k-1}})S^{k-1}\,,\ S_i^0 = \frac{\partial x^{t_0}}{\partial x_i^{t_0}}\,.$$

This is useful if one is interested in the effect a small perturbation at a single source location would have on the future states at multiple locations, x^{t_k}. Alternatively, one could consider the inverse or *adjoint process* [8] of observing some target state in the state space at future times and inferring what states in the initial conditions have a strong influence on that target state [7]. We aim to find 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, that is, a function of the state at the final time,

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

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

$$\frac{\partial \Psi(x^{t_N})}{\partial x_i^{t_0}} = \frac{\partial \Psi(x^{t_N})}{\partial x^{t_N}}\frac{\partial x^{t_N}}{\partial x^{t_0}} = \frac{\partial \Psi(x^{t_N})}{\partial x^{t_N}}S_i^{t_N}\,. \tag{5}$$

Following [7], one can extend (5) for all time indices

$$\frac{\partial \Psi(x^{t_N})}{\partial x_i^{t_0}} = \frac{\partial \Psi(x^{t_N})}{\partial x^{t_N}}\frac{\partial x^{t_N}}{\partial x^{t_{F-1}}}\cdots\frac{\partial x^{t_1}}{\partial x^{t_0}}\frac{\partial x^{t_0}}{\partial x_i^{t_0}}\,,$$
$$\frac{\partial x^{t_k}}{\partial x^{t_{k-1}}} = \frac{\partial \mathcal{M}}{\partial x}(x^{t_{k-1}})\,,\ \frac{\partial x^{t_0}}{\partial x_i^{t_0}} = \delta_i x^{t_0}\ ;\ i \in \{1 \dots M\}\,.$$

Alternatively, by transposing, the adjoint process evolves the sensitivity in reverse order:

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

If the following equations are satisfied [7]:

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

One can show that the adjoint variables or influence functions λ^{t_k} [8] represent the gradients of the cost function with respect to perturbations in the state at earlier times $\lambda^{t_k} = \left(\frac{\partial \Psi(x^{t_N})}{\partial x^{t_k}}\right)^T = \nabla_{x^{t_k}}\Psi(x^{t_N})$. 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* $M^* = \left(\frac{\partial \mathcal{M}}{\partial x}\right)^T$ back to the initial time. As we did in Equation (4), we can also consider the scaled adjoint sensitivity $\widehat{\lambda}$, which can be physically interpreted as the percentage change in the cost function when the variable $x_i^{t_k}$ is changed:

$$\widehat{\lambda}_i^{t_k} = \frac{\partial \Psi(x^{t_N})}{\partial x_i^{t_k}}\frac{x_i^{t_k}}{\Psi(x^{t_N})}\,. \tag{6}$$

Henceforth, we will use $\widehat{\lambda}$ to denote adjoint variables in order to distinguish them from the Lagrange multipliers introduced in the next section.

Large sensitivity values indicate areas of influence, that is, locations where errors or perturbations in the current state (e.g., due to limited sensors) will produce significant changes in the target sites and time as described through the cost function. This is significant because one can, for instance, assess the effect of uncertainty of a particular location on future weather fields.

We illustrate the ASA method on a real test case and employ the WRF model [3], described in Sec. II, which will take the place of $\mathcal{M}$ (see [2] for implementation details). A simplified WRF model has been run through a source-to-source program called Transformation of Algorithm in FORTRAN (TAF) to automatically produce both gradients (M) and adjoints of the gradients (M^*) [9].

IV. Power Grid Optimization

Adjoint sensitivity analysis provides a powerful framework to assess the effect of weather conditions and uncertainty on infrastructures such as the power grid or natural gas networks. In this section, we discuss how to use adjoint analysis in conjunction with optimization sensitivity capabilities to evaluate the effect of spatio-temporal weather patterns on the grid

economic performance over a given geographical region. This information is vital in planning exercises such as transmission/generation expansion and sensor placement since weather drives electricity and natural gas markets (e.g., demands are strongly correlated to ambient temperature) and since weather patterns are becoming more relevant as the share of wind and solar generation is increasing.

Consider the optimization problem

$$\min_{z} \ f(z,\eta) \tag{7a}$$

$$c(z,\eta) = \omega, \quad (\lambda). \tag{7b}$$

where $z \in \Re^{n_z}$ are the decision variables, $\eta \in \Re^{n_\eta}, \omega \in \Re^{\omega}$ are parameters, and $\lambda \in \Re^{m}$ are Lagrange multipliers. The objective function $f : \Re^{n_z} \times \Re^{n_\eta} \to \Re$ and constraint functions $c : \Re^{n_z} \times \Re^{n_\eta} \to \Re^{m}$ are differentiable. The following results can be easily extended to include inequality constraints. The optimization problem can represent different problems, such as economic dispatch, optimal power flow, and transmission/generation expansion, which are parameterized in quantities such as demands and renewable supply (η, ω). These quantities are in turn affected by weather conditions. The following are well-known results of optimization sensitivity.

Theorem 1: Consider that a solution $x_*(\eta,\omega_0), \lambda_*(\eta_0,\omega_0)$ of problem (7) satisfies the linear independence constraint qualification [10]. Then, the multipliers $\lambda_*(\eta_0,\omega_0)$ are unique, and

$$\lambda_*(\eta_0,\omega_0) = \frac{\partial f}{\partial \omega}(z_*(\eta_0,\omega_0),\eta_0).$$

Consequently, up to first order, we have that

$$f(z^*(\eta_0,\omega),\eta_0) \approx f(z^*(\eta_0,\omega_0),\eta_0) + \lambda_*(\eta_0,\omega_0)^T(\omega-\omega_0). \tag{8}$$

We can use this linear function to estimate the effect ω on the cost function as described in Equation (1). Theorem 1 is a basic result of practical significance since Lagrange multipliers are used to establish market prices (e.g., locational marginal prices). Most optimization solvers provide information about Lagrange multipliers since these are computed as part of the solution procedure. We also have the following result.

Theorem 2: Consider that a base solution $x_*(\eta_0,\omega_0), \lambda_*(\eta_0,\omega_0)$ of problem (7) satisfies the linear independence constraint qualification and the strong second-order conditions [11]. Then, the base solution is locally unique, and the following sensitivity matrices exist:

$$\frac{\partial z_*}{\partial \eta}(z_*(\eta_0,\omega_0),\eta_0), \ \frac{\partial \lambda_*}{\partial \eta}(z_*(\eta_0,\omega_0),\eta_0)$$
$$\frac{\partial z_*}{\partial \omega}(z_*(\eta_0,\omega_0),\eta_0), \ \frac{\partial \lambda_*}{\partial \omega}(z_*(\eta_0,\omega_0),\eta_0).$$

Furthermore, there exist nonempty neighborhoods around the base solution in which the solution $z_*(\eta,\omega), \lambda_*(\omega,\eta)$ is unique.

Following the same idea used in (8), we can use the sensitivity matrices to compute first-order estimates of $z_*(\eta,\omega), \lambda_*(\omega,\eta)$. Theorem 2 is a more general result that allows one to quantify the effect of parameter changes on the entire solution vector and not only on the cost function. This approach can be used to handle nonlinear cost functions to compute the sensitivity function $\Psi(\cdot)$ in (1). Some optimization solvers provide sensitivity matrices. Examples include CPLEX, Gurobi, and IPOPT [12], [13].

As an example of the above concepts, consider the following economic dispatch problem [14]:

$$\min \sum_{k=\ell}^{\ell+T} \sum_{j\in\mathcal{G}} c_j \cdot G_{k,j} \tag{9a}$$

$$\text{s.t. } G_{k+1,j} = G_{k,j} + \Delta G_{k,j}, \ k \in \mathcal{T}, j \in \mathcal{G} \tag{9b}$$

$$\sum_{(i,j)\in\mathcal{L}_j} P_{k,i,j} + \sum_{i\in\mathcal{G}_j} G_{k,i} = \sum_{i\in\mathcal{D}_j} D_{k,i} - \sum_{i\in\mathcal{W}_j} W_{k,i}, \ k \in \mathcal{T}, j \in \mathcal{B} \quad (\lambda_{k,j}) \tag{9c}$$

$$P_{k,i,j} = b_{i,j}(\theta_{k,i} - \theta_{k,j}), \ k \in \mathcal{T}, (i,j) \in \mathcal{L} \tag{9d}$$

$$0 \le G_{k,j} \le G_j^{max}, \ k \in \mathcal{T}, j \in \mathcal{G} \tag{9e}$$

$$|\Delta G_{k,j}| \le r^{max}, \ k \in \mathcal{T}, j \in \mathcal{G} \tag{9f}$$

$$|P_{k,i,j}| \le P_{i,j}^{max}, \ k \in \mathcal{T}, (i,j) \in \mathcal{L} \tag{9g}$$

$$|\theta_{k,j}| \le \theta_j^{max}, k \in \mathcal{T}, j \in \mathcal{B} \tag{9h}$$

$$G_{\ell,j} = \text{given}, \ j \in \mathcal{G}. \tag{9i}$$

The objective of this problem is to minimize the regional generation cost for given demand and renewable supply levels. Here, $\mathcal{G}$, $\mathcal{L}$, and $\mathcal{B}$ are the sets of generators, lines, and nodes/buses (intersections of lines) in the geographical region, respectively. $\mathcal{D}_j$ and $\mathcal{W}_j$ are the sets of demand and renewable supply nodes connected to bus j, respectively. The time horizon is given by the set $\mathcal{T} := \{\ell, ..., \ell+T\}$ starting at the time ℓ, where T is the horizon length. Variables $G_{k,j}$ are the generator supply levels for time instant k and bus j. Following a similar notation, $P_{k,j}$ are the transmission line power flows, $\theta_{k,j}$ are the voltage angles, $W_{k,i}$ are the renewable supply flows, and $D_{k,i}$ are the demand levels that are fixed *parameters*. Constraint (9c) is Kirchhoff's law, which holds at each time and bus and which balances flow across the network. The *Lagrange multipliers* of Kirchhoff's law, obtained in the solution of the optimization problem, are the locational marginal prices (LMPs) $\lambda_{k,j}$ for each time instant and node. In Figure 1 we present a typical LMP field for the Illinois grid averaged over an entire year of operation. Note that prices are heterogeneous across the network because of limited transmission capacity (e.g., transmission congestion). Regions with extremely high or low (even negative) prices normally indicate insufficient transmission capacity towards that region (i.e., there exists locational scarcity).

The price $\lambda_{k,j}$ indicates the sensitivity of the regional generation cost to changes in demand or renewable supply levels at time k and node j. This information can be used with the following scenarios,

- By focusing at the network nodes $j \in \mathcal{B}$ with existing renewable supply, one can assess the effect of variations

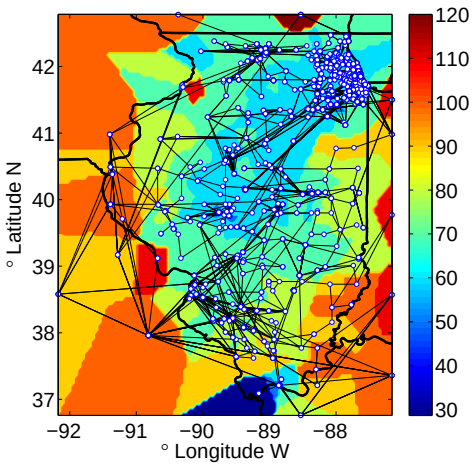


Fig. 1: Spatial field of locational marginal prices ($/MWh) under time-averaged conditions in Illinois grid.

of renewable supply (e.g., wind) on the regional generation cost under existing conditions. In particular, adjoint analysis permits one to assess the effect of the current state of weather variables, such as wind speed, on the future renewable power supply and demand and the regional cost. This assessment is important since most of the uncertainty in weather forecasts is associated with the uncertainty in the current state.

- By considering the hypothetical case in which all network nodes $j \in \mathcal{B}$ can be used to supply wind power, one can assess optimal locations for future renewable generators and the effect of renewable supply patterns on the regional generation cost.
- By examining the weather conditions around the geographical network, one can assess the effect of the prevailing weather conditions at a particular location (not necessarily inside the network) on the regional cost. Since weather fronts evolve slowly over large geographical regions, grid operators can use this information in advance. In addition, one can mitigate weather uncertainty by deploying meteorological stations at locations of maximum sensitivity. Moreover, regions of high sensitivity to weather can indicate transmission congestion and can thus pinpoint regions of high potential for transmission expansion.

If sensitivity matrices are available, the analysis can be extended to compute effects of particular variables of interest. For instance, one can assess the effect of variations of renewable supply directly on the locational marginal prices $\lambda_{k,j}$. This information can be used to expand transmission capacity and/or deploy meteorological stations to homogenize the prices across the network.

V. Large-Scale Numerical Study

In this section, we illustrate the adjoint sensitivity capabilities using the Illinois power grid system. The system comprises 1,900 buses, 2,538 transmission lines, 870 load nodes, and 261 generators. Our data consists of detailed specifications for the network topology, ramp and generation limits, fuel costs, and transmission lines. We analyze the effect of the initial conditions of wind speed on the regional generation cost. To do so, we have extracted locational marginal prices from an economic dispatch formulation for the system reported in [14].

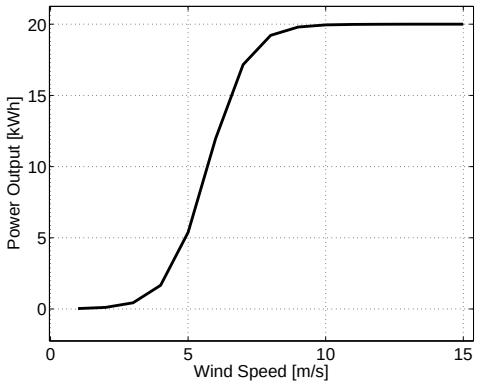


Fig. 2: Power curve function.

The key to using analysis methods from weather forecasts in conjunction with problems such as grid integration and planning is to find the appropriate link between the two frameworks. On the one hand we have access to a numerical weather prediction model that deals with physical quantities such as wind speed and temperature; on the other hand we have an optimization problem that deals with economic metrics and decisions. This study takes advantage of the relation between wind speed and the regional energy cost. The intermediate quantity between these two variables is wind power, which can be estimated directly from wind speed and the technical specifications of wind turbines.

The wind power function or curve represents the power output of a wind turbine as a function of wind speed. The power curve is characterized by a sigmoid profile, since the turbine does not produce any power when there is low wind, saturates at a certain value once the wind speed exceeds a certain threshold, and has a fast-growing activation area between these two extremes. For our tests, we constructed the following power curve starting from the hyperbolic tangent function, scaled and shifted to mimic real-life behavior:

$$\omega(w(t)) = 10(1 + \tanh(0.7w(t) - 4)). \tag{10}$$

The graph of this function is shown in Fig. 2. One can also determine this power function from historical data and regression models. We estimate the sensitivity of the time-dependent cost function Ψ, with respect to initial wind speed $w(t) = \sqrt{U^2 + V^2}$, where U and V are the W-E and S-N wind components, respectively. After defining (10), the sensitivity variables $(\widehat{\lambda})$ are computed as

$$\begin{aligned} \widehat{\lambda}^t &\leftarrow \widehat{\lambda}^t + \lambda(t)\frac{\partial \Psi}{\partial \alpha}\frac{\alpha}{\Psi} \\ &\leftarrow \widehat{\lambda}^t + \lambda(t)\frac{(0.7\,\alpha\,\mathrm{sech}(4 - 0.7w(t))^2)}{w(t)}\frac{\alpha}{1908\,\omega(w(t))}, \end{aligned} \tag{11}$$

where $k = N, \ldots, 1$, $\alpha = \{U, V\}$.

The first term represents the base cost, the second term is the partial derivative of the cost function to either one of the wind components at a certain time, and the third term is used for scaling. 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. The constant 1908 corresponds to the number of locations used in our study; for this experiment we used all the buses for which we had available data. The remaining constants are part of the power curve expression.

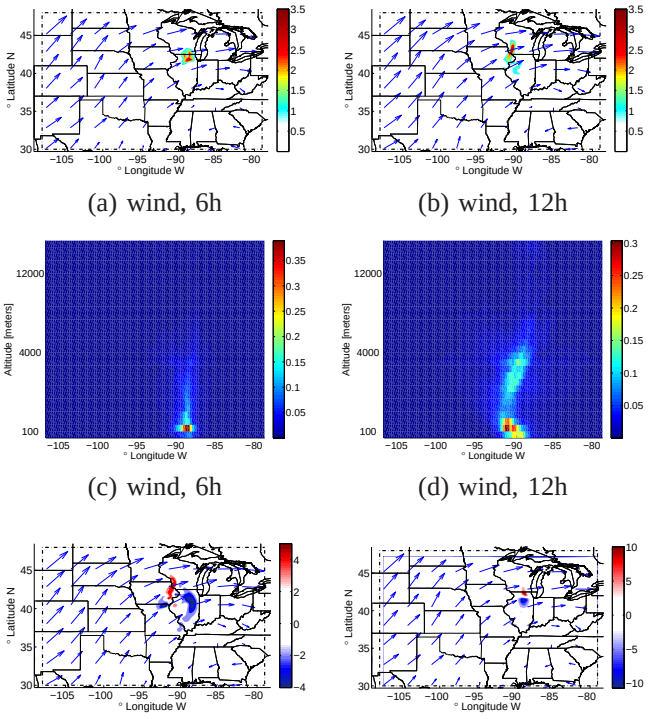
The computational part of the experiment that uses WRF is divided in two stages. The first stage performs a weather forecast for the time period of interest, in our case 24 hours. It is possible to use shorter or longer forecast windows, but the accuracy of the results tends to degrade after simulating 2-3 days. This is due to the high nonlinearity of the processes governing the atmosphere, which gives rise to a chaotic behavior, as well as due to uncertainties in initial conditions. The forecast model is configured to save the values of the variables of interest (wind speed components) at each point in time of interest. This process is called checkpointing, and in our experiments it takes place at every hour. In the second stage, we use the checkpointed values to compute the initial state of the adjoint model (corresponding to the final time of the forecast) and the intermediate adjoint forcing variables, using formula (11). We then run the adjoint model.

One simplifying assumption was made regarding the locations where the adjoint variables are initialized and forced. Since we are interested in wind power, one would usually be interested in studying sensitivities of the locations corresponding to the wind farms that currently produce energy. Since current wind adoption levels are low, we used the locations of the nodes. Consequently, this represents a planning scenario in which we seek to assess the effect of wind power injections at different nodes.

The adjoint model is initialized with potential perturbations at the locations of interest and propagates them backwards in time. After each one hour of simulation, another set of perturbations is forced into the adjoint model based on the computations performed during forecast, and the adjoint model continues to propagate the updated field farther back in time. At the end of the adjoint model run, the adjoint variables indicate areas to which the cost functional is sensitive.

For our tests, a serial version of WRF and its adjoint were compiled without shared- or distributed-memory parallelism capabilities. The compute server used for running the models operates on an 8-core Intel Xeon CPU clocked at 2.66 Ghz with 32 GB of RAM. The space discretization of the simulation domain was set at 25 kilometers (approx. 15 miles) for each grid point, while the time step was set at 150 seconds. Running the experiment took less than 1 hour using these parameters. We observed no clear benefits in the accuracy of our results when reducing the time step size to 30 seconds.

We employed WRF with real data and performed simulations



(a) wind, 6h (b) wind, 12h

(c) wind, 6h (d) wind, 12h

(e) temperature, 6h (f) temperature, 12h

Fig. 3: Regional cost sensitivity with respect to the wind speed (a-d) and potential temperature (e-f), 6 and 12 hours before the final time (June 11th, 2000 - 12:00)

on four scenarios. Two scenarios take place in June 2000, one on the 10th of the month and another on the 14th, both starting at 12:00 CST. The other two scenarios take place in October 2000, the 18th and 20th respectively, from 06:00 CST. While June is associated with a high demand of electricity, October is characterized by lower demands. We present two scenarios from each month in order to account for social factors such as energy demands during weekends (June 10) and business days (June 14, October 18, October 20). This approach, along with the different meteorological conditions of each scenario, leads to different regional cost sensitivities.

In Fig. 3.a and 3.b we show the evolution of the vertically integrated sensitivities with respect to wind, 6 and 12 hours in retrospect from the final forecast time. In other words, the sensitivity at the final time (11) is propagated backwards 6 and 12 hours and gives a measure of the influence of the initial condition on the final target state 6 and 12 hours ahead. The larger the value, the more sensitive is the final-time target solution to the current state. The high sensitivity regions illustrated in this study indicate areas with high impact on the future wind speed conditions. The integrated vertical profile of the sensitivities at 6 hours and 12 hours in retrospect, as seen from the south, are illustrated in Fig. 3.c and 3.d, respectively. This information can be interpreted in the following ways:

- From a numerical point of view, high-sensitivity regions indicate regions that need to be resolved accurately by

the NWP models. In other words, resolution should be increased in these regions.

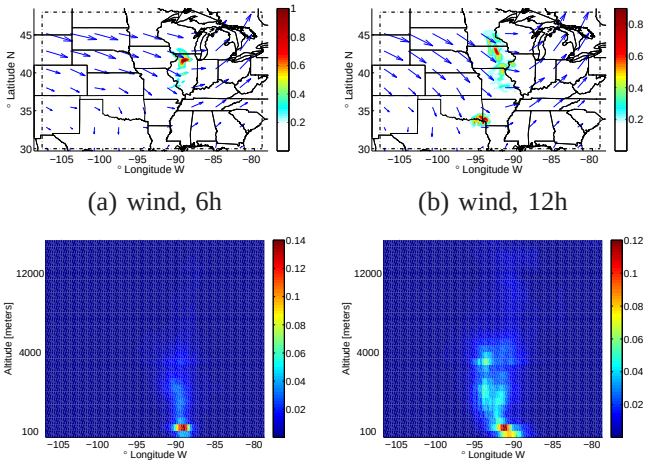
- From an uncertainty quantification point of view, high-sensitivity regions indicate locations where forecast errors have the largest impacts on the system. This information can be used to determine optimal locations for meteorological stations to mitigate this uncertainty.
- From a generation expansion point of view, regions of high sensitivity indicate locations where wind farms should be installed. In addition, these regions identify locations for natural gas generators to provide necessary ramping capacity. We note that adjoint sensitivity information can augment traditional resource maps used for wind farm planning since it is equally important to install wind generators in regions with high wind speeds but also with low uncertainty in neighboring regions.
- From a transmission expansion point of view, regions of high sensitivity that match regions of high price indicate nodes at which transmission congestion limits wind adoption.
- From a operational point of view, regions of high sensitivity indicate locations that will affect market clearing tasks such as unit commitment and economic dispatch due to higher sensitivity to uncertainty. Thus, if sensitivity information can be provided in advance to the ISOs, they can prepare to face high uncertainties of wind power variation by allocating reserves or by committing peaking units.

The sensitivity to wind illustrated in Figs. 3-6 demonstrate that on different days, different meteorological sources affect the target area. This analysis points to the dynamic size of the domain necessary for such a simulation to efficiently achieve accurate forecasts. We highlight the variability of these directions under different seasonal conditions. We observe that the highest sensitivity is consistently observed in the western part of the state as indicated by the wind directions. In addition, we note that regions of high sensitivity are not necessarily inside the network region of interest since wind fronts move across large geographical regions.

We also note the variability in the vertical column, a consequence of the fact that wind components are not two-dimensional but three-dimensional fields. The cost function at our target sites is influenced in a different manner by the evolution of wind at different height layers. We notice that the highest sensitivity is observed in the first layers, corresponding to the 100-4000 meter range. From a forecasting point of view, this indicates that a high uncertainty at these heights can translate into a large impact on cost. This is of significance since few sensors exist to measure conditions at those heights. From a generation point of view, the vertical sensitivity profiles indicate that the highest wind power generation is expected above 50 meters, as is the current practice.

In Fig. 3.e and 3.f we show the sensitivities with respect to the ambient temperature. This illustrates the effect of temperature fields on the cost function. Positive and negative values indicate locations in the temperature field that, if perturbed (or forecast incorrectly), increase or decrease the cost. This is important since it indicates that other physical variables can affect wind speed and thus regional cost. Consequently, care should be taken to mitigate uncertainty and forecast those variables accurately.

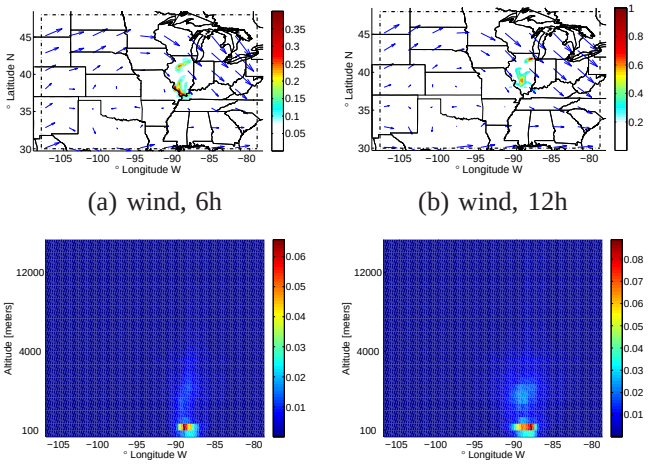
The two scenarios from June (see Figs 3 and 4) exhibit different dynamics, indicating a complex relationship between weather variables and the grid conditions. A similar conclusion can be drawn for the sensitivities computed in October (see Figs. 5 and 6).



(a) wind, 6h (b) wind, 12h

(c) wind, 6h (d) wind, 12h

Fig. 4: Regional cost sensitivity with respect to wind speed, 6 and 12 hours before the final time (June 15, 2000 - 12:00)



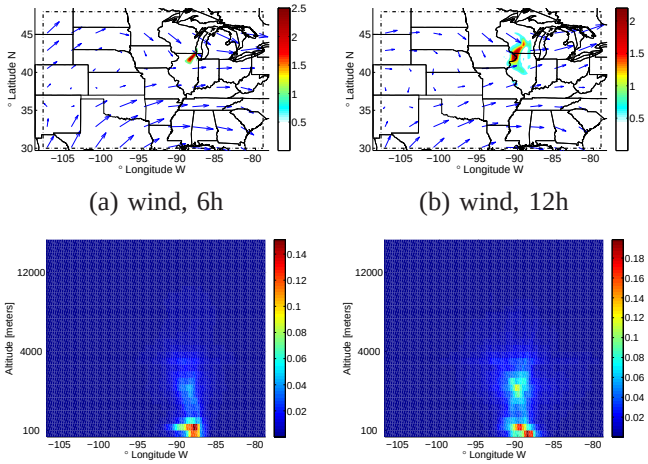
(a) wind, 6h (b) wind, 12h

(c) wind, 6h (d) wind, 12h

Fig. 5: Regional cost sensitivity with respect to wind speed, 6 and 12 hours before the final time (October 19, 2000 - 06:00)

VI. Conclusions

We have presented a framework for adjoint sensitivity of numerical weather prediction models. We have found that



(a) wind, 6h (b) wind, 12h

(c) wind, 6h (d) wind, 12h

Fig. 6: Regional cost sensitivity with respect to wind speed, 6 and 12 hours before the final time (October 21, 2000 - 06:00)

adjoint analysis provides valuable information that can aid in planning and operation tasks for the power grid. We discuss how to use optimization sensitivity capabilities to map physical weather variables to power grid economic metrics. Particular applications of the framework include wind farm and meteorological sensor placement and generation/transmission expansion planning. A numerical case study has been provided to illustrate the developments.

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 LCRC for the use of the Fusion cluster.

References

[1] C. Monteiro, R. Bessa, V. Miranda, A. Boterrud, J. Wang, and G. Conzelmann, "Wind power forecasting: state-of-the-art 2009," INESC Porto and Argonne National Laboratory, Tech. Rep., 2009.

[2] E. Constantinescu and A. Sandu, "Extrapolated implicit-explicit time stepping," *SIAM Journal on Scientific Computing*, vol. 31, no. 6, pp. 4452–4477, 2010.

[3] 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," NCAR, Tech. Rep. Tech Notes-475+ STR, 2008.

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

[5] E. Constantinescu, V. 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*, vol. 26, no. 1, pp. 431–441, 2011.

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

[7] 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)*, vol. 42, no. 23, pp. 5787–5804, 2008.

[8] R. Errico, "What is an adjoint model?" *Bull. Am. Met. Soc.*, vol. 78, pp. 2577–2591, 1997.

[9] R. Giering and T. Kaminski, "Recipes for adjoint code construction," *ACM Transactions on Mathematical Software (TOMS)*, vol. 24, no. 4, pp. 437–474, 1998.

[10] J. Nocedal and S. Wright, *Numerical Optimization*. New York, NY: Springer, 1999.

[11] S. M. Robinson, "Strongly regular generalized equations," *Mathematics of Operations Research*, vol. 5, pp. 43–61, 1980.

[12] A. Wächter and L. T. Biegler, "On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming," *Mathematical Programming*, vol. 106, pp. 25–57, 2006.

[13] H. Pirnay, R. Lopez-Negrete, and L. T. Biegler, "Optimal sensitivity based on IPOPT," *Submitted for Publication*, 2011.

[14] V. M. Zavala, A. Botterud, E. M. Constantinescu, and J. Wang, "Computational and economic limitations of dispatch operations in the next-generation power grid," *IEEE Conference on Innovative Technologies for Efficient and Reliable Power Supply*, 2010.