IISE Transactions

ISSN: 2472-5854 (Print) 2472-5862 (Online) Journal homepage: https://www.tandfonline.com/loi/uiie21

When wind travels through turbines: A new statistical approach for characterizing heterogeneous wake effects in multi-turbine wind farms

Mingdi You, Eunshin Byon, Jionghua (Judy) Jin & Giwhyun Lee

To cite this article: Mingdi You, Eunshin Byon, Jionghua (Judy) Jin & Giwhyun Lee (2017) When wind travels through turbines: A new statistical approach for characterizing heterogeneous wake effects in multi-turbine wind farms, IISE Transactions, 49:1, 84-95, DOI: 10.1080/0740817X.2016.1204489

To link to this article: https://doi.org/10.1080/0740817X.2016.1204489

Accepted author version posted online: 28 Jun 2016.
Published online: 17 Oct 2016.

IISE TRANSACTIONS
2017, VOL. 49, NO. 1, 84–95
http://dx.doi.org/10.1080/0740817X.2016.1204489

Taylor & Francis
Taylor & Francis Group

When wind travels through turbines: A new statistical approach for characterizing heterogeneous wake effects in multi-turbine wind farms

Mingdi You[a], Eunshin Byon[a], Jionghua (Judy) Jin[a] and Giwhyun Lee[b]

[a]Department of Industrial & Operations Engineering, University of Michigan, Ann Arbor, MI, USA; [b]Korea Army Academy, Yeong-Cheon, Korea

ABSTRACT

Modern utility-scale wind farms consist of a large number of wind turbines. In order to improve the power generation efficiency of wind turbines, accurate quantification of power generation levels of multi-turbines is critical, in both wind farm design and operational controls. One challenging issue is that the power output levels of multiple wind turbines are different, due to complex interactions between turbines, known as wake effects. In general, upstream turbines in a wind farm absorb kinetic energy from wind. Therefore, downstream turbines tend to produce less power than upstream turbines. Moreover, depending on weather conditions, the power deficits of downstream turbines exhibit heterogeneous patterns. This study proposes a new statistical approach to characterize heterogeneous wake effects. The proposed approach decomposes the power outputs into the average pattern commonly exhibited by all turbines and the turbine-to-turbine variability caused by multi-turbine interactions. To capture the wake effects, turbine-specific regression parameters are modeled using a Gaussian Markov random field. A case study using actual wind farm data demonstrates the proposed approach's superior performance.

ARTICLE HISTORY
Received 15 October 2015
Accepted 5 June 2016

KEYWORDS
Gaussian Markov random field; random effects; spatial modeling; spline regression; wind energy

1. Introduction

Wind energy is one of the fastest growing renewable energy sources in many parts of the world. In the United States, wind energy installation increased by more than a factor of ten in the past decade, from 4.2 GW in 2001 to almost 66 GW by the end of 2014 (American Wind Energy Association, 2015). The U.S. Department of Energy's Wind Vision study scenario envisions that wind energy will supply 35% of the domestic electrical power by 2050 (U.S. Department of Energy, 2015).

Unlike fossil fuel-based power systems, which generally operate in steady state, wind power systems operate in highly stochastic conditions (Byon *et al.*, 2010; Lee *et al.*, 2015a). As wind speed has been deemed the major factor that influences power production, many data-driven methods including neural network (Barbounis *et al.*, 2006), support vector regression (Yampikulsakul *et al.*, 2014), and other statistical methods (Damousis *et al.*, 2004; Sánchez, 2006; Chen *et al.*, 2010; Byon *et al.*, 2016) have been developed to quantify the power output at a single turbine, given a wind speed. This wind-to-power relationship is called a power curve in the wind industry. Recently, turbulence intensity, air density, and other weather factors have also received attention in studying their effects on the power generation efficiency (Lee *et al.*, 2013, Lee *et al.*, 2015a).

Due to economic reasons and transmission constraints, modern utility-scale wind farms place turbines close together, especially when the geography is favorable. One interesting, yet challenging, aspect is that the power generation patterns of multiple turbines in a wind farm are different from that of a stand-alone turbine; in fact, there are wide differences in power outputs due to the complex interactions, known as wake effects, among turbines in a utility-scale wind farm (Kusiak and Song, 2010). The explanation is simple. As wind passes through, upstream turbines absorb kinetic energy from the wind. As a result, the downstream turbines installed behind the upstream turbines have less kinetic energy to convert, thus generating less power than the upstream turbines. Therefore, maximizing the net energy production in the design layouts (i.e., where to install turbines; Kusiak and Song (2010)) and operational controls (e.g., yaw and pitch control; Johnson and Thomas, (2009)) relies heavily on accurately quantifying the wake effects throughout an industrial-scale wind farm.

Some initial attempts have been made to understand the wake effects in wind farm operations. The wake effects literature falls roughly into three categories. Studies in the first category focus on analytically predicting wind speeds at downstream turbines, based on physical understanding (Jensen, 1983; Katic *et al.*, 1986). Physics-based engineering models such as Jensen's model (Jensen, 1983), also known as the Park model, estimate the wind speed deficits at downstream turbines in an explicit functional form of the wind speeds at upstream turbines and other factors such as a distance between turbines. For mathematical tractability, these models are typically built upon simplifying assumptions and, as such, they cannot capture the complicated nature of wake effects (Staid, 2015).

Studies in the second category use sophisticated simulation models, such as computational fluid dynamics, to obtain detailed wind flow in wind farms (Porté-Agel *et al.*, 2010; Fleming *et al.*, 2014). Recently, Large Eddy Simulation, which investigates the wake characteristics including the turbulence generated from wind turbines in fine scales, has proven

CONTACT Eunshin Byon ebyon@umich.edu
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/uiie.

valuable in reducing modeling uncertainties (Wu and Porté-Agel, 2011). However, running the numerical simulation model is extremely time-consuming. For example, Fleming *et al.* (2014) use SOWFA (Simulator fOr Wind Farm Applications), which is a high-fidelity simulation software developed by the National Renewable Energy Laboratory (NREL), and simulate the wake effects in a wind farm of size of 2142 m × 378 m with a single NREL 5-MW baseline turbine. It takes 34.4 hours to simulate 1000 seconds of turbine operations using the NREL Red Mesa supercomputer with 256 processors. Therefore, the application of this numerical approach has been limited to specific cases for detailed analysis; e.g., at certain wind conditions (Fleming *et al.*, 2014).

Lastly, the third approach develops statistical methods to estimate the spatial wind field in a wind farm. Recently, Pourhabib *et al.* (2016) employ vector autoregressive models to predict wind speeds using historical wind conditions collected from anemometers mounted on each turbine's nacelle. However, the quantification of wake effects such as power deficits, or deficits in incoming wind speeds that turbines face, is not addressed because their focus is predicting downwind speeds at anemometers.

Despite the difference among these existing approaches, they commonly focus on estimating the spatial variation of a local wind field in a wind farm. Although these studies have contributed to wind industry, they have limitations in understanding the wake effects because they do not directly quantify the heterogeneity in turbines' power generations.

To better understand the relationship between wake effects and the differences in power generation, this article proposes a new modeling approach that directly quantifies different power generation levels at multiple interacting turbines. The proposed approach has two components. The first component characterizes the general nonlinear relationship between the power generation and weather conditions, which are commonly exhibited in all of the turbines in a wind farm. We call this a global trend and model it with spline functions in the article. The second component captures the turbine-to-turbine variations caused by the wake effects. We call this spatial variations in this article. In particular, the differences in power outputs between turbines are heterogeneous, depending on weather conditions (to be detailed in Section 2). To characterize such heterogeneous turbine interactions, we model the spatial variations with spline functions of the weather conditions, where the spline regression parameters are modeled with the Gaussian Markov Random Field (GMRF; Rue and Held (2005)). Our implementation results with wind farm data collected during actual operations indicate that the proposed approach can successfully quantify the heterogeneous wake effects, demonstrating its superior performance over other alternative methods.

The remainder of this article is organized as follows. Section 2 describes the dynamic characteristics of wake effects. The proposed approach is presented in Section 3. Section 4 discusses the implementation details. Section 5 presents the case study for evaluating the proposed model and compares its performance with other benchmark models. Finally, we summarize the article in Section 6.

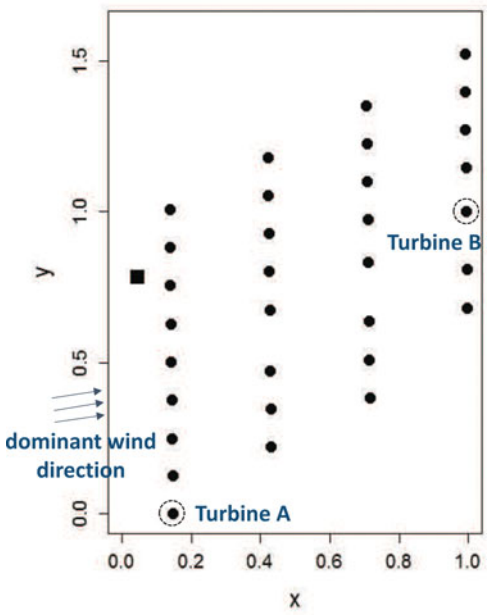


Figure 1. Wind farm layout.

2. Characteristics of wake effects

We begin with describing dynamic characteristics of wake effects. Figure 1 shows a partial layout of the actual wind farm with 30+ turbines used in this article (for confidentiality, the geographical coordinates have been rotated). The *x*- and *y*-axes, scaled in 0–1 in the *x*-axis and 0–1.5 in the *y*-axis, depict the relative locations of the turbines in the longitude and latitude in the rotated coordinates. The solid circles represent the wind turbines, and the solid square is the meteorology tower (henceforth met-tower) that collects meteorological measurements including wind speed, wind direction, turbulence intensity, air density, and humidity. The height of the met-tower and the hub heights of the wind turbines are the same in this wind farm.

The wind farm considered in this study has only one met-tower, which is placed facing the dominant wind direction. Therefore, the met-tower can measure the free-flow wind conditions at most times. However, when wind blows in non-dominant directions, the met-tower falls into the wake region, and the data collected at the met-tower do not represent the free-flow wind. Therefore, in our analysis we limit the data in the dominant wind direction. Under this direction, Turbine A in Fig. 1 is an upstream turbine, and Turbine B in Fig. 1 becomes a downstream turbine.

To understand the impact of ambient wind conditions on power generations, Figure 2 shows the scatter plots of power generations from Turbine A and B versus the wind speed. The *y*-axis is the 10-minute average power generation in the standardized 0–1000 scale (we re-scale the power to 0–1000 to retain the data confidentiality), and the *x*-axis is the 10-minute average wind speed collected at the met-tower. The "+" marks indicate the power outputs from Turbine A, and the "○" marks correspond to the outputs from Turbine B. Figure 2 shows that both Turbine A and B tend to generate more power as the wind speed increases up to about 14–15 m/s and then generate a steady level of power due to their blade pitch controls, which reduce energy

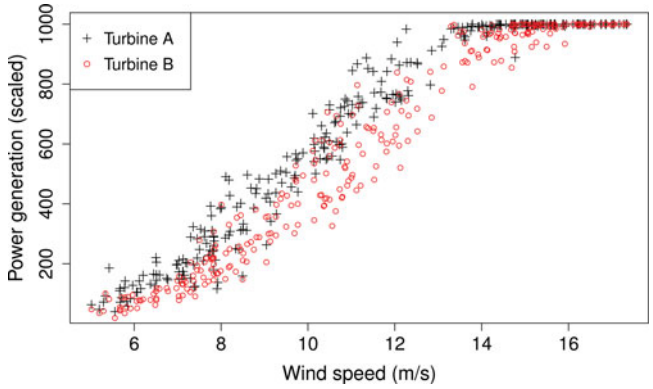


Figure 2. Heterogeneous power generation patterns at two turbines.

absorption to protect their blades in strong wind conditions (Lee *et al.*, 2013; Yampikulsakul *et al.*, 2014).

Although all turbines in a wind farm exhibit nonlinear patterns similar to those shown in Fig. 2, the power generation differs among multiple turbines, due to their complex interactions in the spatial neighborhood. In particular, downstream turbines exhibit power deficits (i.e., smaller levels of power generation) due to the wake effects, and the power deficits show heterogeneous patterns, depending on the wind condition. As Fig. 2 shows, the difference in power outputs between Turbines A and B tends to increase as the wind speed increases and then diminishes under high wind speeds.

To understand the heterogeneous wake effects over a range of wind conditions in multiple turbines, Fig. 3(a) depicts the power generation in one of the samples in the dataset when the wind speed is 12 m/s. The colors of each circle indicate the power generation levels at each turbine. A darker color in the solid circle represents a larger power generation. Note that in general, turbines on the left (upstream turbines) produce more power than those on the right (downstream turbines). On the contrary, the differences among power generations from multiple wind turbines are not significant when the wind speed is very low (Fig. 3(b)) or high (Fig. 3(c)). The rationale is that when the wind speed is low (Fig. 3(b)), upstream turbines' blades rotate slowly and only a small amount of energy is extracted from the wind. As a result, the energy loss due to the wake effects for the downstream turbines is not significant. Moreover, when the wind travels from upstream turbines to downstream turbines, it can recover energy to some extent. When the wind speed is high (Fig. 3(c)), upstream turbines use pitch control to generate a steady level of power, called the rated power, and they do not extract the maximum energy they can. As such, downstream turbines still face incoming wind with sufficient energy to convert to the rated power.

3. Mathematical model

This section provides a new modeling framework that addresses the heterogeneity in the power generation performance of multiple turbines, caused by the spatial variations coupled with wind conditions. Let $Y_{n,i}$ $(n = 1, \ldots, N; i = 1, 2, \ldots, I)$ be the power output of the nth turbine at the ith sample in the dataset, where N is the number of wind turbines in the wind farm and I is the total number of samples in the dataset. Let $\mathbf{x}_i = [x_i^1, x_i^2, \ldots, x_i^K]^T$ denote a vector of weather covariates (e.g., wind speed, turbulence intensity) at the ith sample, measured at the met-tower, where K is the number of weather factors considered in the model.

We model the power generation at multiple turbines by decomposing the power output $Y_{n,i}$ into the global trend component $G(\cdot)$ and the spatial component $S(\cdot)$ as

$$Y_{n,i} = G(\mathbf{x}_i; \beta) + S(\mathbf{x}_i; \gamma_n) + \epsilon_{n,i}, \quad n = 1, \ldots, N; \; i = 1, \ldots, I, \tag{1}$$

where $\epsilon_{n,i}$ represents the white noise, which is assumed to be an independent Gaussian random variable with a mean of zero and variance σ^2. Figure 2 shows the similarities in the general power generation patterns of multiple turbines. The first term $G(\cdot)$ explains the average pattern that turbines commonly exhibit in response to weather condition $\mathbf{x}_i$. In $G(\cdot)$, all of the turbines share the same parameter vector β. The second term $S(\cdot)$ handles the turbine-to-turbine variability with the turbine-specific random effect vector, γ_n.

For notational simplicity, we suppress the subscript i in the subsequent discussions when no confusion arises and use the model

$$Y_n = G(\mathbf{x}; \beta) + S(\mathbf{x}; \gamma_n) + \epsilon_n, \quad n = 1, \ldots, N. \tag{2}$$

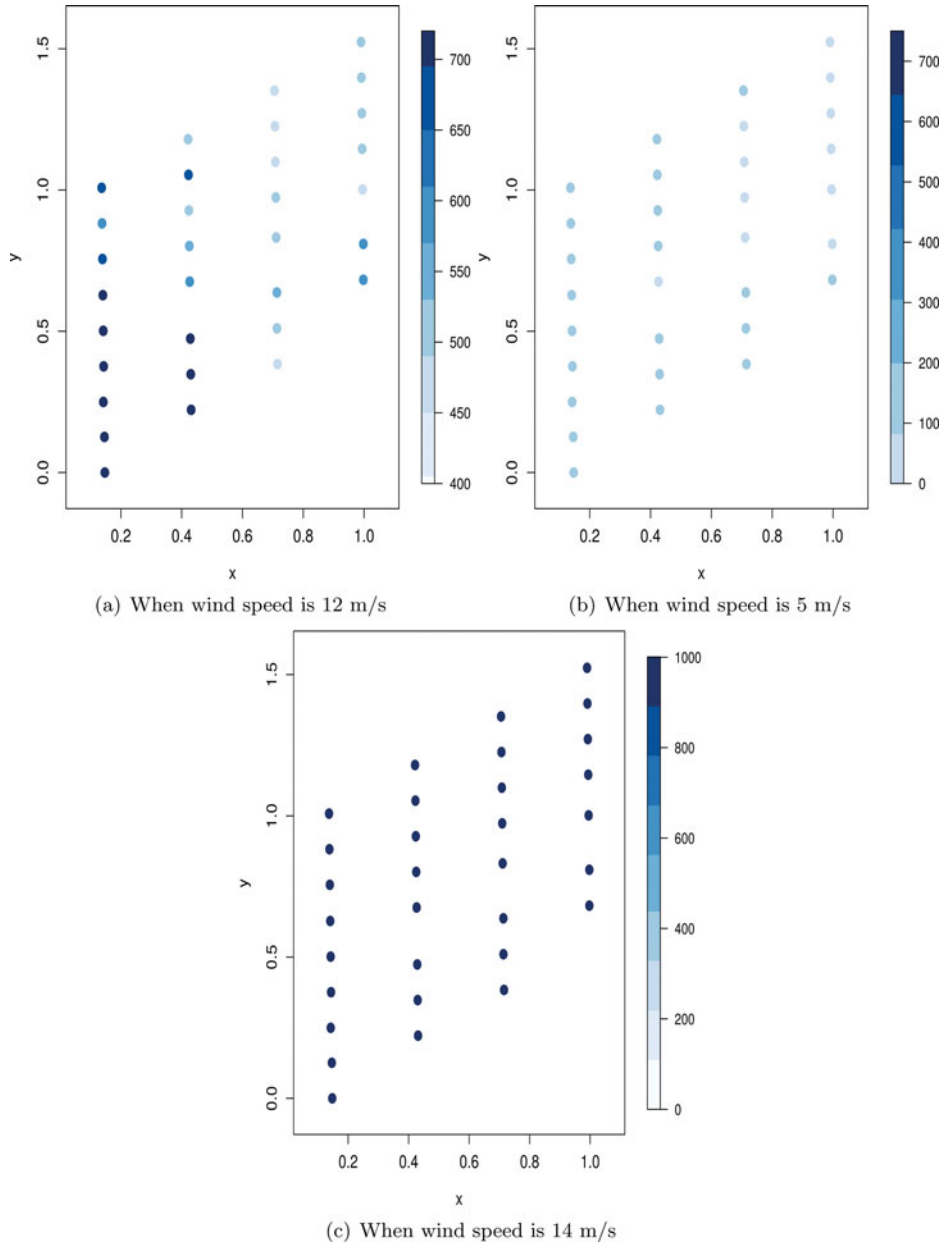
3.1. Modeling the global pattern

First, we model the global component $G(\mathbf{x}; \beta)$ in Equation (2) to represent the nonlinear pattern between the power output and weather conditions as shown in Fig. 2. In the literature, several nonlinear regression methods, including the kernel-based method (Lee *et al.*, 2015b), support vector regression (Yampikulsakul *et al.*, 2014), and neural networks (Li *et al.*, 2001), are proposed to capture the nonlinear relationship between weather covariates and power generation for a single turbine. Even though they can possibly capture the global patterns that multi-turbines commonly exhibit, they have limitations in incorporating turbine-to-turbine interactions. In this study, we use the spline regression (Friedman *et al.*, 2009; Lee *et al.*, 2013; Choe *et al.*, 2015) in the global term, as it can be easily combined with the spatial component, while successfully capturing the nonlinearity.

Among various spline models, we use the basis spline, or B-spline, as it empirically shows a good performance in the dataset we used in this study. Specifically, with the input data with K weather factors, $\mathbf{x} = [x^1, x^2, \ldots, x^K]^T$, we model $G(\mathbf{x}; \beta)$ as

$$G(\mathbf{x}; \beta) = \sum_{k=1}^{K} \mathbf{x}^{g,k} \beta^k, \tag{3}$$

where $\mathbf{x}^{g,k}$ is a row vector in the design matrix for the B-spline basis corresponding to the k^{th} weather factor x^k, and β^k is a column vector consisting of associated spline regression coefficients. We include the detailed description of the B-spline basis in the Appendix.



(a) When wind speed is 12 m/s

(b) When wind speed is 5 m/s

(c) When wind speed is 14 m/s

Figure 3. Power generation at multi-turbines under different wind speeds.

3.2. Modeling the spatial variations

This section models the spatial variations. To capture the turbine-to-turbine variations, one common method is to use a random effect δ_n and model the power output Y_n as

$$Y_n = G(\mathbf{x}; \beta) + \delta_n + \epsilon, \quad n = 1, 2, \ldots, N, \tag{4}$$

where $\delta_n \sim N(0, \tau^2)$. The use of typical random effects, however, has two deficiencies in characterizing the wake effects. First, it assumes homogeneous turbine-to-turbine variations across weather conditions; i.e., the probabilistic characteristic of δ_n, $n = 1, \ldots, N$, remains the same regardless of the weather condition. The power deficit at a downstream turbine, however, shows a heterogeneous pattern as seen in Fig. 2, which tends to be large in a mid-wind speed range but vanishes in both low- and high-wind speed ranges. In other words, the distribution of δ_n should depend on the weather condition. Second, it assumes independence among units (turbines in our case); i.e., it assumes δ_n to be independent of δ_m, $n \neq m$. As such, typical random effects models cannot capture multi-turbine interactions.

To overcome these limitations and address the heterogeneous turbine interactions in a range of weather conditions, our approach is to replace δ_n with $S(\mathbf{x}; \gamma_n)$ that depends on weather condition $\mathbf{x}$ and to make $S(\mathbf{x}; \gamma_n)$ inter-dependent. Figure 2 indicates that the power difference between the two turbines shows a piecewise pattern. That is, the power difference is negligible at the low wind speed range and then increases as wind speed increase up to about 12 m/s. It then decreases and finally diminishes to zero at the high wind speed range. To flexibly capture such a piecewise pattern, we employ a spline function for $S(\mathbf{x}; \gamma_n)$. In particular, similar to the global term, we use a B-spline function as follows:

$$S(\mathbf{x}; \gamma_n) = \sum_{k=1}^{K} \mathbf{x}^{s,k} \gamma_n^k, \tag{5}$$

where $\mathbf{x}^{s,k}$ is a row vector of the design matrix for the B-spline basis associated with input x^k. Note that $\mathbf{x}^{s,k}$ in $S(\mathbf{x};\ \gamma_n)$ can differ from $\mathbf{x}^{g,k}$ in Equation (3) if different B-spline degrees and knots are used for the global and spatial terms. γ_n^k is the column vector consisting of corresponding regression parameters. The regression parameter vector γ_n^k is turbine-specific for addressing the turbine-to-turbine variations, whereas the parameter vector β^k in Equation (3) for the global term is shared by all turbines.

Next, to capture the spatial dependence caused by the wake effects, we treat the regression parameters in γ_n^k as random effects and make them inter-dependent. One possible way is to model each regression parameter using GMRF (Cressie, 2015). We note, however, that treating each regression parameter as a random effect and individually modeling it with GMRF is computationally expensive and could lead to an unnecessarily complicated model. Therefore, we introduce a single vector of random effects $\eta^k = [\eta_1^k, \eta_2^k, \ldots, \eta_N^k]^T$ for each weather factor x^k, and set the regression parameter vector γ_n^k in Equation (5) as $\gamma_n^k = \eta_n^k \cdot \gamma^k$. Then, the spatial term in Equation (5) becomes

$$S(\mathbf{x};\ \eta_n^{1:K},\ \gamma) = \sum_{k=1}^{K} \eta_n^k \mathbf{x}^{s,k} \gamma^k, \tag{6}$$

where $\eta_n^{1:K}$ is a set of $\eta_n^1, \eta_n^2, \ldots, \eta_n^K$. With this simplification, the regression parameter vector γ^k becomes the same for every turbine, whereas the variations among individual turbines are captured by η_n^k for each input factor, x^k.

To illustrate how the spatial model in Equation (6) can simplify the model in Equation (5), suppose that $K = 1$ and the dimension of the row vector $\mathbf{x}^{s,1}$ is 4. With 30 turbines, the model in (5) includes $30 \times 4 = 120$ parameters for γ_n^1, and all of the 120 parameters are treated as random effects. Due to the large number of parameters that need to be estimated, our analysis with real data indicates that the resulting estimations are unstable. On the other hand, the simplified model in (6) uses 30 parameters for η_n^1 as random effects and four fixed parameters for γ^1. Consequently, the model in (6) significantly reduces the model complexity. Moreover, this simplification provides useful interpretations that γ^k addresses the heterogeneous power deficits in a range of weather condition, while η_n^k maintains the underlying idea of modeling the turbine-to-turbine variations.

To represent the multi-turbine interactions, we model η_n^k using GMRF. Specifically, the distribution of η_n^k can be specified by the full conditional probability density function of the Gaussian form:

$$\eta_n^k | \{\eta_m^k : m \in \mathcal{N}(n)\} \sim N\left(\sum_{m \in \mathcal{N}(n)} c_{n,m} \eta_m^k,\ \tau_k^2\right),$$
$$n = 1, 2, \ldots, N, \tag{7}$$

for the kth weather factor. Here, $\mathcal{N}(n)$ denotes the nth turbine's neighboring turbines, and $c_{n,m}$ captures the dependence intensity between turbine n and turbine m, $n \neq m$. In the spatial modeling literature (Besag, 1974), the neighborhood is typically defined for the regular lattice data. However, the layout of a wind farm in reality can be irregular. We will detail how we define the neighborhood turbines in Section 4.

Next, we define the dependence intensity $c_{n,m}$ in Equation (7). Noting that both wind direction and distances among turbines affect the dependence intensity between two turbines, we use the directional spatial dependence intensity proposed in Kaiser *et al.* (2002),

$$c_{n,m} = \delta_1 \sin^2(\theta_{n,m}) \left(\frac{1}{d_{n,m}}\right)^h + \delta_2 \cos^2(\theta_{n,m}) \left(\frac{1}{d_{n,m}}\right)^h, \tag{8}$$

for $n, m = 1 \ldots, N$, $n \neq m$, where $\theta_{n,m}$ is the angle between the wind direction and the line connecting turbine n and turbine m, and $d_{n,m}$ is the distance between two turbines. The dependence intensity parameter $c_{n,m}$ decomposes the spatial dependence between the two turbines into two components: the first term in Equation (8) corresponds to the dependence in the direction orthogonal to the wind direction, and the second term corresponds to the dependence in the wind direction. Coefficients δ_1 and δ_2 measure the contributions in each direction. The exponent h scales how quickly the dependence intensity decreases as the distance increases.

The full conditional density, specified in Equation (7), allows us to obtain the joint density function of $\boldsymbol{\eta}^k = [\eta_1^k, \eta_2^k, \ldots, \eta_N^k]^T$ as

$$f(\boldsymbol{\eta}^k) = (2\pi)^{-\frac{N}{2}} |Q^k|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\boldsymbol{\eta}^k)^T Q^k \boldsymbol{\eta}^k\right], \tag{9}$$

where the precision matrix $Q^k = \tau_k^{-2}(\mathbb{1}_N - C)$ (Cressie, 2015). Here, $\mathbb{1}_N$ is an $N \times N$ identity matrix, and C is an $N \times N$ matrix whose (n, m)th off-diagonal element is $c_{n,m}$ defined in Equation (8) and diagonal elements are zero.

For a proper GMRF model, the precision matrix Q^k needs to be symmetric and positive definite. First, the matrix Q^k is symmetric because of $c_{n,m} = c_{m,n}$ in Equation (8). Next, Lemma 1 derives a sufficient condition to ensure the positive definiteness of precision matrix Q^k (see Appendix for the proof).

Lemma 1. *If*

$$\max\{|\delta_1|, |\delta_2|\} < \min\left\{1 \Big/ \left(\sum_{m=1, m\neq n}^{N} \left(1/d_{n,m}\right)^h\right),\ n = 1, \ldots, N\right\},$$

the precision matrix Q^k in Equation (9) is positive definite.

We use the result of Lemma 1 in defining the prior distributions of δ_1 and δ_2 in the next section. Figure 4 summarizes the proposed modeling framework.

3.3. Parameter estimations

We estimate the model parameters in a Bayesian hierarchical inference framework. To facilitate computation, we use conjugate priors for some parameters. For the regression coefficients in β^k and γ^k in the global and spatial terms, we use the multivariate Gaussian distribution as their priors with zero mean and covariance matrix $\sigma_1^2 \cdot \mathbb{1}$, where $\mathbb{1}$ is the identity matrix with an appropriate dimension. For σ^2 and each τ_k^2 $(k = 1, \ldots, K)$, we use the inverse gamma distribution with parameters α_0, κ_0, and α_k, κ_k $(k = 1, \ldots, K)$ as their priors. For δ_1 and δ_2, no conjugate priors are available. Therefore, we use the uniform distribution for their priors. In particular,

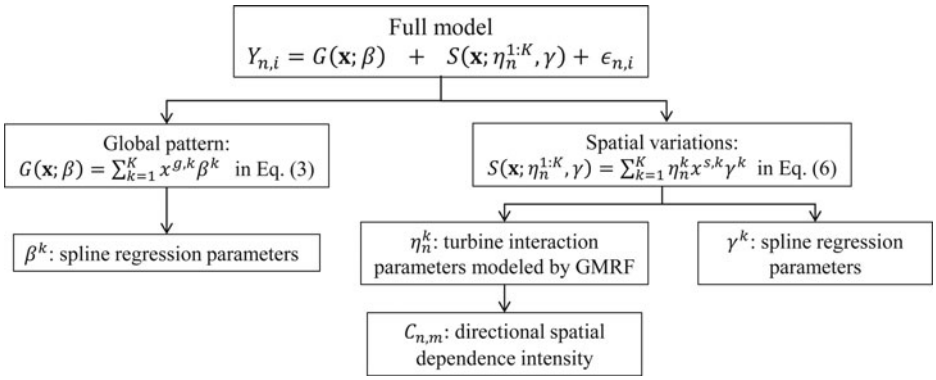


Figure 4. The global–spatial modeling framework for wake effects analysis.

based on Lemma 1, the uniform distributions are bounded by $\pm\min\{1/(\sum_{m=1,m\neq n}^{N}(1/d_{n,m})^h), n = 1,\ldots,N\}$ to ensure that their posteriors are proper. Below we summarize the proposed Bayesian hierarchical framework for wake effects analysis, after recovering the subscript i to denote the ith sample in the dataset.

Level 1: $Y_{n,i}|\mathbf{x}_i, \beta^k, \gamma^k, \eta_n^k$,

$$\sigma^2 \sim N\left(\sum_{k=1}^{K}\left(\mathbf{x}_i^{g,k}\beta^k + \eta_n^k\mathbf{x}_i^{s,k}\gamma^k\right), \sigma^2\right), \tag{10}$$

$n = 1,\ldots,N, i = 1,\ldots,I.$

Level 2: $\eta^k \sim MVN\left(0, \tau_k^{-2}(\mathbb{1} - C)\right), k = 1,\ldots,K.$ (11)

Level 3: $\beta^k \sim MVN(0, \sigma_1^2 \times \mathbb{1}), k = 1,\ldots,K,$ (12)

$$\gamma^k \sim MVN(0, \sigma_1^2 \times \mathbb{1}), k = 1,\ldots,K, \tag{13}$$

$$\delta_1, \delta_2 \sim \text{uniform}\left(-\min\left\{1/\left(\sum_{m=1,m\neq n}^{N}\left(1/d_{n,m}\right)^h\right), n = 1,\ldots,N\right\}, + \min\left\{1/\left(\sum_{m=1,m\neq n}^{N}\left(1/d_{n,m}\right)^h\right), n = 1,\ldots,N\right\}\right), \tag{14}$$

$$\tau_k^2 \sim IG(\alpha_k, \kappa_k), k = 1,\ldots,K, \tag{15}$$

$$\sigma^2 \sim IG(\alpha_0, \kappa_0). \tag{16}$$

Theorem 1 states that with the hierarchal structure in Models (10)–(16) and the result in Lemma 1, we can obtain a proper posterior density of the model parameters; i.e., the joint posterior density is integrable (see Appendix for the proof).

Theorem 1. *Let θ denote a vector of parameters $\beta^k, \gamma^k, \delta_1, \delta_2, \sigma^2$, and τ_k^2 $(k = 1, 2,\ldots,K)$. Assume that the prior densities of $\beta^k, \gamma^k, \delta_1, \delta_2, \sigma^2$, and τ_k^2 $(k = 1, 2,\ldots,K)$ are independent of one another, and the boundary conditions for δ_1 and δ_2 specified in Lemma 1 are satisfied. Then the joint posterior density of θ and η^k $(k = 1,\ldots,K)$ is proper.*

The posterior density for each parameter involves high-dimensional integrations, and it is not expressed as a closed form. As such, we use Markov Chain Monte Carlo (MCMC) methods to obtain the posterior means of the parameters. Given the input covariate vector, we obtain the fitted power at turbine n $(n = 1,\ldots,N)$, using

$$\hat{Y}_{n,i} = G(\mathbf{x}_i;\hat{\beta}) + S(\mathbf{x}_i;\hat{\gamma}_n) \tag{17}$$

$$= G(\mathbf{x}_i;\hat{\beta}) + S(\mathbf{x}_i;\hat{\eta}_n^{1:K},\hat{\gamma}) \tag{18}$$

$$= \sum_{k=1}^{K}\left(\mathbf{x}_i^{g,k}\hat{\beta}^k + \hat{\eta}_n^k\mathbf{x}_i^{s,k}\hat{\gamma}^k\right), \tag{19}$$

where $\hat{\beta}^k$, $\hat{\gamma}^k$, and $\hat{\eta}_n^k$ denote the posterior means of the corresponding parameters.

4. Implementation details

This section discusses the details for implementing the proposed approach. From the original dataset, we extract the data in the dominant wind direction shown in Fig. 1, when the measurements collected at the met-tower exhibit free-stream wind conditions. The resulting dataset with no missing measurements consists of 300 samples, and each sample includes the power outputs at all turbines in the wind farm and the met-tower's meteorological measurements of weather conditions.

As input factors, we consider the 10-minute average wind speed and the 10-minute average turbulence intensity. We also included other environmental factors such as humidity and air density in our preliminary analysis; however, their inclusion did not actually improve the estimation accuracy. The outputs in our model are the 10-minute average power generation levels at individual turbines in the wind farm. In our implementation, we do not consider the wind travel time across the wind farm, as it appears to be insignificant in the mid-size wind farm considered in this study.

For each input factor, we use the B-spline functions with different degrees in the global and spatial components. In the global component, to address the nonlinear pattern between the power generation and wind speed shown in Fig. 2, we consider a B-spline function with degree 2 or higher for wind speed. For the turbulence intensity, a simpler model appears to be capable of

modeling the relationship between power and turbulence intensity. Therefore, we use a B-spline of degree 1 or 2 for the turbulence intensity. For the B-spline functions in the spatial component, the degree is equal to, or smaller than, the degree in the global component considered in our implementation.

In defining the knots for wind speeds, we note that the wind speeds in the dataset range between 5 m/s and 17.5 m/s. Thus, we choose these two numbers as the boundary knots for the B-splines. In deciding internal knots, we consider that the rated wind speed of the wind turbines in our dataset is 15 m/s. The wind turbines produce a steady power level when wind speed is between 15 m/s and 17.5 m/s. Ideally, our B-splines should capture this piecewise pattern. Therefore, we select 15 m/s as one internal knot. For other internal knots, we use an equivalent stepwise distance, which results in the internal knot locations of 7.5 m/s, 10 m/s, 12.5 m/s, and 15 m/s. Next, in defining the knots for the turbulence intensities, we note that the turbulence intensities in the dataset range between 0.2 and 1.5. Thus, we use these two numbers as the boundary knots. In contrast with the rated wind speed, there are no specific numbers for the turbulence intensity available. Thus, we consider two internal knots with an equal stepwise distance—i.e., 0.63 and 1.07—which generally closely capture the piecewise patterns between the turbulence intensity and power generations. For each input factor, we use the same internal knots in the basis function for both the global and spatial components.

In defining the neighborhood turbines, the first- and second-order neighborhoods are typically considered. In our implementations, we define the neighborhood turbines as follows:

1. For wind turbines that are not placed on the wind farm's boundaries, we define the first-order neighborhood turbines as the eight nearest surrounding wind turbines. The second-order neighborhood turbines of a specific turbine are defined as the first-order neighborhood turbines of this turbine's first-order neighborhood turbines.
2. For the turbines on the boundary, we use a symmetric property. In other words, if a turbine, say Turbine D, is a first-order neighborhood turbine to Turbine C, Turbine C becomes the first-order neighborhood turbine to Turbine D (See Fig. 5). In addition, to define neighborhood turbines among turbines on the boundary, we set a threshold distance such that the turbines within the threshold distance become neighbors. In our implementation, the threshold is defined based on the wind farm size and the average distance among the neighborhood turbines within the boundary.

We consider both first- and second-order neighborhood turbines in our study. However, our preliminary analysis that used a second-order neighborhood did not improve the estimation performance. As such, we present the results using the first-order neighborhood for specifying $\mathcal{N}(n)$ in Equation (7) over the use of the second-order neighborhood. We use 0.5 for h in defining the dependency intensity in Equation (8) (Section 5 includes a sensitivity study with different values of h).

In the MCMC implementations, we do not assume any prior knowledge, and attempt to use non-informative priors. We set σ_1 in Equations (12) and (13) to be 1000, to ensure a sufficiently large coverage for β^k and γ^k. To specify the parameters in the inverse gamma priors in Equations (15) and (16), we employ non-informative priors by using the same shape and scale parameters (Gelman, 2006). Specifically, we use $\alpha_k = \kappa_k = 0.1$ $(k = 0, 1, \ldots, K)$ (Banerjee *et al.*, 2014). The MCMC is conducted by calling the WinBUGS software from R using the function *bugs()* from the library "R2WinBUGS" (Sturtz *et al.*, 2005). We run the simulation for 10 000 iterations with first 5000 iterations as burn-in iterations.

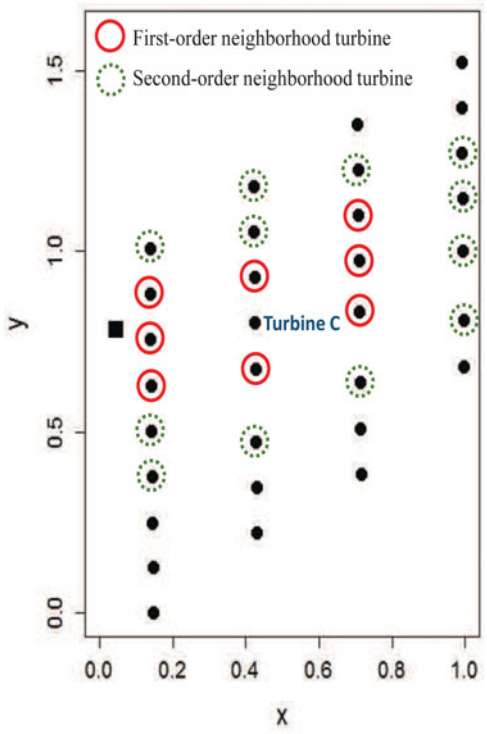


Figure 5. Illustrations of the first- and second-order neighborhood turbines of Turbine C.

We employ a 10-fold Cross-Validation (CV) to evaluate the estimation accuracy of the proposed approach, using three criteria: the Root Mean Square Error (RMSE), the Mean Absolute Error (MAE), and the Mean Absolute Percentage Error (MAPE) defined as

$$RMSE = \sqrt{\frac{\sum_{n=1}^{N}\sum_{i=1}^{I_0}(Y_{n,i} - \hat{Y}_{n,i})^2}{N \times I_0}} \tag{20}$$

$$MAE = \frac{\sum_{n=1}^{N}\sum_{i=1}^{I_0}|Y_{n,i} - \hat{Y}_{n,i}|}{N \times I_0} \tag{21}$$

$$MAPE = \frac{\sum_{n=1}^{N}\sum_{i=1}^{I_0}|Y_{n,i} - \hat{Y}_{n,i}|/Y_{n,i}}{N \times I_0}, \tag{22}$$

where I_0 is the size of a testing set, and $Y_{n,i}$ and $\hat{Y}_{n,i}$ are the actual and estimated power generation levels at the nth turbine and the ith sample in the testing set. Recall that in our implementation, the power generation is re-scaled to 0–1000 to retain the confidentiality of data source.

5. Case study

This section discusses the computational results of the proposed model. We also compare its estimation accuracy with other alternative methods.

Table 1. Average and standard deviation of RMSE, MAE, and MAPE from 10-fold CV

Model	*RMSE*	*MAE*	*MAPE*
G(W2)S(W1)	58.46 (6.86)	39.31 (4.77)	0.12 (0.02)
G(W3)S(W1)	58.45 (6.74)	39.54 (4.81)	0.11 (0.01)
G(W3)S(W2)	60.53 (8.29)	41.06 (5.50)	0.11 (0.02)
G(W2T1)S(W1)	57.12 (5.72)	39.26 (4.22)	0.12 (0.02)
G(W2T1)S(W2)	56.95 (5.86)	39.29 (4.33)	0.12 (0.02)
G(W2T1)S(W1T1)	57.15 (5.72)	39.15 (4.37)	0.12 (0.02)
G(W2T1)S(W2T1)	**56.68 (5.83)**	**38.62 (4.44)**	**0.11 (0.02)**
G(W2T2)S(W1T1)	57.10 (5.29)	38.92 (4.12)	0.11 (0.02)

Note: Each number in parentheses is the standard deviation of RMSEs, MAEs, and MAPEs from 10 testing sets.

5.1. Computational results

We implement the proposed model with different B-spline degrees in both global and spatial terms. Table 1 summarizes the estimation results of eight models in terms of the average RMSE, MAE, and MAPE of 10 testing sets in the 10-fold CV and their standard deviations. The shorthand names in the first column of Table 1 specify the input covariates included in the model and the degrees of the spline functions in the global and spatial components. For example, G(W2T1)S(W2T1) refers to the model using the B-splines of degrees 2 and 1 for wind speed and turbulence intensity, respectively, in both global pattern $G(\cdot)$ and spatial component $S(\cdot)$. The first three models in Table 1, namely, G(W2)S(W1), G(W3)S(W1), and G(W3)S(W2), do not include the turbulence intensity as an input factor. The comparisons between the first three models and other models in Table 1 suggest that including turbulence intensity in the input vector slightly improves the estimation accuracy. However, the fitting errors from different models are within one standard deviation, and no model significantly outperforms others. Therefore, we believe that any model would be acceptable. Overall, the model G(W2T1)S(W2T1), highlighted in bold in Table 1, provides the smallest average error in the testing sets, so we use it in subsequent discussions.

Figure 6 depicts the fitted results at Turbines A and B using model G(W2T1)S(W2T1). It should be noted that multiple estimations are made at the same wind speed in the figure due to the different values of turbulence intensity. In general, the estimated power at Turbine B at the downwind location is smaller than that of Turbine A at the upwind location, which is consistent with our observations. Moreover, the model can capture the heterogeneous power deficits at the downstream turbine; the difference of estimated power outputs between the two turbines (or the power deficit at Turbine B) increases as wind speed increases, but as the wind speed nears 14–15 m/s, the difference becomes smaller.

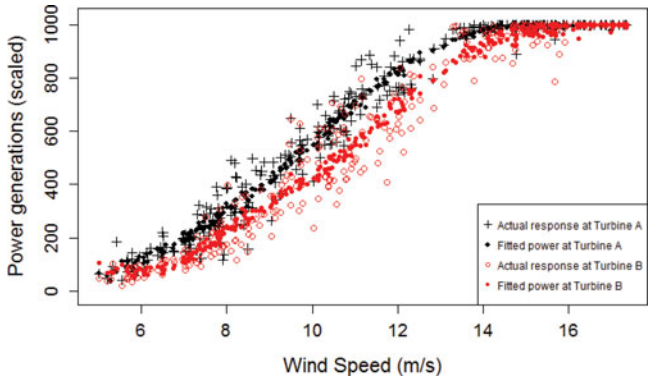


Figure 6. Comparison of actual and fitted power at Turbines A and B using model G(W2T1)S(W2T1).

Table 2. Sensitivity analysis on h with model G(W2T1)S(W2T1)

	$h = 0.2$	$h = 0.5$	$h = 1$	$h = 1.5$	$h = 2$
RMSE	56.71	56.68	56.87	56.55	56.79
MAE	38.73	38.62	38.55	38.86	38.67
MAPE	0.11	0.11	0.11	0.11	0.11

5.2. Sensitivity analysis

This section describes a sensitivity study for the exponent parameter h in Equation (8) in defining the dependence intensity. Table 2 shows the results with different values for h in model G(W2T1)S(W2T1) and reports the average RMSE, MAE, and MAPE of the 10 testing sets. We note that the estimation results are not sensitive to the choice of the parameter h. Therefore, we keep using $h = 0.5$ in our subsequent analysis.

5.3. Comparison with alternative methods

We compare the estimation capability of the proposed modeling approach with other alternative models, including a physics-based engineering model and statistical models. The first model we consider is Jensen's model (Jensen, 1983; Katic *et al.*, 1986), one of the most widely used physics-based engineering models in the global wind industry. Jensen's model analytically predicts the incoming wind speed at each turbine as follows:

$$(Jensen): V_{down} = (1 - \kappa_{deficit})V_{up}, \tag{23}$$

where V_{up} and V_{down} are the incoming wind speeds at an upstream turbine and a downstream turbine, respectively, and $\kappa_{deficit}$ is a wind speed deficit factor at a downstream turbine. Jensen's model (23) specifies the relationship between two turbines. We use the extended version of Jensen's model that incorporates multiple wake interactions discussed in Jensen (1983). The value of $\kappa_{deficit}$ depends on a wake effect decay parameter and other wind farm-specific factors, such as the distance between turbines and a rotor diameter. We use a wake effect decay parameter of 0.075, as suggested in Katic *et al.* (1986), DTU Wind Energy (2015), and Staid (2015).

Note that Jensen's model does not estimate the power generation level, as it focuses on estimating the incoming wind conditions at multiple turbines. As our turbines have the same specification (i.e., same turbine model from the same manufacturer), we assume similar power generation patterns, given the same incoming wind speed at each turbine. Then, we construct a power curve for one of the upstream turbines (Turbine A in Fig. 1), which is not in the wake region, and apply the resulting power curve to other turbines.

We also consider other statistical models. The first and simplest model contains only the global component

$$(GLB): Y_n = \sum_{k=1}^{K} \mathbf{x}^k \beta^k + \epsilon, \quad n = 1, 2, \ldots, N, \tag{24}$$

where $\mathbf{x}^k$ is a row vector of the design matrix of the B-spline basis with the kth weather factor x^k measured at the met-tower. The spline regression parameter vector β^k is the same for all wind turbines, implying that the same regression function is used for all turbines.

The second statistical model individually fits the spline regression for each turbine as

$$(IND): Y_n = \sum_{k=1}^{K} \mathbf{x}^k \beta_n^k + \epsilon, \quad n = 1, 2, \ldots, N, \tag{25}$$

where β_n^k differs for individual turbines, and the model separately builds a regression function using the data from each turbine. This model generates N (i.e., the number of turbines in a wind farm) regression functions.

The third statistical model uses the typical random effects, δ_n,

$$(RND): Y_n = \sum_{k=1}^{K} \mathbf{x}^k \beta^k + \delta_n + \epsilon, \quad n = 1, 2, \ldots, N, \tag{26}$$

where $\delta_n \sim N(0, \tau^2)$. All turbines share the same spline regression parameters β^k as in GLB, but δ_n represents the turbine-to-turbine variability. This model assumes independence among δ_n, $n = 1, \ldots, N$.

As the model G(W2T1)S(W2T1) provides the best estimation results in Section 5.1, we use the B-spline functions of degrees 2 and 1 for wind speed and turbulence intensity, respectively, in these three statistical models and estimate their parameters using MCMC. Table 3 summarizes the comparison results. The model names in rows 4, 5, and 6 of the first column can be interpreted as in Table 1. We summarize our observations as follows:

1. Among all methods, the proposed modeling approach, G(W2T1)S(W2T1), highlighted in bold in Table 3, produces the smallest estimation errors in terms of the three criteria. We also observe that overall, statistical models outperform Jensen's model. In particular, the estimation errors in Jensen's model are about double those in the statistical-based approaches.
2. Understandably, the simplest model, GLB(W2T1), generates the highest estimation errors among the four statistical models, as it does not account for the turbine-to-turbine variations.
3. Even though the differences between our approach and other statistical models, IND and RND, are not significant in terms of the fitting errors, our model provides better interpretations in practical applications:
   - 3.1 The model with the random effects, RND(W2T1), captures the variations among turbines, but it assumes homogeneous power deficits across the weather conditions and the independence among the turbines.
   - 3.2 The result from the individual model, IND(W2T1), is difficult to use for understanding the multi-turbine dependencies, as it presents a local modeling approach that individually analyzes the data from each turbine. Therefore, the IND model does not provide wind farm–level insights. On the contrary, the proposed model captures the interactions among turbines in an integrative manner, while preserving a strong fitting capability. This feature of our proposed model is critical for wind farm–level problems such as wind farm design and control optimization.

Table 3. Performance comparison

Model	*RMSE*	*MAE*	*MAPE*
Jensen's model	124.18	92.09	0.20
G(W2T1)S(W2T1)	**56.68 (5.83)**	**38.62 (4.44)**	**0.11 (0.02)**
GLB(W2T1)	65.01 (6.24)	45.34 (4.97)	0.13 (0.02)
IND(W2T1)	56.91 (5.81)	39.07 (4.31)	0.12 (0.02)
RND(W2T1)	59.52 (5.98)	43.26 (4.10)	0.12 (0.02)

Note: Numbers inside the parentheses are the standard deviations of fitting errors from the 10 testing sets. In Jensen's model, CV is not employed and the numbers are obtained from the fitting results in 300 samples.

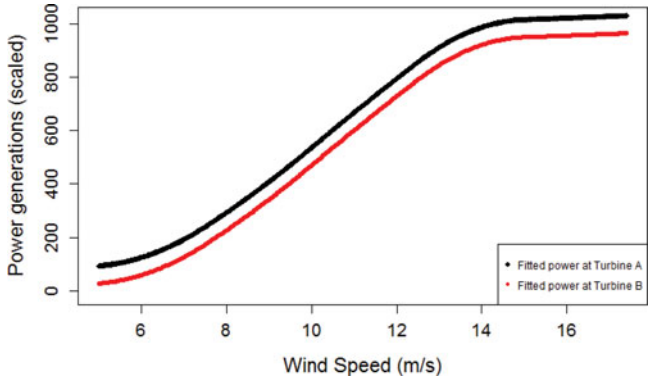


Figure 7. Estimated power curves for Turbines A and B using model RND(W2).

To further illustrate the major difference between the proposed approach and RND, Fig. 7 shows the fitted power curves for the two turbines, Turbines A and B, using RND with degree 2, RND(W2). Here, we exclude the turbulence intensity from the input vector and only include the wind speed as a covariate to see the limitation of RND more clearly. Figure 7 shows that the fitted curves for the two turbines are parallel, implying that the difference in the fitted power outputs is same across wind conditions. These results do not agree with actual power generation patterns, showing that RND cannot successfully capture the heterogeneous wake interactions among turbines.

5.4. Quantification of turbine-to-turbine variability

This section quantifies the variations in power generation levels at multiple turbines. We compute the power deficit at the nth turbine of the ith sample as

$$\hat{D}_{n,i} = \max_{m=1,\cdots,N} \hat{P}_{m,i} - \hat{P}_{n,i}. \tag{27}$$

That is, the power deficit at the the nth turbine represents the difference between its estimated power output and the maximum estimated output among N turbines.

To illustrate, we consider the wind speed as the only input weather factor and draw the box plots of power deficits at multiple wind speeds from 5 to 16 m/s with a stepwise of 1 m/s as shown in Fig. 8. Each box plot shows the minimum, first quartile, median, third quartile, and maximum power deficits at each wind speed. Figure 8 shows that the range of power deficits

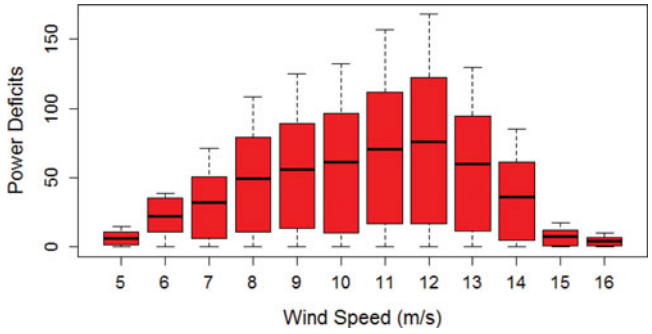


Figure 8. Box plots of estimated power deficits in multiple turbines using G(W2)S(W1).

increases when the wind speed increases up to 12 m/s. This implies that the turbine-to-turbine variability increases as the wind speed increases. When the wind speed is 12 m/s, the power deficit ranges from 0 to 184. Considering that the estimated power generation at Turbine A (the upstream turbine) is about 793, a turbine downstream generates 23% ($= 184/793 \times 100\%$) less power than Turbine A. When the wind speed is over 12 m/s, the turbine-to-turbine variability becomes smaller.

Overall, these results show good agreement with the observed data on the heterogeneous variability of power outputs discussed in Section 2. The results also indicate that controlling the power generation level at upstream turbines to absorb less energy from the wind can possibly increase the power generation levels at downstream turbines, thus optimizing overall performance at a wind farm level (Annoni *et al.*, 2014).

6. Summary

The accurate assessment of the wake effects is a fundamental step in optimally designing a wind farm layout and controlling wind turbines. We switch the conventional perspective of wake effects analysis that focuses on estimating the local wind field in a wind farm to directly estimating the power outputs of multiple turbines. This study proposes a new integrative approach for understanding the effects of unobserved wake effects on the heterogeneous power generation levels of multiple turbines. The approach decomposes the power outputs into global and spatial components, respectively, to model both the average pattern globally shared by all turbines and the turbine-specific variations. In representing the global pattern, B-splines are used to address the nonlinear relationship between power and weather conditions. For the spatial variations, we also use B-spline functions where the spline regression parameters are modeled through GMRF to capture the interactions among neighboring turbines. A case study using actual data from a multi-turbine wind farm suggests that the proposed statistical approach successfully characterizes the heterogeneous wake effects, compared with other alternative methods.

In our analysis, we only use the data from the dominant wind direction when the met-tower measures the free-wind flow condition. Due to the high cost of placing met-towers, typical wind farms have one met-tower (or at most a few in a large-size wind farm). In the future, we plan to extend our work to account for the wake effects in different wind directions by using data measured at multiple met-towers when such data become available to us and/or using data collected at turbine anemometers. Another direction is to consider other spline models. In particular, we will consider splines with the locations of knots determined by the data, rather than being fixed *a priori*, for example, by using the Bayesian multivariate adaptive regression splines (Lee *et al.*, 2013).

We will also extend the proposed approach to the reliability analysis of wind turbines (Byon, 2013; Ko and Byon, 2015; Choe *et al.*, 2016). Turbines in a downwind location are not only impacted by the wind speed deficits, thus capturing less energy, but they are also subject to increased loads due to the turbulence generated by the upstream turbines. Therefore, downstream turbines can deteriorate faster than upstream turbines. Capitalizing on the work in this article, wind farm layout design and turbine controls will be analyzed to achieve optimal generation efficiency and ensure reliability in industrial wind farms.

Acknowledgements

We thank the Editor, Department Editor, and anonymous reviewers for their constructive comments on various aspects of this work.

Funding

This work was partially supported by the National Science Foundation grants CMMI-123310, CMMI-1362513 and CMMI-1536924.

Notes on contributors

Mingdi You received a B.S. degree in mathematics from Southeast University, Nanjing, China, in 2011 and an M.S. degree in industrial & operations engineering from the University of Michigan, Ann Arbor, in 2013. He is currently working toward a Ph.D. degree in the Department of Industrial & Operations Engineering at the University of Michigan. His research interests include applied statistics and spatio temporal data analytic and reliability engineering with applications in wind energy. He is a student member of IIE and INFORMS.

Eunshin Byon received a Ph.D. degree in industrial and systems engineering from Texas A&M University, College Station, in 2010. She is an assistant professor with the Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor. Her research interests include data analytics, quality and reliability engineering, condition monitoring, operations and maintenance optimization, and uncertainty quantification. She is a member of IIE, INFORMS, IEEE, and ASQ.

Jionghua (Judy) Jin is a professor in the Department of Industrial and Operations Engineering at the University of Michigan. She received her Ph.D. degree from the University of Michigan in 1999. Her research focuses on data fusion for improving systems operational quality, reliability and efficiency. Her research emphasizes the integration of applied statistics, signal processing, reliability, system modeling, and decision-making theory. She has received numerous awards including a NSF CAREER Award in 2002, a PECASE Award in 2004, and 10 Best Paper Awards since 2005. She is a Fellow of IIE and ASME, a senior member of ISI; and a member of ASQ, IEEE, INFORMS, and SME.

Giwhyun Lee received an M.S. degree in statistics from Seoul National University, Seoul, South Korea, in 2006 and a Ph.D. degree in industrial and systems engineering from Texas A&M University, College Station, in 2013. She is an Assistant Professor in the Department of Military Operations Research in Korean Army Academy at Yeong-Cheon, South Korea. Her research interests include data analytics, monitoring and prediction of wind power, and fault detection and diagnosis.

References

American Wind Energy Association. (2015) AWEA U.S. wind industry first quater 2015 market report. Technical report, American Wind Energy Association, Washington DC.

Annoni, J., Seiler, P., Johnson, K., Fleming, P. and Gebraad, P. (2014) Evaluating wake models for wind farm control, in *Proceedings of the American Control Conference (ACC) 2014*, Portland, OR, USA, pp. 2517–2523.

Banerjee, S., Carlin, B.P. and Gelfand, A.E. (2014) *Hierarchical Modeling and Analysis for Spatial Data*, CRC Press, Boca Raton, FL.

Barbounis, T.G., Theocharis, J.B., Alexiadis, M.C. and Dokopoulos, P.S. (2006) Long-term wind speed and power forecasting using local recurrent neural network models. *IEEE Transactions on Energy Conversion*, **21**(1), 273–284.

Besag, J. (1974) Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society, Series B*, **36**(2), 192–236.

Byon, E. (2013) Wind turbine operations and maintenance: A tractable approximation of dynamic decision-making. *IIE Transactions*, **45**(11), 1188–1201.

Byon, E., Choe, Y. and Yampikulsakul, N. (2016) Adaptive modeling and prediction in time-variant processes with application to wind power systems. *IEEE Transactions on Automation Science and Engineering*, **13**(2), 997–1007.

Byon, E., Ntaimo, L. and Ding, Y. (2010) Optimal maintenance strategies for wind turbine systems under stochastic weather conditions. *IEEE Transactions on Reliability*, **59**(2), 393–404.

Chen, P., Pedersen, T., Bak-Jensen, B. and Chen, Z. (2010) ARIMA-based time series model of stochastic wind power generation. *IEEE Transactions on Power Systems*, **25**(2), 667–676.

Choe, Y., Byon, E. and Chen, N. (2015) Importance sampling for reliability evaluation with stochastic simulation models. *Technometrics*, **57**(3), 351–361.

Choe, Y., Pan, Q. and Byon, E. (2016) Computationally efficient uncertainty minimization in wind turbine extreme load assessments. *ASME Journal of Solar Energy Engineering: Including Wind Energy and Building Energy Conservation*, **138**(4), 041012.

Cressie, N. (2015) *Statistics for Spatial Data*, John Wiley & Sons, New York NY.

Damousis, I.G., Alexiadis, M.C., Theocharis, J.B. and Dokopoulos, P.S. (2004) A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation. *IEEE Transactions on Energy Conversion*, **19**(2), 352–361.

DTU Wind Energy. (2015) Wind resources for energy production of wind turbines. Available at: http://www.wasp.dk/wasp#details__wakeeffectmodel (accessed September 18, 2015).

Fleming, P.A., Gebraad, P.M.O., Lee, S., van Wingerden, J.-W., Johnson, K., Churchfield, M., Michalakes, J., Spalart, P. and Moriarty, P. (2014) Evaluating techniques for redirecting turbine wakes using SOWFA. *Renewable Energy*, **70**, 211–218.

Friedman, J., Hastie, T. and Tibshirani, R. (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, New York, NY.

Gelman, A. (2006) Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis*, **1**(3), 515–534.

Jensen, N.O. (1983) A note on wind generator interaction. Technical Report. Riso-M-2411, Riso National Laboratory, Roskilde, Denmark.

Johnson, K.E. and Thomas, N. (2009) Wind farm control: Addressing the aerodynamic interaction among wind turbines, in *Proceedings of the 2009 American Control Conference*, St. Louis, MO, pp. 2104–2109.

Kaiser, M.S., Daniels, M.J., Furakawa, K. and Dixon, P. (2002) Analysis of particulate matter air pollution using Markov random field models of spatial dependence. *Environmetrics*, **13**(5-6), 615–628.

Katic, I., Højstrup, J. and Jensen, N. (1986) A simple model for cluster efficiency, in *European Wind Energy Association Conference and Exhibition*, Rome, Italy, pp. 407–410.

Ko, Y.M. and Byon, E. (2015) Reliability evaluation of large-scale systems with identical units. *IEEE Transactions on Reliability*, **64**(1), 420–434.

Kusiak, A. and Song, Z. (2010) Design of wind farm layout for maximum wind energy capture. *Renewable Energy*, **35**(3), 685–694.

Lee, G., Byon, E., Ntaimo, L. and Ding, Y. (2013) Bayesian spline method for assessing extreme loads on wind turbines. *The Annals of Applied Statistics*, **7**(4), 2034–2061.

Lee, G., Ding, Y., Genton, M.G. and Xie, L. (2015a) Power curve estimation with multivariate environmental factors for inland and offshore wind farms. *Journal of the American Statistical Association*, **110**(509), 56–67.

Lee, G., Ding, Y., Xie, L. and Genton, M.G. (2015b) A kernel plus method for quantifying wind turbine performance upgrades. *Wind Energy*, **18**(7), 1207–1219.

Li, S., Wunsch, D.C., O'Hair, E.A. and Giesselmann, M.G. (2001) Using neural networks to estimate wind turbine power generation. *IEEE Transactions on Energy Conversion*, **16**(3), 276–282.

Ortega, J. M. (1987) *Matrix Theory*, Plenum Press, New York, NY.

Porté-Agel, F., Lu, H. and Wu, Y.-T. (2010) A large-eddy simulation framework for wind energy applications, in *Proceedings of the Fifth International Symposium on Computational Wind Engineering*, Chapel Hill, NC, USA.

Pourhabib, A., Huang, J.Z. and Ding, Y. (2016) Short-term wind speed forecast using measurements from multiple turbines in a wind farm. *Technometrics*, **58**(1), 138–147.

Rue, H. and Held, L. (2005) *Gaussian Markov Random Fields: Theory and Applications*, CRC Press, Boca Raton, FL.

Sánchez, I. (2006) Recursive estimation of dynamic models using cook's distance, with application to wind energy forecast. *Technometrics*, **48**(1), 61–73.

Staid, A. (2015) Statistical modeling to support power system planning, Ph.D. thesis, Johns Hopkins University, Baltimore, MD.

Sturtz, S., Ligges, U. and Gelman, A.E. (2005) R2WinBUGS: A package for running WinBUGS from R. *Journal of Statistical Software*, **12**(3), 1–16.

U.S. Department of Energy. (2015) Wind vision: A new era for wind power in the United States. Technical Report. DOE/GO-102015-4557, Washington DC.

Wu, Y.-T. and Porté-Agel, F. (2011) Large-eddy simulation of wind-turbine wakes: Evaluation of turbine parametrisations. *Boundary-Layer Meteorology*, **138**(3), 345–366.

Yampikulsakul, N., Byon, E., Huang, S., Shawn, S. and You, M. (2014) Condition monitoring of wind turbine system with nonparametric regression-based analysis. *IEEE Transactions on Energy Conversion*, **29**(2), 288–299.

Appendix

Description of B-spline basis and design matrix

For each weather factor x, suppose its domain is $[x_{\min}, x_{\max}]$. With a pre-specified degree D, let $x_{min} = \kappa_1 = \kappa_2 = \cdots = \kappa_d < \kappa_{d+1} < \cdots < \kappa_{T-d} = \kappa_{T-d+1} = \cdots = \kappa_T = x_{max}$ denote a non-decreasing sequence of the knots in the domain, where T is the total number of knots including the boundary knots. The B-spline basis can be defined recursively as (Friedman *et al.*, 2009)

$$b_{0,l}(x) = \begin{cases} 1, & \kappa_l \le x < \kappa_{l+1} \\ 0, & \text{otherwise} \end{cases}, \tag{A1}$$

and for $1 \le d \le D$,

$$b_{d,l}(x) = \frac{x - \kappa_l}{\kappa_{l+d-1} - \kappa_l} b_{d-1,l}(x) + \frac{\kappa_{l+d} - x}{\kappa_{l+d} - \kappa_{l+1}} b_{d-1,l+1}(x). \tag{A2}$$

Then, the row of the design matrix with input x becomes $[b_{D,1}(x), b_{D,2}(x), \ldots, b_{D,T}(x)]^T$. In our modeling, for the kth input factor x^k with the degree D^k and the total number of knots T^k, $\mathbf{x}^{g,k}$ in Equation (3) becomes $\mathbf{x}^{g,k} = [b_{D^k,1}(x^k), b_{D^k,2}(x^k), \ldots, b_{D^k,T^k}(x^k)]^T$. Similarly, $\mathbf{x}^{s,k}$ in Equation (5) can be defined.

Proof of Lemma 1

Proof. It suffices to show the positive definiteness of $\mathbb{1}_N - C$. Suppose the condition

$$\max\{|\delta_1|, |\delta_2|\} < \min\left\{1 \Big/ \left(\sum_{m=1,m\neq n}^{N} \left(1/d_{n,m}\right)^h\right), \; n = 1, 2, \ldots, N\right\},$$

holds. Then, for $\forall n = 1, 2, \ldots, N$,

$$1 > \max\{|\delta_1|, |\delta_2|\} \sum_{m=1,m\neq n}^{N} \left(\frac{1}{d_{n,m}}\right)^h \tag{A3}$$

$$= \sum_{m=1,m\neq n}^{N} \max\{|\delta_1|, |\delta_2|\} \left(\sin^2(\theta_{n,m}) \left(\frac{1}{d_{n,m}}\right)^h + \cos^2(\theta_{n,m}) \left(\frac{1}{d_{n,m}}\right)^h\right) \tag{A4}$$

$$\geq \sum_{m=1,m\neq n}^{N} \left(|\delta_1| \sin^2(\theta_{n,m}) \left(\frac{1}{d_{n,m}}\right)^h + |\delta_2| \cos^2(\theta_{n,m}) \left(\frac{1}{d_{n,m}}\right)^h\right) \tag{A5}$$

$$\geq \sum_{m=1,m\neq n}^{N} \left(\left|\delta_1 \sin^2(\theta_{n,m}) \left(\frac{1}{d_{n,m}}\right)^h + \delta_2 \cos^2(\theta_{n,m}) \left(\frac{1}{d_{n,m}}\right)^h\right|\right) \tag{A6}$$

$$= \sum_{m=1,m\neq n}^{N} |c_{n,m}|. \tag{A7}$$

As a result, $\sum_{m=1,m\neq n}^{N} |c_{n,m}| < 1, \forall n$. Note that the diagonal elements of $\mathbb{1}_N - C$ are 1. Consequently, the matrix $\mathbb{1}_N - C$ is diagonally dominant. This concludes that $\mathbb{1}_N - C$ is positive definite under the condition

$$\max\{|\delta_1|, |\delta_2|\} < \min\left\{1 \Big/ \left(\sum_{m=1,m\neq n}^{N} \left(1/d_{n,m}\right)^h\right), \; n = 1, 2, \ldots, N\right\}.$$

(Ortega, 1987). □

Proof of Theorem 1

Proof. For notational simplicity, we consider $K = 1$ and omit the subscript and superscript k in the following proof. The extension to the case with $K > 1$ is straightforward due to the independence assumption among the parameters in θ in their priors. We use $p(\cdot)$ to denote the prior distribution and $L(\cdot)$ to denote the likelihood function. Y and X denote the response vector and covariate matrix, respectively.

The joint posterior density satisfies

$$f(\theta, \eta|Y, X) \propto L(Y|X, \theta, \eta)p(\eta|\theta)p(\theta) \tag{A8}$$

$$\propto L(Y|X, \theta, \eta)p(\eta|\tau^2, \delta_1, \delta_2)p(\delta_1)p(\delta_2) \times p(\sigma^2)p(\tau^2)p(\beta)p(\gamma). \tag{A9}$$

To show that the posterior is proper, the integration of the right-hand side in (A9) should be finite. First, in the likelihood, we get

$$L(Y|X, \theta, \eta) = (2\pi\sigma^2)^{-NI/2} \times \exp\left\{-\frac{1}{2\sigma^2} \sum_{n,i} [Y_{n,i} - (x_i^g\beta + \eta_n x_i^s\gamma)]^2\right\} \tag{A10}$$

$$\leq (\sigma^2)^{-NI/2}, \tag{A11}$$

where the inequality in Equation (A11) holds due to the fact that $[Y_{n,i} - (x_i^g\beta + \eta_n x_i^s\gamma)]^2 \geq 0$.

Therefore, it suffices to show that the integration of the following term is finite; that is,

$$\int_{\theta,\eta} (\sigma^2)^{-NI/2} p(\eta|\tau^2, \delta_1, \delta_2)p(\delta_1)p(\delta_2)p(\sigma^2) \times p(\tau^2)p(\beta)p(\gamma)d\eta d\theta < \infty. \tag{A12}$$

Due to the independence assumption in priors, we have

$$\int_{\theta,\eta} (\sigma^2)^{-NI/2} p(\eta|\tau^2, \delta_1, \delta_2)p(\delta_1)p(\delta_2) \times p(\sigma^2)p(\tau^2)p(\beta)p(\gamma)d\eta d\theta \tag{A13}$$

$$= \left[\int_{\eta,\tau^2,\delta_1,\delta_2} p(\eta|\tau^2, \delta_1, \delta_2)p(\delta_1)p(\delta_2)p(\tau^2)d\eta d\tau^2 d\delta_1 d\delta_2\right] \times \left[\int_{\beta} p(\beta)d\beta\right] \times \left[\int_{\gamma} p(\gamma)d\gamma\right] \times \left[\int_{\sigma} (\sigma^2)^{-NI/2} p(\sigma^2)d\sigma\right]. \tag{A14}$$

Note that priors of β and γ are multivariate normal, so we have $\int_\beta p(\beta)d\beta = 1$ and $\int_\gamma p(\gamma)d\gamma = 1$. For σ^2, we use the inverse gamma distribution as its prior with parameters α_0 and κ_0. Then, we have

$$(\sigma^2)^{-NI/2} p(\sigma^2) \propto (\sigma^2)^{-NI/2-\alpha_0-1} \exp\left\{-\frac{\kappa_0}{\sigma^2}\right\}, \tag{A15}$$

which is proportional to the density of the inverse gamma distribution with parameters $NI/2 + \alpha_0$ and κ_0. Therefore, the integration $\int_\sigma (\sigma^2)^{-NI/2} p(\sigma^2)d\sigma$ is finite.

Lastly, the first term in Equation (A14) becomes

$$\int_{\eta,\tau^2,\delta_1,\delta_2} p(\eta|\tau^2, \delta_1, \delta_2)p(\delta_1)p(\delta_2)p(\tau^2)d\eta d\tau^2 d\delta_1 d\delta_2 \tag{A16}$$

$$\propto \int_{\delta_1,\delta_2} \left(\int_{\tau^2} \left(\int_{\eta} p(\eta|\tau^2, \delta_1, \delta_2)d\eta\right) p(\tau^2)d\tau^2\right) d\delta_1 d\delta_2. \tag{A17}$$

When δ_1 and δ_2 satisfy the boundary condition specified in Lemma 1, η is a valid multivariate Gaussian random variable. Thus, we have $\int_\eta p(\eta|\tau^2, \delta_1, \delta_2)d\eta = 1$. Moreover, with the inverse Gamma distribution as a prior of τ^2, we have $\int_{\tau^2} p(\tau^2)d\tau^2 = 1$. Finally, when we integrate δ_1 and δ_2 over a bounded set specified in Lemma 1, $\int_{\delta_1,\delta_2} d\delta_1 d\delta_2$ becomes finite, which shows that the joint posterior density is proper. □