Probabilistic Characterization of Wind Diurnal Variability for Wind Resource Assessment

Youngchan Jang, Eunshin Byon, *Member, IEEE*

***Abstract*—As wind energy penetration is expected to grow in the future, wind resource assessment becomes important in modern power grid operations. Selecting an appropriate wind farm site can benefit from understanding nonstationary characteristics of wind speeds. In particular, wind speed exhibits a diurnal pattern and the pattern varies, day-by-day and site-by-site. The goal of this study is to develop a new probabilistic modeling approach for quantifying variation in the wind diurnal pattern for assessing wind resource at unmonitored locations. Specifically, we formulate the coefficient of wind model as a latent random process and incorporate both day-to-day and spatial variability into the latent process. The estimation performance of the proposed approach is validated with actual data collected in west Texas. The results demonstrate that our approach can capture both spatially- and daily-varying patterns and quantify the uncertainty successfully.**

***Index Terms*—Bayesian inference, latent process, spatial analysis, wind energy**

I. Introduction

Selecting an appropriate wind farm site is vital for the success of wind energy in both financial and operational aspects. Among several factors to be considered for assessing potential wind farm site suitability, wind speed is clearly one of the most important factors. In general, windy areas are desired for installing a new wind farm. However, due to the wind's nonstationary characteristics, quantifying the wind variability is also inevitable for effective power grid operations [1].

One of the most effective ways for comprehensive wind resource assessment is to construct a wind model that characterizes its diurnal pattern. However, weather measurement at a candidate site may not be necessarily available in practice. For such a non-observational site, a meteorological tower can be installed to collect wind speed measurement. For more reliable wind resource assessment, it is essential to collect and analyze long-term wind resource (e.g., a year). However, installing new meteorological towers to collect long-term data is expensive and time-demanding for practical purposes.

Instead, one can collect wind measurement at a target site for a short-period of time. Such data collection activity is called measurement campaign [2], [3]. When a weather station, which collects long-term data, exists at a location close to the target site, the relationship of wind speeds between the two sites can be established and the speed at the target site can be estimated using the measurement at the weather station. Kwon [4] formulates the wind velocity at the target site as a linear function of the velocity at the weather station and estimates the linear function using the measurement collected at two locations. Jung *et al.* [5] further extend the approach in [4] and propose the Bayesian framework to handle various types of uncertainties due to limited data collected during the short-term measurement campaign. Similarly, Martinez-Cesena *et al.* [6] use the linear model between the annual average mean wind speeds at the target and measured locations. Even though these studies do not require long-term collection of measurement at the target location, short-term measurement is still needed. Moreover, they generally focus on quantifying the annual distribution of wind speed without considering its time-varying and nonstationary characteristics.

Another approach is to use numerical weather prediction (NWP) model. Zhang *et al.* [2] compare several NWP-based wind resource assessment methods using three datasets, including the Modern-Era Retrospective Analysis for Research and Applications dataset which is a low-resolution dataset, the Wind Integration National Dataset (WIND) which is a high-resolution dataset based on the Weather Research and Forecasting (WRF) model, and short-term campaign measurement. It was shown that the analog ensemble method, which integrates the low-resolution NWP dataset with the short-term measurement, provides the best estimate of wind distribution in most sites considered in their study, whereas the WIND is suitable for estimating the distribution of the difference between two consecutive hourly wind speeds. Jimenez *et al.* [7] compare two different weather prediction models at six locations, including offshore, onshore and island sites. In the study by [8], six different WRF simulations are conducted with different initial and boundary conditions and their estimation performances are compared with measured data at thirteen weather stations in Portugal. Their study shows that the new initial and boundary datasets improve the prediction accuracy over the old datasets. However, running NWP models requires considerable computational burden, and appropriate initial and boundary conditions need to be set *a priori*.

Unlike these prior studies, we consider a case where wind measurement at nearby locations are available. Recent advances in sensing technology make meteorological measurement increasingly available at many locations. This motivates us to assess wind resource when measurement at the target site does not exist, but data near the target site is available. Some recent studies propose a spatial model for predicting wind speed at a non-observational location. Lenzi *et al.* [9] apply the Gaussian process (GP) to wind measurements collected at neighbor locations at each time point. Byon *et al.* [10] spatially

Y. Jang and E. Byon are with the Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI, 48109, USA, E-mail: mapsossa@umich.edu; ebyon@umich.edu. This work was supported by the National Science Foundation under Grants IIS-1741166 and ECCS-1709094. (Corresponding author: Eunshin Byon.)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSTE.2020.2965444, IEEE Transactions on Sustainable Energy

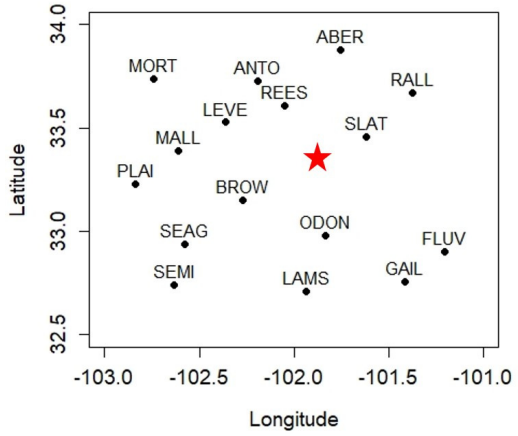


Fig. 1: Layout of multiple stations in west Texas region.

interpolate wind speeds at neighbor monitored stations for estimating the speed at a target unmonitored station. Although the models in [9], [10] provide estimates at non-observational sites, they only provide a snapshot estimate at each time instance, rather than providing wind models at the target site. As such, the spatial snapshot approach cannot fully describe the wind variability over time.

To illustrate, Fig. 1 shows a partial layout of actual meteorological stations in West Texas Mesonet [11]. West Texas Mesonet is an integrated network of meteorological monitored stations designed to observe meteorological conditions in the west Texas region [11]. The x- and y-axes in Fig. 1 represent the longitude and latitude, respectively. The solid circles denote the meteorological stations where wind speed data is collected. Suppose that the red star in Fig. 1 is a potential wind farm site where wind measurement is not available. Fig. 2 shows hourly average wind speeds at three stations, ANTO, REES, and SEMI, during the first week of January 2008. It is observed that overall wind speed pattern during each day shows a diurnal cycle and the diurnal patterns among the three stations are highly correlated. In particular, the patterns in ANTO and REES are more similar than that in SEMI, because ANTO and REES are more closely located. Therefore, we can borrow information of time-series characteristics at neighbor stations to assess wind resources at the target site.

It is also observed that diurnal cycles change day-by-day. For example, the pattern on the first day is quite different from that on the last day in Fig. 2. Fig. 3 further shows daily patterns at BROW during January 2008, where the thick curve represents the average diurnal pattern during January. Although there is commonality, wind patterns substantially differ day-by-day. As such, one cannot fully characterize wind variability with the average pattern only. Therefore, the wind resource assessment requires thorough understanding of the spatially- and daily-varying nonstationary characteristics.

This study develops a systematic approach to estimate diurnal patterns of wind speed and to quantify estimation uncertainties by using measurements collected at spatially dispersed nearby stations. We present a new modeling approach that formulates the time-varying pattern with daily- and spatially-varying coefficients. The parameters in the proposed model are estimated in a Bayesian hierarchical framework.

The main contribution of this study is two-fold: (1) Unlike the aforementioned studies that use the short-term measurement campaign data and/or NWP data, our approach uses wind measurement collected at nearby locations; (2) In contrast to the studies in [9], [10], the proposed approach provides a probabilistic wind model, which enables us to fully characterize the time-varying pattern of wind speed and quantify the uncertainties. The resulting model can generate scenarios of wind speed trajectories, which can be used for investment decision-making in wind power projects [6]. A case study is carried out using actual data collected in West Texas Mesonet. The implementation results demonstrate that the proposed approach is capable of successfully characterizing the wind variability at unmonitored sites, which provides useful insights for wind resource assessment.

The remainder of this paper is organized as follows. Section II discusses the proposed modeling approach and parameter estimation procedure. Section III presents a case study and Section IV concludes the paper.

II. Methodology

A. Integrative Modeling Approach

This section develops an integrative framework for quantifying the day-to-day and spatial variability in wind's diurnal pattern at non-observational locations. The underlying idea is as follows. We formulate the wind model using trigonometric functions to characterize a nonstationary pattern. Considering that diurnal patterns at neighbor locations should exhibit similarity and could change over different days, we make the model coefficients spatially correlated and daily-varying.

Let $Y(s,d,t)$ denote the wind speed at a location s at time t on day d. In this study, an hourly average measurements are considered, but the proposed approach can be applied to data with different temporal resolutions. To address the cyclic diurnal pattern, the wind speed, $Y(s,d,t)$, is formulated using L pairs of trigonometric functions [12], [13] as follows.

$$Y(s,d,t) = \mu(s,d,t) + \epsilon(s,d,t) \tag{1}$$

with

$$\mu(s,d,t) = \beta_0(s,d) + \sum_{\ell=1}^{L}\left[\beta_{1,\ell}(s,d)\sin\frac{2\ell\pi t}{24} + \beta_{2,\ell}(s,d)\cos\frac{2\ell\pi t}{24}\right], \tag{2}$$

where $\epsilon(s,d,t) \sim N(0,\sigma^2)$ is a Gaussian random noise. As a remark, an additional seasonal cycles can be included in $\mu(s,d,t)$ [10]. But such model assumes that the diurnal pattern remains the same in different seasons, which may not hold in practice. Instead, we suggest building monthly models with the formulation in (2), so heterogeneous diurnal patterns which could vary, depending on seasons, can be captured.

To capture day-to-day and location-to-location variations, we formulate each model coefficient as a latent process and decompose it into day-specific and site-specific random effects

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSTE.2020.2965444, IEEE Transactions on Sustainable Energy

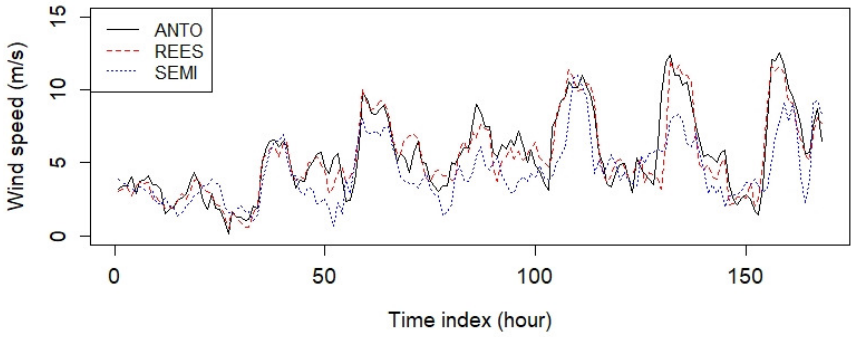


Fig. 2: Wind speed patterns at three stations, ATNO, REES, and SEMI in west Texas in the first week of January 2008.

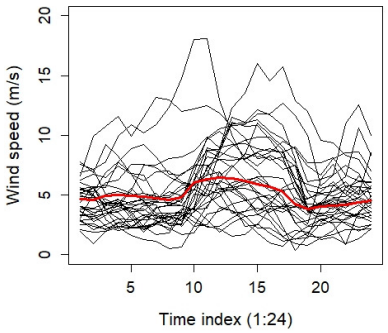


Fig. 3: Day-to-day variations at BROW on January 2008.

(Fig. 4). The day-specific latent process captures the day-to-day variation, whereas the site-specific random effects quantify the spatial correlation among multiple sites. Specifically, let $\beta(s,d)$ denote a vector of model coefficients, i.e.,

$$\beta(s,d) = [\beta_0(s,d), \beta_{1,1}(s,d), \beta_{2,1}(s,d), \ldots, \beta_{1,L}(s,d), \beta_{2,L}(s,d)]^T. \quad (3)$$

and $\beta^i(s,d)$ denote the i^{th} coefficient of $\beta(s,d)$. Depending on the flexibility to specify $\beta^i(s,d)$, three different models, referred to as integrative model 1, 2 and 3 (shortly, IM1, IM2 and IM3) are proposed as follows.

- IM1 assumes that the day-specific parameter on day d are dependent on that on day $d-1$, implying that wind diurnal patterns on the two consecutive days are similar each other.
- IM2 allows more flexibility in describing day-to-day variations. It lets the day-specific parameter randomly vary day-by-day. Thus, IM2 is appropriate when the daily pattern changes significantly.
- IM3 further allows the spatial correlation structure to be heterogeneous on different days, in contrast to IM1 and IM2 which implicitly assume the homogeneous spatial correlation structure.

Below we describe each model in more detail.

1) Integrative model 1 (IM1): To quantify spatial and daily variations, $\beta^i(s,d)$ is decomposed into two components as

$$\beta^i(s,d) = \beta^i_S(s) + \beta^i_D(d) \quad (4)$$

where $\beta^i_D(d)$ is a day-specific coefficient and $\beta^i_S(s)$ is a location-specific coefficient. Fig. 4 shows the overall framework of IM1 model.

First, to capture the spatial correlation, the site-specific coefficient $\beta^i_S(s)$ in (4) is formulated as a spatially-varying parameter [14], [**?**]. It should be noted that the pattern at closely located sites is similar to one another, as observed in Fig. 2. Accordingly, $\beta^i_S(s_j)$ should be similar to $\beta^i_S(s_k)$ for closely located sites, s_j and s_k. To characterize such spatial dependency, $\beta^i_S(s)$ is modeled with the latent GP [16] as

$$\beta^i_S(s) \sim \mathcal{GF}(\mu_i(s), C_i), \quad (5)$$

for $i = 1, 2, \cdots, 2L+1$, where μ_i and C_i denote the mean and covariance functions for $\beta^i_S(s)$, respectively.

By the consistency property of GP, a collection of $\beta^i_S(s)$'s jointly follow multivariate normal distribution [16]. Suppose that there are N monitored stations, $s = s_1, s_2, \cdots, s_N$. Let $\beta^i_{S,obs}$ denote an $N \times 1$ vector of $\beta^i_S(s)$'s, i.e., $\beta^i_{S,obs} = [\beta^i_S(s_1), \beta^i_S(s_2), \cdots, \beta^i_S(s_N)]^T$. Then, we have

$$\beta^i_{S,obs} \sim MVN(0, \Sigma_i), \quad (6)$$

where Σ_i denotes an $N \times N$ covariance matrix whose $(j,k)^{th}$ component, $c_i(s_j, s_k)$, is the covariance function between stations s_j and s_k. Here $c_i(s_j, s_k)$ is a positive definite kernel function. Among several choices for modeling $c_i(s_j, s_k)$, one of the commonly used covariance functions is the Matérn covariance function, defined as

$$c_i(s_j, s_k) = \frac{\tau_i^2}{2^{\nu-1}\Gamma(\nu)}(\kappa_i||x_j - x_k||)^\nu K_\nu(\kappa_i||x_j - x_k||), \quad (7)$$

where x_j is the location of station s_j, τ_i^2 is the marginal variance, K_ν is the modified Bessel function of second kind of order $\nu > 0$, $\Gamma(\cdot)$ is the Gamma function, $||\cdot||$ is the Euclidean distance, and κ_i is the decay parameter [17]. The parameter, ν, is a smoothness parameter, affecting differentiability of the underlying process. In general, ν is fixed to 1 for computational convenience [18]. Due to its flexibility and computational advantage, the Matérn covariance function is employed in our implementation, however, other covariance functions can be employed in the proposed framework [19].

Using the coefficients, $\beta^i_S(s)$'s ($s = s_1, s_2, \cdots, s_N$), the coefficient at an unmonitored site is obtained. Let s_0 denote a non-observational location. Given the parameter vector, $\beta^i_{S,obs}$, the coefficient, $\beta^i_S(s_0)$, at the non-observational site becomes normally distributed as

$$\beta^i_S(s_0)|\beta^i_{S,obs} \sim N(\mu_i(s_0), \tau_i^2(s_0)), \quad (8)$$

with

$$\mu_i(s_0) = c_i(s_0)^T \cdot \Sigma_i^{-1} \cdot \beta^i_{S,obs}, \quad (9)$$

$$\tau_i^2(s_0) = \tau_i^2 - c_i(s_0)^T \Sigma_i^{-1} c_i(s_0), \quad (10)$$

where $c_i(s_0) = [c_i(s_0, s_1), c_i(s_0, s_2), \cdots, c_i(s_0, s_N)]^T$ is an $N \times 1$ vector for $i = 1, 2, \cdots, 2L+1$. The results implies that

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSTE.2020.2965444, IEEE Transactions on Sustainable Energy

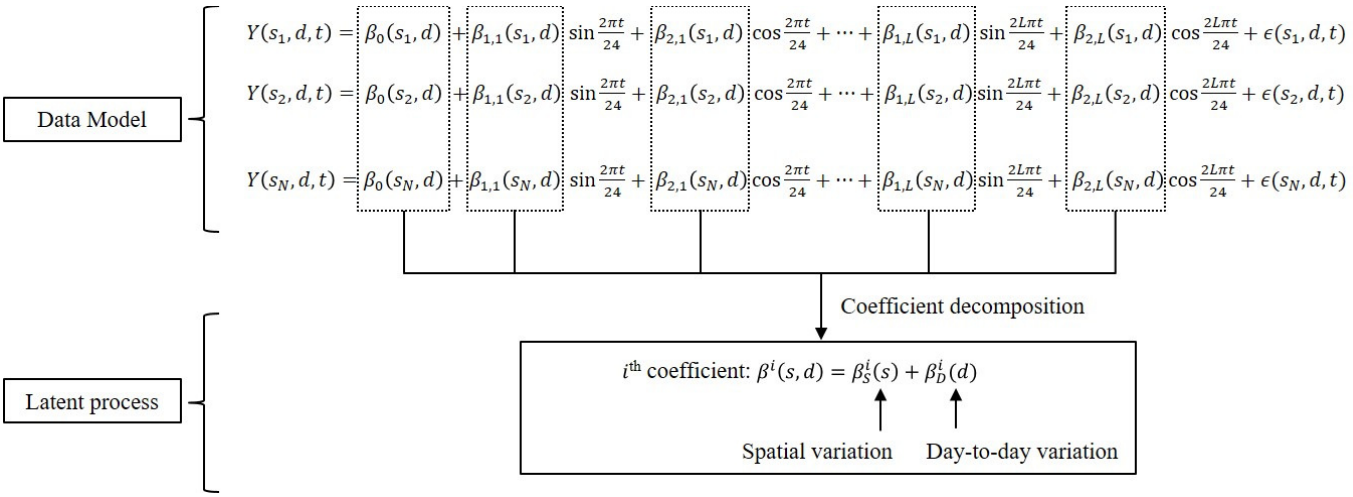


Fig. 4: Overall framework of the integrated model 1 (IM1).

once the spatial parameters at monitored sites are estimated, the parameter at an unmonitored site can be estimated accordingly. While we present the estimation procedure at a single unmonitored site in (8)-(10), the results can be extended for simultaneously estimating parameters at multiple sites [16].

Next, in describing the day-to-day variability, we observe that the daily pattern in one day tends to be similar to the pattern in the next day (Fig. 2). Therefore, the day-specific parameter, $\beta^i_D(d)$, could be highly correlated with $\beta^i_D(d-1)$. To characterize such temporal correlation, we apply the autoregressive (AR) process to $\beta^i_D(d)$ as

$$\beta^i_D(d) = \beta^i_{D,0} + \rho_i \beta^i_D(d-1) + \epsilon^i_D, \tag{11}$$

where ϵ^i_D denotes the random noise, $\epsilon^i_D \sim N(0, \delta^2_{1i})$. Here, we present AR1 for simplicity, but a higher order AR process can be employed.

2) Integrative model 2 (IM2): The second model, IM2, uses the same decomposition structure in (4) with the same spatial term. However, unlike IM1 that restricts the day-specific parameter to be temporally correlated, IM2 lets $\beta^i_D(d)$ be fully random day-by-day.

Specifically, $\beta^i_D(d)$ is formulated as random effects. Let β^i_D denote $D \times 1$ vector of $\beta^i_D(d)$'s, i.e., $\beta^i_D = [\beta^i_D(1), \beta^i_D(2), \ldots, \beta^i_D(D)]^T$. Then we have

$$\beta^i_D \sim MVN(0, \delta^2_{2i} I), \tag{12}$$

for $i = 1, 2, \ldots, 2L+1$, where I is a $D \times D$ identity matrix and δ^2_{2i} is the corresponding variance term.

3) Integrative model 3 (IM3): In IM1 and IM2, the spatial correlation is assumed to be homogeneous in different days (note that $\beta^i_S(s)$ is the same for all d's in (4)). To allow the heterogeneous spatial correlation structure on different days, IM3 breaks down $\beta^i(s,d)$ into two components as follows.

$$\beta^i(s,d) = \beta^i_S(s,d) + \beta^i_D(d). \tag{13}$$

Note that the spatial effect, $\beta^i_S(s,d)$, also depends on d, unlike $\beta^i_S(s)$ that depends on s only in IM1 and IM2 (see (4)).

As in IM1 and IM2, $\beta^i_S(s,d)$ is modeled as a latent GP. Let $\beta^i_{S,obs}(d)$ denote an $N \times 1$ vector of $\beta^i_S(s,d)$'s at monitored stations on day d, i.e., $\beta^i_{S,obs}(d) = [\beta^i_S(s_1,d), \beta^i_S(s_2,d), \cdots, \beta^i_S(s_N,d)]^T$. Then,

$$\beta^i_{S,obs}(d) \sim MVN(0, \Sigma_i(d)), \tag{14}$$

Here, $\Sigma_i(d)$ is an $N \times N$ covariance matrix on day d. It is assumed that $\beta^i_{S,obs}(d)$ is independent of $\beta^i_{S,obs}(d')$ for $d \neq d'$.

For the day-specific parameter $\beta^i_D(d)$, it is modeled as random effects in (12) as in IM2. Table I summarizes the decomposition structure in three integrative models.

B. Implication

Before discussing the parameter estimation procedure, it is worthwhile to discuss the primary difference between the proposed approach and the snapshot approach in [9], [10]. The snapshot approach directly formulates the correlation among $Y(s,d,t)$'s through interpolation techniques such as GP and kriging. For example, GP is applied to the measurement at monitored locations to estimate wind speed at the unmonitored location at each time instant. The salient feature of the proposed approach is that we characterize the correlation structure through the latent process, $\beta^i(s,d)$'s, instead of $Y(s,d,t)$'s. Assuming that the day-specific effect and spatial random effect, are independent, the covariance of model coefficients in IM1 is given by

$$\begin{aligned} &Cov(\beta^i(s_j,d), \beta^i(s_k,d')) \\ &= Cov(\beta^i_D(d), \beta^i_D(d')) + Cov(\beta^i_S(s_j), \beta^i_S(s_k)) \end{aligned} \tag{15}$$

where $Cov(\beta^i_S(s_j), \beta^i_S(s_k))$ is given in (7) and $Cov(\beta^i_D(d), \beta^i_D(d'))$ in AR1 [20] is

$$Cov(\beta^i_D(d), \beta^i_D(d')) = \delta^2_{1i} \frac{\rho_i^{|d-d'|}}{1-\rho_i^2}. \tag{16}$$

Therefore, we get

$$Cov(\beta^i(s_j,d), \beta^i(s_k,d')) = \begin{cases} \tau_i^2 + \delta^2_{1i}\frac{1}{1-\rho_i^2}, & \text{for } j=k, d=d' \\ c_i(s_j,s_k) + \delta^2_{1i}\frac{1}{1-\rho_i^2}, & \text{for } j\neq k, d=d' \\ \tau_i^2 + \delta^2_{1i}\frac{\rho_i^{|d-d'|}}{1-\rho_i^2}, & \text{for } j=k, d\neq d' \\ c_i(s_j,s_k) + \delta^2_{1i}\frac{\rho_i^{|d-d'|}}{1-\rho_i^2}, & \text{for } j\neq k, d\neq d' \end{cases} \tag{17}$$

TABLE I: Decomposition structure in latent process

	IM1	IM2	IM3
Decomposition model	$\beta^i(s,d) = \beta^i_S(s) + \beta^i_D(d)$	$\beta^i(s,d) = \beta^i_S(s) + \beta^i_D(d)$	$\beta^i(s,d) = \beta^i_S(s,d) + \beta^i_D(d)$
Latent spatially-varying parameter	$\beta^i_S(s) \sim GP$ (homogeneous spatial pattern)	$\beta^i_S(s) \sim GP$ (homogeneous spatial pattern)	$\beta^i_S(s,d) \sim GP(d)$ (heterogeneous spatial pattern)
Latent daily-varying parameter	$\beta^i_D(d) \sim AR$ (temporarily correlated diurnal pattern)	$\beta^i_D(d) \sim$ *random effects* (temporarily uncorrelated diurnal pattern)	$\beta^i_D(d) \sim$ *random effects* (temporarily uncorrelated diurnal pattern)

where $c_i(s_j, s_k)$ is the $(j,k)^{th}$ component of the covariance matrix, Σ_i, in (14). Note that $\beta^i_D(d)$ and $\beta^i_D(d')$ are strongly correlated when d and d' are closer, thereby making $\beta^i(s_j, d)$ and $\beta^i(s_j, d')$ similar to each other. Likewise, $\beta^i_S(s_j)$ and $\beta^i_S(s_k)$ at closely located s_j and s_k have larger covariance, $c_i(s_j, s_k)$. The covariance structures in IM2 and IM3 can be similarly specified. We omit them to save space.

C. Parameter estimation

This section discusses the parameter estimation procedure. We focus our discussion on estimating parameters in IM1. The parameter estimation in IM2 and IM3 can be performed in a straightforward way. We use wind measurement at N monitored stations at time $t = 1, 2, \ldots, T$ (e.g., $T = 24$ for hourly collected data) during D days ($d = 1, 2, \cdots, D$). Because the day-specific and site-specific parameters are formulated as latent processes, the proposed model has a multi-level hierarchical structure. The first level formulates the data model in (1). The second level specifies the latent processes for the spatially- and daily-varying parameters. The last level provides a prior density for hyperparameters.

We estimate the parameters in the Bayesian inference framework [21]. Let $\mathcal{D}$ denote the dataset used for inference, and let Θ denote a set of all parameters in the model. The joint posterior density of Θ is given by

$$\begin{aligned} p(\Theta|\mathcal{D}) \propto & \left[\Pi_{s,d,t} f(Y(s,d,t)|\mu(s,d,t),\sigma^2)\right] p(\sigma^2) \\ & \times \Pi_i \left[MVN(0, \mathbf{\Sigma}_i) p(\rho_i) p(\delta^2_{1i}) p(\tau^2_i) p(\kappa_i)\right], \end{aligned} \tag{18}$$

where $f(Y(s,d,t)|\mu(s,d,t),\sigma^2)$ represents the likelihood of wind speed with $Y(s,d,t) \sim N(\mu(s,d,t),\sigma^2)$, $p(\rho^i)$ and $p(\delta^2_{1i})$ imply the priors in the latent AR process for the daily-varying coefficient in (12), and $MVN(0, \mathbf{\Sigma}_i)$ denotes the latent Gaussian process in (14) for spatially-varying coefficients. Lastly, $p(\rho_i)$, $p(\delta^2_{1i})$, $p(\tau^2_i)$, $p(\kappa_i)$, and $p(\sigma^2)$ denote prior densities for their corresponding parameters and hyperparameters.

The posterior mean from the posterior density in (18) is used for estimating parameters. Obtaining the posterior density requires multi-dimensional integration, and it is not derived in a closed form. Therefore, simulation-based methods such as Markov chain Monte Carlo (MCMC) can be used to approximate the posterior density. However, implementing MCMC demands expensive computational cost, so we use an approximation method. In particular, we employ the integrated nested Laplace approximations (INLA) in our analysis [19]. For more details on the INLA approximation procedure, please refer to [19]. In our analysis, 'R-INLA' package in the statistical software, R [18] is used. In our implementation, priors are specified as suggested in INLA. When parameters are estimated in the Bayesian hierarchical framework, it has been known that the deviance information criterion (DIC) is useful for choosing a model order, L, in our model [22]. For more details on DIC, please refer to [22].

Once the parameters are estimated, the estimated distribution of the wind speed at the target site s_0 is provided by

$$Y(s_0,d,t) \sim N(\hat{\mu}(s_0,d,t), \hat{\sigma}^2(s_0,d,t)) \tag{19}$$

with

$$\begin{aligned} \hat{\mu}(s_0,d,t) = & \hat{\beta}_0(s_0,d) \\ & + \sum_{\ell=1}^{L} \left[\hat{\beta}_{1,\ell}(s_0,d)\sin\frac{2\ell\pi t}{24} + \hat{\beta}_{2,\ell}(s_0,d)\cos\frac{2\ell\pi t}{24}\right], \end{aligned} \tag{20}$$

where $\hat{\beta}(\cdot)$'s in (20) denote the posterior means for the corresponding parameters and $\hat{\sigma}^2(s_0,d,t)$ is the posterior variance of $Y(s_0,d,t)$.

III. CASE STUDY

We use wind measurement collected at 16 stations in West Texas Mesonet in this study. The shortest and average distances between two adjacent stations are 18.6 km and 36.3 km, respectively. The location information of the stations can be found in [11].

As discussed earlier, we suggest building monthly diurnal models to account for heterogeneous diurnal patterns in different seasons. Wind resource assessment requires quantification of year-long wind pattern. Due to the time limitation, we were not able to estimate year-long pattern, but we implement our method using four months data, including January, April, July and November in 2008. The original data contains 5-minute average wind speeds at a height of 10 meter above the surface. In this study, hourly-averaged wind speeds are used.

A. Alternative two-step approach

This section presents an alternative approach that extends the spatial snapshot approach [10]. First, the snapshot estimate, $\tilde{Y}(s_0,d,t)$, for the target station, s_0, is estimated by spatially interpolating wind speeds at neighbor stations, $Y(s,d,t)$ ($s = s_1, s_2, \cdots, s_N$), through the ordinary kriging as

$$\tilde{Y}(s_0,d,t) = \mathbf{w}^T \mathbf{Y}_{d,t}, \tag{21}$$

for each d and t, $d = 1, 2, \ldots, D$, and $t = 1, 2, \cdots, T$. Here $\mathbf{Y}_{d,t}$ is a vector whose i^{th} component is $Y(s_i,d,t)$ and $\mathbf{w} = [w_1, \ldots, w_N]^T$ is the weight matrix, defined as

$$\mathbf{w} = \mathbf{C}^{-1}\mathbf{c} - \frac{\mathbf{C}^{-1}\mathbf{1}\mathbf{1}^T\mathbf{C}^{-1}\mathbf{c}}{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}} + \frac{\mathbf{C}^{-1}\mathbf{1}}{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}}, \tag{22}$$

where $\mathbf{C}$ is an $N \times N$ covariance matrix among $\mathbf{Y}_{d,t}$, $\mathbf{c}$ is an $N \times 1$ dimensional covariance between $\mathbf{Y}_{d,t}$ and $\tilde{Y}(s_0, d, t)$ and $\mathbf{1}$ is an $N \times 1$ dimensional vector with 1 elements. Although we present the ordinary kriging, other kriging models (e.g., universal kriging), or GP, can be employed for performing the spatial interpolation.

Once the snapshot estimate is estimated, the time series model at the unmonitored site, s_0, is fitted as follows.

$$\tilde{Y}(s_0, d, t) = \beta_0(s_0, d) + \sum_{\ell=1}^{L} \left[\beta_{1,\ell}(s_0, d) \sin \frac{2\ell\pi t}{24} + \beta_{2,\ell}(s_0, d) \cos \frac{2\ell\pi t}{24}\right] + \sum_{i=1}^{p} \gamma_h(s_0, d)\tilde{Y}(s_0, d, t-i) + \epsilon(s_0, d, t), \quad (23)$$

where p denotes the model order in the AR process, which is decided based on the Akaike information criterion (AIC). The noise term, $\epsilon(s_0, d, t)$, is assumed to be an independent Gaussian random variable. We estimate parameters using maximum likelihood estimation.

Below this two-step alternative approach is summarized.

Step 1 At each time point, obtain a snapshot estimate by spatially interpolating the wind measurements collected at neighbor stations using (21)-(22).

Step 2 Fit the linear model with the snapshot estimates, using (23).

B. Implementation results

For evaluating the estimation performance, we divide the 16 stations into two sets: training set (in-sample) and testing set (out-of-sample). The training set includes measurements at 15 stations, representing observational sites, whereas the testing set contains data collected at the remaining station which represents a non-observational site. Using data from the 15 stations in the training set, we estimate the wind speed at the testing station and evaluate the prediction performance by comparing its estimated and measured wind speeds. This procedure is repeated 16 times to get all estimation results for 16 testing stations. Therefore, with four months data, our case study includes 16 stations $\times$ 4 months = 64 testing cases.

We measure the estimation performance with several criteria. First, for evaluating the point estimation capability, root mean square error (RMSE) is used. We also employ the continuous ranked probability score (CRPS) [23], [24], defined as follows, when parameters are estimated in the Bayesian framework [23], [24].

$$CRPS = \frac{1}{DT} \sum_{d=1}^{D} \sum_{t=1}^{T} \Big[\frac{1}{m} \sum_{j=1}^{m} |\hat{Y}^{(j)}(s_0, d, t) - Y(s_0, d, t)| - \frac{1}{2m^2} \sum_{j=1}^{m} \sum_{k=1}^{m} |\hat{Y}^{(j)}(s_0, d, t) - \hat{Y}^{(k)}(s_0, d, t)|\Big] \quad (24)$$

where m is the number of posterior samples in the posterior predictive density and $\hat{Y}^{(j)}(s_0, d, t)$ denotes the j^{th} samples. For the alternative approach, CRPS measure presented in [23] is used. The smaller CRPS indicates better performance.

Tables II and III summarize the RMSE and CRPS results for four months, respectively, for 16 testing stations where each testing station is considered as a non-observational site. Overall, the estimation performance at testing stations located in the center of monitored stations (e.g., REES and BROW) is generally better than that at boundary stations (e.g., ABER, GAIL and MORT). This is understandable because the central stations have more informative spatial information from their neighbor stations than boundary stations. Overall, IM3 consistently provides the smallest values in most testing sites in both criteria. For example, on average it generates 14% lower RMSE and 18% lower CRPS, compared to the alternative approach on January.

It is worthwhile to mention that, although IM1 and IM2 generate performance comparable to the alternative approach, they do so with much lower model complexity. In the alternative approach, three parameters need to be estimated to get a snapshot estimate using the ordinary kriging at each time instant. Therefore, it requires $3DT + 2L + p + 2$ parameters in total for each month. With $D = 31$, $T = 24$, $L = 5$, and $p = 2$, it uses 2,246 parameters. On the contrary, IM1 uses $4(2L + 1) + 1$ parameters (4 parameters, τ_i, κ_i, $\beta^i_{D,0}$, and ρ_i, for each i and variance parameter, σ^2), whereas IM2 uses $3(2L + 1) + 1$ parameters (3 parameters, τ_i, κ_i, and δ^2_{2i}, for each i and variance parameter). Thus, IM1 and IM2, respectively, employ 45 and 34 parameters only, which account for about 1.5% and 2% of the alternative's. With such remarkably smaller number of parameters, they lead to the estimation performance similar to the alternative approach. The number of parameters required in IM3 is larger than those in IM1 and IM2, because its spatial parameters differ day-by-day, however, IM3 still reduces the model complexity over the alternative approach by about 69%.

Another advantage of the proposed approach is that it can better quantify the estimation uncertainty. The proposed approach can obtain the posterior predictive density and prediction interval (PI) in the Bayesian framework. Fig. 5 presents the measured and estimated speeds at a testing station, BROW, for the first week of January, along with PI. The bold central lines denote the predicted values and the dotted upper and lower lines denote the 90% PIs. It shows that most estimated speeds in IM1, IM2, and IM3 belong to the PIs.

However, it is difficult, if not possible, to accommodate all uncertainties in the alternative approach, because it characterizes the spatial and temporal correlation separately through the two-step procedure. Therefore, we alternatively treat the snapshot estimates as real values and build the PI with the model in (23). Fig. 5(d) shows that the resulting PIs are unduly narrow and thus, several data points are located outside the intervals, indicating underestimated uncertainties. As a result, the coverage rate of the alternative model is much lower than ours. Here, the coverage rate implies the ratio of the number of estimates within PIs to the total number of estimates. Ideally, the coverage rate should be close to the nominal rate. For example, in January, the average coverage rate of the alternative approach remains at 67.6% for the 90% PI, whereas the coverage rates from IM1, IM2, and IM3 are 89.3%, 89.5%, and 86.7%, respectively, which are close to the nominal rate. It is also worthwhile to mention that the PIs from our approach are wider, because it fully quantifies

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSTE.2020.2965444, IEEE Transactions on Sustainable Energy

uncertainties for estimating spatial and day-to-day variability.

To further assess probabilistic estimation performance, reliability diagram [9] is employed. To construct the reliability diagram, an indicator variable that compares an actual speed, $Y(s_0, d, t)$ with its α-quantile forecast, $\hat{Y}^{(\alpha)}(s_0, d, t)$, for $0 \leq \alpha \leq 1$ is obtained as

$$I^{(\alpha)}_{s_0,d,t} = \begin{cases} 1, & \text{if } Y(s_0, d, t) \leq \hat{Y}^{(\alpha)}(s_0, d, t) \\ 0, & \text{if } Y(s_0, d, t) > \hat{Y}^{(\alpha)}(s_0, d, t), \end{cases} \tag{25}$$

Then, similar to the PI coverage, the empirical coverage in the reliability diagram is obtained as

$$\hat{a}^{(\alpha)}_{s_0} = \frac{1}{DT} \sum_{d=1}^{D} \sum_{t=1}^{T} I^{(\alpha)}_{s_0,d,t}. \tag{26}$$

The reliability diagram compares the empirical coverage with the nominal coverage. The empirical coverage becomes close to the nominal coverage, α, when the probabilistic estimation is performed appropriately. We compare the empirical coverage at nominal levels from 5% to 95% on increments of 5% in Fig. 6. The average reliability diagram in Fig. 6(c) is constructed by taking the average of empirical coverage from all 16 testing stations. The empirical coverage from IM1, IM2, and IM3 align with the diagonal line in all cases, while those from the two-step approach deviate from the diagonal line. This result demonstrates a stronger probabilistic assessment capability of the proposed approach over the alternative one.

In summary, although the proposed integrative and alternative two-step approaches provide comparable point estimation capability, our approach quantifies wind variability better so its estimated density is more accurate. Our strong probabilistic assessment performance is mainly due to the fact that our approach can capture different types of uncertainties arising from spatial and diurnal variations in an integrative way.

Among the studied models, the performance of IM1 and IM2 were comparable in most cases. This result coincides with our observation in Fig. 2 where the daily pattern is similar on consecutive days. Therefore, when the diurnal patterns do

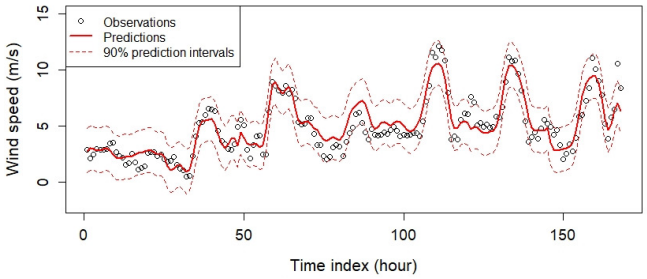
TABLE II: Comparison of RMSEs at 16 testing stations (unit: m/s, SD in the last row represents standard deviation)

	January				April				July				November			
	IM1	IM2	IM3	Two-step	IM1	IM2	IM3	Two-step	IM1	IM2	IM3	Two-step	IM1	IM2	IM3	Two-step
ABER	1.28	1.28	1.09	1.33	1.41	1.36	1.08	1.35	0.97	0.98	0.87	0.95	0.91	0.90	0.85	0.97
REES	0.96	0.96	0.78	0.96	1.07	1.07	0.88	1.08	1.07	1.08	0.92	1.04	0.77	0.78	0.69	0.80
RALL	1.21	1.21	1.10	1.10	1.23	1.25	1.10	1.23	0.99	0.99	0.87	0.91	1.02	1.04	0.97	1.02
ANTO	1.20	1.20	0.97	1.18	1.20	1.20	0.96	1.16	0.94	0.94	0.83	0.93	0.96	0.97	0.83	0.96
SLAT	1.33	1.34	1.23	1.37	1.44	1.46	1.31	1.45	1.20	1.20	1.08	1.17	1.18	1.18	1.16	1.20
LEVE	1.12	1.11	0.91	1.14	1.12	1.12	0.93	1.13	0.83	0.82	0.70	0.83	0.75	0.75	0.67	0.76
MORT	1.36	1.37	1.12	1.45	1.42	1.42	1.25	1.52	1.03	1.01	0.92	1.00	1.13	1.13	0.95	1.16
BROW	0.92	0.92	0.83	0.94	0.97	0.97	0.90	0.98	0.95	0.95	0.90	1.01	0.85	0.84	0.79	0.85
MALL	0.97	0.97	0.76	0.98	1.01	1.01	0.84	0.99	0.83	0.83	0.71	0.80	0.85	0.86	0.71	0.86
ODON	1.12	1.12	1.00	1.05	1.44	1.44	1.32	1.35	0.80	0.80	0.76	0.78	0.95	0.98	0.86	0.90
FLUV	1.48	1.49	1.19	1.15	1.55	1.56	1.30	1.26	1.09	1.10	1.00	1.06	1.35	1.36	1.22	1.20
PLAI	1.19	1.18	0.94	1.18	1.29	1.31	1.04	1.20	0.93	0.93	0.79	0.88	0.99	0.99	0.79	0.96
GAIL	1.48	1.47	1.31	1.38	1.63	1.61	1.40	1.52	1.16	1.17	1.10	1.07	1.21	1.21	1.17	1.11
SEAG	1.21	1.20	1.00	1.19	1.33	1.34	1.09	1.29	0.90	0.89	0.77	0.86	1.01	1.01	0.90	0.99
LAMS	1.16	1.14	0.96	1.10	1.31	1.31	1.10	1.25	0.96	0.96	0.88	0.97	1.06	1.02	0.86	0.91
SEMI	1.30	1.31	1.06	1.18	1.37	1.35	1.09	1.22	1.01	1.00	0.85	0.95	1.03	1.03	0.90	0.96
Average	1.20	1.20	1.01	1.17	1.30	1.30	1.10	1.25	0.98	0.98	0.87	0.95	1.00	1.00	0.90	0.98
(SD)	(0.17)	(0.17)	(0.15)	(0.15)	(0.19)	(0.19)	(0.17)	(0.16)	(0.11)	(0.11)	(0.12)	(0.11)	(0.16)	(0.11)	(0.17)	(0.13)

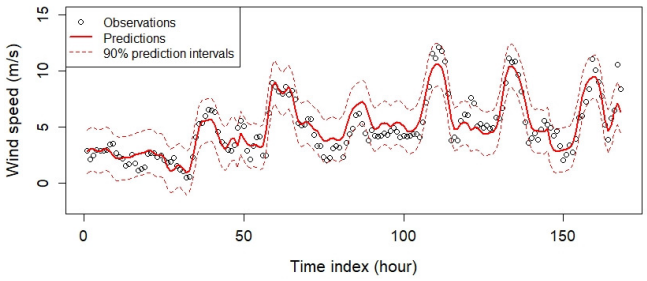
TABLE III: Comparison of CRPSs at 16 testing stations (unit: m/s, SD in the last row represents standard deviation)

	January				April				July				November			
	IM1	IM2	IM3	Two-step	IM1	IM2	IM3	Two-step	IM1	IM2	IM3	Two-step	IM1	IM2	IM3	Two-step
ABER	0.71	0.71	0.58	0.77	0.78	0.74	0.59	0.75	0.53	0.54	0.48	0.53	0.51	0.50	0.46	0.55
REES	0.54	0.54	0.42	0.53	0.60	0.60	0.48	0.59	0.60	0.60	0.52	0.62	0.44	0.43	0.38	0.46
RALL	0.68	0.68	0.61	0.53	0.68	0.69	0.60	0.68	0.55	0.55	0.48	0.54	0.57	0.58	0.53	0.61
ANTO	0.66	0.66	0.54	0.68	0.67	0.67	0.54	0.66	0.52	0.52	0.45	0.52	0.53	0.54	0.46	0.57
SLAT	0.74	0.75	0.68	0.84	0.80	0.81	0.74	0.87	0.68	0.67	0.61	0.71	0.67	0.67	0.66	0.74
LEVE	0.61	0.61	0.49	0.65	0.62	0.62	0.51	0.64	0.45	0.46	0.38	0.47	0.43	0.43	0.38	0.44
MORT	0.77	0.76	0.62	0.86	0.80	0.80	0.70	0.90	0.57	0.55	0.51	0.59	0.63	0.63	0.52	0.69
BROW	0.52	0.52	0.46	0.54	0.55	0.55	0.49	0.56	0.53	0.53	0.50	0.60	0.48	0.47	0.43	0.50
MALL	0.55	0.55	0.42	0.56	0.57	0.57	0.46	0.56	0.46	0.47	0.39	0.44	0.48	0.48	0.40	0.50
ODON	0.61	0.61	0.54	0.66	0.79	0.79	0.73	0.79	0.45	0.45	0.43	0.45	0.53	0.53	0.48	0.54
FLUV	0.86	0.85	0.67	0.66	0.88	0.89	0.74	0.72	0.61	0.61	0.57	0.64	0.77	0.77	0.69	0.71
PLAI	0.65	0.66	0.52	0.68	0.70	0.71	0.56	0.68	0.52	0.53	0.44	0.51	0.55	0.55	0.44	0.57
GAIL	0.82	0.82	0.74	0.86	0.93	0.92	0.82	0.96	0.66	0.66	0.64	0.66	0.68	0.68	0.67	0.66
SEAG	0.68	0.68	0.56	0.73	0.73	0.73	0.59	0.78	0.49	0.49	0.43	0.51	0.57	0.57	0.51	0.61
LAMS	0.63	0.64	0.54	0.65	0.72	0.73	0.61	0.73	0.53	0.52	0.48	0.58	0.60	0.58	0.49	0.55
SEMI	0.73	0.72	0.59	0.69	0.75	0.74	0.60	0.72	0.56	0.57	0.47	0.56	0.57	0.58	0.57	0.57
Average	0.67	0.67	0.56	0.68	0.72	0.72	0.61	0.72	0.54	0.54	0.49	0.56	0.56	0.56	0.50	0.58
(SD)	(0.10)	(0.10)	(0.09)	(0.11)	(0.11)	(0.11)	(0.11)	(0.12)	(0.07)	(0.07)	(0.07)	(0.08)	(0.09)	(0.09)	(0.10)	(0.09)

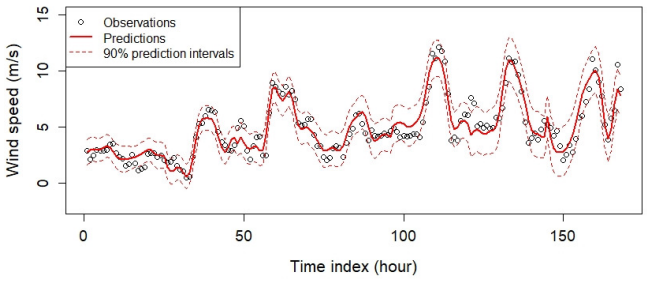
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSTE.2020.2965444, IEEE Transactions on Sustainable Energy



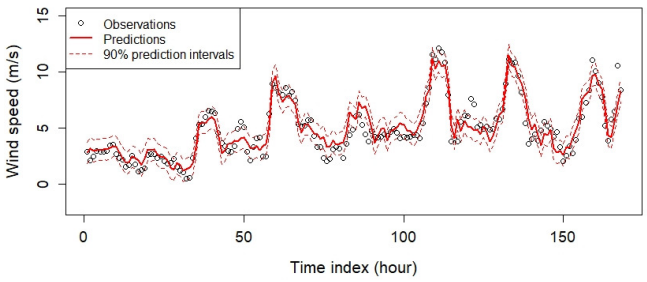
(a) IM1



(b) IM2

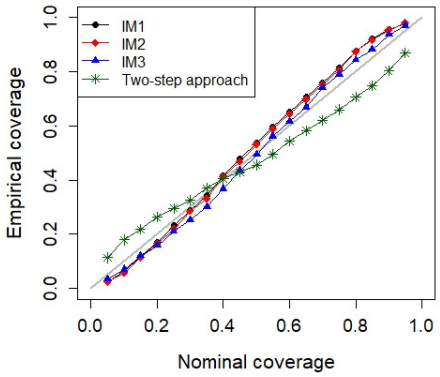


(c) IM3

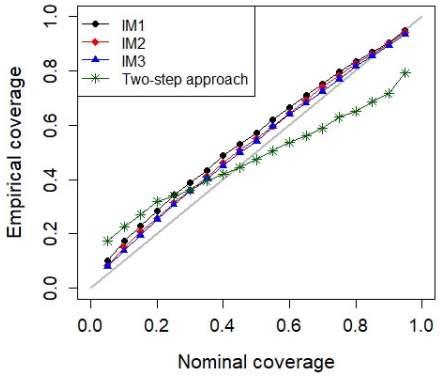


(d) Two-step approach

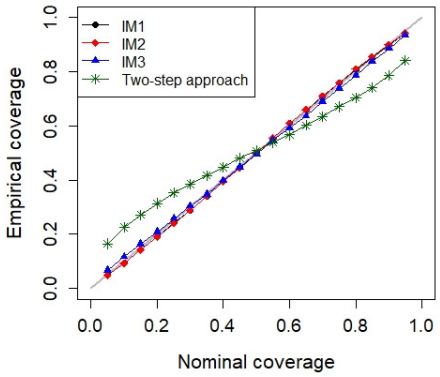
Fig. 5: Comparison of observed and predicted wind speeds and prediction intervals at the testing station, BROW, in January



(a) Testing station: BROW



(b) Testing station: MORT



(c) Average

Fig. 6: Reliability diagram in January

not rapidly change, either the AR formulation in (11) or random effect formulation in (12) would provide similar results. When the diurnal pattern changes considerably, we suggest the random effect formulation used in IM2. Regarding the spatial correlation, it has been known that the dominating wind direction substantially affects the correlation structure [21]. Therefore, IM3 would perform better than IM1 and IM2 when wind direction varies substantially even during the same month.

IV. Conclusion and Future Plans

This study develops a probabilistic model for assessing wind resource by characterizing the spatial and temporal correlations through the model parameters. Specifically, the model parameters are treated as latent spatially- and daily-varying random processes. Such collective treatment enables the proposed integrative approach to provide compelling capabilities for evaluating the wind variability at non-observational locations. A case study using actual data demonstrates that the proposed approach is capable of fully quantifying wind

variations, which provides insights for selecting wind farm locations.

In the future, we plan to extend our analysis to account for wind direction, other environmental factors [13], a site's local characteristics (e.g., terrain type) and information from physics-based NWP models for improving the prediction accuracy and applicability of the proposed approach.

REFERENCES

[1] F. Bouffard and F. D. Galiana, "Stochastic security for operations planning with significant wind power generation," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 306-316, 2008.

[2] J. Zhang, C. Draxl, T. Hopson, L. Delle Monache, E. Vanvyve, and B.-M. Hodge, "Comparison of numerical weather prediction based deterministic and probabilistic wind resource assessment methods," *Appl. Energy*, vol. 156, pp. 528-541, 2015.

[3] G. Lee, E. Byon, L. Ntaimo, and Y. Ding, "Bayesian spline method for assessing extreme loads on wind turbines," *Ann. Appl. Stat.*, vol. 7, no. 4, pp. 2034-2061, 2013.

[4] S.-D. Kwon, "Uncertainty analysis of wind energy potential assessment," *Appl. Energy*, vol. 87, no. 3, pp. 856-865, 2010.

[5] S. Jung, O. A. Vanli, and S.-D. Kwon, "Wind energy potential assessment considering the uncertainties due to limited data," *Appl. Energy*, vol. 102, pp. 1492-1503, 2013.

[6] E. A. Martinez-Cesena and J. Mutale, "Wind power projects planning considering real options for the wind resource assessment," *IEEE Trans. Sustain. Energy*, vol. 3, no. 1, pp. 158-166, 2011.

[7] B. Jimenez, F. Durante, B. Lange, T. Kreutzer, and J. Tambke, "Offshore wind resource assessment with WAsP and MM5: comparative study for the German Bight," *Wind Energy*, vol. 10, no. 2, pp. 121-134, 2007.

[8] D. Carvalho, A. Rocha, M. Gómez-Gesteira and C. S. Santos, "WRF wind simulation and wind energy production estimates forced by different reanalyses: Comparison with observed data for Portugal," *Appl. Energy*, vol. 117, pp. 116-126, 2014.

[9] A. Lenzi, P. Pinson, L. H. Clemmensen, and G. Guillot, "Spatial models for probabilistic prediction of wind power with application to annual-average and high temporal resolution data," *Stoch Environ Res Risk Assess*, vol. 31, no. 7, pp. 1615-1631, 2017.

[10] E. Byon, E. Pérez, Y. Ding, and L. Ntaimo, "Simulation of wind farm operations and maintenance using discrete event system specification," *Simulation*, vol. 87, no. 12, pp. 1093-1117, 2011.

[11] "West Texas Mesonet," accessed in May, 2008. [Online]. Available: http://www.mesonet.ttu.edu/

[12] T. Gneiting, K. Larson, K. Westrick, M. G. Genton, and E. Aldrich, "Calibrated probabilistic forecasting at the stateline wind energy center: The regime-switching space-time method," *J. Am. Stat. Assoc.*, vol. 101, no. 475, pp. 968-979, 2006.

[13] A. S. Hering and M. G. Genton, "Powering up with space-time wind forecasting," *J. Am. Stat. Assoc.*, vol. 105, no. 489, pp. 92-104, 2010.

[14] J. Lindström, A. A. Szpiro, P. D. Sampson, A. P. Oron, M. Richards, T. V. Larson, and L. Sheppard, "A flexible spatio-temporal model for air pollution with spatial and spatio-temporal covariates," *Environ. Ecol. Stat.*, vol. 21, no. 3, pp. 411-433, 2014.

[15] M. You, E. Byon, J. J. Jin, and G. Lee, "When wind travels through turbines: A new statistical approach for characterizing heterogeneous wake effects in multi-turbine wind farms," *IISE Trans.*, vol. 49, no. 1, pp. 84-95, 2017.

[16] C. E. Rasmussen, "Gaussian processes in machine learning," in *Advanced lectures on machine learning.* Springer, 2004, pp. 63–71.

[17] M. Stein, "Interpolation of spatial data: some theory for Kriging, Springer Science & Business Media, 1999.

[18] F. Lindgren and H. Rue, "Bayesian spatial modelling with R-INLA," *J. Stat. Softw.*, vol. 63, no. 19, 2015.

[19] H. Rue, S. Martino, and N. Chopin, "Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations, "*J. Royal Stat. Soc. Series B*, vol. 71, no. 2, pp. 319-392, 2009.

[20] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control.* Wiley, 2015.

[21] M. You, B. Liu, E. Byon, S. Huang, and J. J. Jin, "Direction-dependent power curve modeling for multiple interacting wind turbines," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1725-1733, 2018.

[22] D. J. Spiegelhalter, N. G. Best, B. P. Carlin, and A. Van Der Linde, "Bayesian measures of model complexity and fit," *J. Royal Stat. Soc. Series B*, vol. 64, no. 4, pp. 583-639, 2002.

[23] T. Gneiting, A. E. Raftery, A. H. Westveld III, and T. Goldman, "Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation," *Mon. Weather Rev.*, vol. 133, no. 5, pp. 1098-1118, 2005.

[24] E. P. Grimit, T. Gneiting, V. Berrocal, and N. A. Johnson, "The continuous ranked probability score for circular variables and its application to mesoscale forecast ensemble verification," *Q. J. R. Meteorolog. Soc.*, vol. 132, no. 621C, pp. 2925-2942, 2006.

PLACE PHOTO HERE

Youngchan Jang received his B.S. degree in mechanical engineering from Korea Military Academy, Korea, and M.S. degree in Industrial and Operations Engineering from the University of Michigan, Ann Arbor, MI, USA. He is currently working toward the Ph.D. degree in the Department of Industrial and Operations Engineering, University of Michigan. His research interest includes applied statistics, spatial data analytics, and quality and reliability engineering with applications on the power system.

PLACE PHOTO HERE

Eunshin Byon (S'09-M'20) is an associate professor in the Department of Industrial and Operations Engineering at the University of Michigan. She received her B.S. and M.S. in Industrial and Systems Engineering from the Korea Advanced Institute of Science and Technology (KAIST) and Ph.D. in Industrial and Systems Engineering from Texas A&M University. Her research interests include optimizing operations and management of wind power systems, data analytics, quality and reliability engineering, and uncertainty quantification. She is a member of IIE, INFORMS, and IEEE.