Spatio-temporal asymmetry of local wind fields and its impact on short-term wind forecasting.

Ahmed Aziz Ezzat, Mikyoung Jun, and Yu Ding, *Senior Member, IEEE*

***Abstract*—The massive amounts of spatio-temporal data collected in today's wind farms have created a necessity for accurate spatio-temporal models. Despite the growing recognition for non-separable spatio-temporal models, a significant reliance on separable, symmetric models is still the norm in today's renewable industry. We discover that the broad use of separable models is due to the handling of wind data in a setting that does not reveal their fine-scale spatio-temporal structure. The contribution of this research is two-fold. First, we devise a special pair of spatio-temporal "lens" that allows us to see the fine-scale spatio-temporal variations and interactions, and subsequently, we conclude that local wind fields exhibit strong signs of non-separability and asymmetry. Using one year of turbine-specific wind measurements, we show that asymmetry can in fact be detected in more than 93% of the time. Second, making use of the spatio-temporal lens, we propose an enhanced procedure for short-term wind speed forecast. Substantial improvements in forecast accuracy in both wind speed and wind power were observed. When combined with certain intelligent methods such as support vector machine, additional improvements are possible.**

***Index Terms*—Asymmetry, kriging, spatio-temporal statistics, wind energy, forecasting.**

I. Introduction

THE installation of large-scale wind farms has been growing at a fast pace to meet the world's increasing demand in clean energy. According to the 2015 Wind Vision report by the U.S. Department of Energy, wind power supplies are projected to meet 20% of the United States' electricity demand by 2030 [1]. The increase in size and number of wind farms goes hand-in-hand with growing complexities in the associated operational analytics such as wind forecasts and power estimations, which further impacts decision making in turbine control, maintenance, and economic dispatch. These activities can benefit greatly from a better understanding of the spatio-temporal wind dynamics on a wind farm. Modern wind farms are in fact equipped with sensors measuring wind characteristics in a fine spatio-temporal scale, providing an unprecedented opportunity to look into local wind dynamics.

Spatio-temporal models have emerged in the geo-statistic field to model random processes evolving through space and time. Let $Y(\mathbf{s},t)$ be a random process such that $(\mathbf{s},t) \subset \mathbb{R}^d \times \mathbb{R}$, then the key aspect is to model the covariance structure of Y through imposing a positive-definite parametric function form $\tilde{K}(\cdot,\cdot,\cdot,\cdot)$ such that $Cov\{Y(\mathbf{s}_1,t_1), Y(\mathbf{s}_2,t_2)\} = \tilde{K}(\mathbf{s}_1,\mathbf{s}_2,t_1,t_2)$; for details about covariance structures or covariance functions, please refer to the classical texts [2], [3]. By assuming stationarity, the covariance structure, also referred to as a covariance function, only depends on the spatial and temporal lags, $\mathbf{u} = \mathbf{s}_1 - \mathbf{s}_2$ and $h = t_1 - t_2$, respectively, and as such, it is denoted by $K(\mathbf{u},h)$; examples of covariance structure K are given later in Equations (3) and (4). Intuitively, the covariance structure K defines the similarity between a pair of spatio-temporal data points $Y(\mathbf{s}_1,t_1)$ and $Y(\mathbf{s}_2,t_2)$ based on their spatial and temporal separations, $\mathbf{u}$ and h, respectively.

Until the beginning of the 21st century, separable spatio-temporal models dominated the related literature [3]. A covariance structure is said to be separable if it factors into the product of purely spatial and temporal components such that $K(\mathbf{u},h) = K^s(\mathbf{u}) \cdot K^t(h)$. Despite their computational benefits, separable models have limiting modeling capabilities which seldom align with reality. For example, separability overlooks the interaction between the spatial and temporal components and implies full-symmetry in the spatio-temporal covariance structure [3]. A number of works have pointed out that full-symmetry is not a physically justifiable assumption for large-scale atmospheric processes in which there exists a dominant air or water flow over time, making the correlation in one direction often stronger than in other directions [4]–[9]. A covariance structure is symmetric if $K(\mathbf{u},h) = K(-\mathbf{u},h) = K(\mathbf{u},-h) = K(-\mathbf{u},-h)$ [4]. In other words, the correlation between sites $\mathbf{s}_1$ and $\mathbf{s}_2$ at times t_1 and t_2, is the same as that between $\mathbf{s}_1$ and $\mathbf{s}_2$ at times t_2 and t_1. Separability implies symmetry, but the converse cannot be guaranteed [7].

The limiting capabilities of separable models in capturing the complex correlation patterns in real-world processes motivated the geo-statistic community in the past decade to establish classes of non-separable, but still symmetric, models that capture spatio-temporal interactions [4], [10], as well as asymmetric non-separable models that, in addition to the interactions, account for the lack of spatio-temporal symmetry [6]–[8], [11]. By establishing an analogy between the atmospheric processes studied in the aforementioned literature and the wind field on a farm, one expects that the lack of symmetry is extended to the farm-level wind dynamics, and thus, anticipates a successful application of non-separable asymmetric models in short-term wind forecasting. Our review of the literature, however, has suggested the contrary – to the best of our knowledge, the physical phenomenon of spatio-temporal asymmetry has not been reported or made use of in the wind energy applications [12]–[14].

Triggered by that observation, our goal is to explore the

A. Ezzat and Y. Ding are with the Department of Industrial and Systems Engineering, Texas A&M University, College Station, TX, 77840 USA.

M. Jun is with the Department of Statistics, Texas A&M University, College Station, TX, 77840 USA.

Manuscript received XXXX, 2017; revised XXXX, 2017.

existence of this physical phenomenon in the context of a local wind field and the reasons behind its disregard in the related literature. We understand that the large-scale atmospheric processes referenced above are usually measured at a few, spatially distant sites over the span of several days. These measurements may be compatible with the spatial and temporal resolution in those large-scale processes but such setting is too coarse for understanding the wind dynamics on a farm. We are fortunate to have in possession a year worth of hourly wind data from 200 turbine-mounted anemometers scattered over a wind farm of roughly 15-by-10 miles. This anemometer network provides a coverage of spatial resolution of one mile and a temporal resolution of one hour, supposedly sufficient to allow finer-scale wind dynamics to be unearthed.

We show, however, that a straightforward use of this spatio-temporal dense data would lead to the conclusion that local wind fields are approximately symmetric; a conclusion that explains the widespread use of separable models. Our later findings, however, suggest that this conclusion is misleading because it overlooks the physical nature of local wind fields, where the local wind dynamics take place on a much more granular scale compared to large-scale processes. In order to see the fine-scale asymmetric patterns, unnoticed previously in the spatio-temporal data, we need some magnifying mechanisms that look deeper into the spatial and temporal scales, analogous to the use of optical lens for seeing tiny specimens and features transparent to naked eyes. The magnifying mechanisms are pertinent to the data in space and time, thereby referred to as a pair of "spatio-temporal" lens. These lens are apparently not a physical lens but realized instead through statistically-motivated computational algorithms. Using the lens, we conclude that, contrary to common practice, asymmetry does exist in the wind farm data.

In light of these findings and making use of the devised spatio-temporal lens, we propose an enhanced short-term wind forecasting procedure, which is based on an asymmetric non-separable model that takes into account the wind directionality information. Short-term wind forecasting is an important task in wind farm analytics [15] and is a corner stone for subsequent operations such as power estimations, turbine control and others [16]. The current literature for short-term forecast either uses simplistic methods such as persistence forecasts, or time series methods that do not account for spatial correlations such as autoregressive (AR) and autoregressive moving average (ARMA) models and support vector machines (SVMs) [17]–[19], or separable spatio-temporal models that take into account both the spatial and temporal aspects, but over-simplify the spatio-temporal structure by ignoring spatio-temporal interactions and asymmetries [14]. We show that significant improvements in forecast accuracy, in both wind speed and power, are achievable when an appropriate non-separable spatio-temporal model is adopted. When combined with certain intelligent methods such as support vector machine, additional improvements can be expected.

II. Data Description

The data used in this study consists of one year of spatio-temporal measurements, between 2010 and 2011, from an onshore wind farm in the United States. Due to confidentiality reasons, we randomly selected 200 turbines, noting that this represents a vast majority of the whole turbine set. The data consists of turbine-specific hourly wind speeds measured by the anenometers mounted on each turbine at the hub height of 80 meters. In addition, one year of hourly wind speed and direction measurements are available at a mast located in the northeastern part of the wind farm. We shift the coordinates of the turbines by a constant, so that the relative positions of the turbines and the mast remain faithful to the actual layout but their true geographic information is kept confidential; see Figure 1. The farm is located on a rather flat terrain.

The histogram of wind speeds recorded at the mast, as well as the rose plot that illustrates the wind speed and direction are presented in Figure 2. The yearly median and standard deviation of the wind speeds at the wind farm are $6.60\ ms^{-1}$ and $3.29\ ms^{-1}$, respectively. The most dominant wind occurring in the wind farm throughout the year is westerly, i.e., it blows from west to east. The average yearly wind direction is at $\bar{\theta} = 264.24°$.

At any given time, we can characterize the wind measured at a turbine by its wind direction and the wind speed. By considering measurements from multiple turbines, a vector of wind directions and speeds can be formed at each wind turbine and it evolves as a time series. It is in these time series vectors that the spatio-temporal correlation information is stored, which we need to decipher. Because wind has a transport effect due to its directionality, this spatio-temporal correlation is dependent on the wind direction, which could rapidly fluctuate in as short as 6 hours, as shown later in Section III. In other words, the spatio-temporal correlation in the along-wind direction is stronger than in other directions, creating the so-called "spatio-temporal asymmetry". Fortunately enough, the spatio-temporal resolution of our data allows us to characterize the fine-scale direction-dependent spatio-temporal correlations and interactions in local wind dynamics. Such characterization of the directional wind dynamics enables us to improve short-term wind speed forecasts, and subsequently, enhance turbine-specific power estimations. These practical implications are discussed in Section IV.

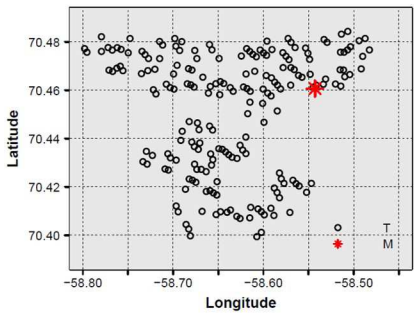


Fig. 1. Wind farm layout:"T" refers to turbines and "M" refers to the mast.

III. Asymmetry Detection & Quantification

An empirical method to quantify asymmetry is proposed in [6] in terms of spatio-temporal semi-variogram, which is a

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.2018.2789685, IEEE Transactions on Sustainable Energy

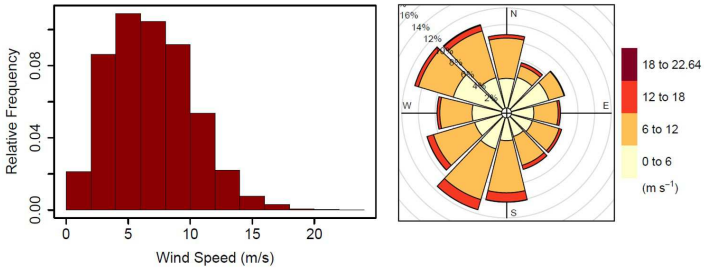


Fig. 2. Left panel: histogram of wind speeds as recorded at the mast. Right panel: wind rose plot as recorded at the mast.

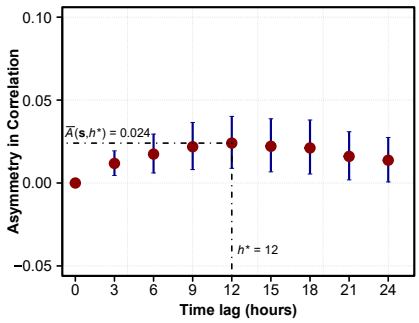


Fig. 3. The 25^{th}, 50^{th} and 75^{th} percentiles of asymmetry values $A(\mathbf{s},h)$ versus different time lags h.

negative-definite measure of dissimilarity in geostatistics [3]. Specifically, let $Y(\mathbf{s},t)$ represent the wind speeds at multiple locations $\{\mathbf{s}_i\}_{i=1}^n$ at discrete times $\{t_j\}_{j=1}^m$ such that $N = n \times m$, then the spatio-temporal empirical semi-variogram of Y between $\mathbf{s}_1$ and $\mathbf{s}_2$ at time lag h is defined as:

$$g(\mathbf{s}_1,\mathbf{s}_2,h) = \frac{1}{2(m-h-1)} \sum_{j=1}^{m-h-1} \{Y(\mathbf{s}_1,t_j+h) - Y(\mathbf{s}_2,t_j)\}^2. \quad (1)$$

Introduce two semi-variograms between $\mathbf{s}_1$ and $\mathbf{s}_2$: $g(\mathbf{s}_1,\mathbf{s}_2,h)$ and $g(\mathbf{s}_2,\mathbf{s}_1,h)$. Both of them represent the dissimilarity between the two spatial sites, but $g(\mathbf{s}_1,\mathbf{s}_2,h)$ means that measurements taken at $\mathbf{s}_2$ are h time lag behind that at $\mathbf{s}_1$, whereas $g(\mathbf{s}_2,\mathbf{s}_1,h)$ means that measurements at $\mathbf{s}_1$ are behind those at $\mathbf{s}_2$. Assuming Y is stationary in space-time with a positive-definite covariance function K, then $\mathbb{E}[g(\mathbf{s}_1,\mathbf{s}_2,h) - g(\mathbf{s}_2,\mathbf{s}_1,h)] = K(\mathbf{s}_2-\mathbf{s}_1,h) - K(\mathbf{s}_1-\mathbf{s}_2,h)$ [6]. As such, when the two semi-variogram quantities are the same, the wind field is said to be symmetric because their expected difference, which is equal to the expected difference in covariances, is zero. But when there is a dominant wind blowing from $\mathbf{s}_1$ towards $\mathbf{s}_2$, the propagation of wind velocities from $\mathbf{s}_1$ towards $\mathbf{s}_2$ would generate a significantly positive value for the difference $a(\mathbf{s}_1,\mathbf{s}_2,h) := g(\mathbf{s}_1,\mathbf{s}_2,h) - g(\mathbf{s}_2,\mathbf{s}_1,h)$, indicating a lack of symmetry. In a more general sense, when there is a dominant wind direction, denoted by θ (in the above example, from $\mathbf{s}_1$ towards $\mathbf{s}_2$), a non-zero $a(\mathbf{s}_1,\mathbf{s}_2,h)$ exhibits itself as a positive value most of the time. In our research, we signify the dominant wind direction through an extra input, i.e., $a(\mathbf{s}_1,\mathbf{s}_2,h,\theta)$, and use it to detect the existence of asymmetry in wind data and quantify its strength.

A. Yearly Scenario

We first show the conclusion that would result when attempting to quantify asymmetry in a local wind field by using the metric described above and following the procedure proposed in [6], which was originally designed for a large-scale atmospheric process.

To account for non-stationarity in $\mathbb{E}[Y(\mathbf{s},t)]$, we fit a parametric diurnal trend for the wind speed data using least squares estimation. Specifically, we use the diurnal trend function originally proposed by [20], which captures the global temporal variation in wind speeds. The fitted trend is then subtracted from the actual wind speed data to get the residuals, which will be used in quantifying asymmetry.

Next, for every pair of turbines i and j located at $\mathbf{s}_i$ and $\mathbf{s}_j$ such that $\mathbf{s}_i$ is west of $\mathbf{s}_j$, we compute $g(\mathbf{s}_i,\mathbf{s}_j,h) - g(\mathbf{s}_j,\mathbf{s}_i,h)$ using the residuals in place of Y in (1). We repeat this computation for every pair of turbines and for different time lags ranging from 0 to 24 hours. All of the computed quantities are then transformed into the correlation scale. For the ℓ-th pair of turbines, the resulting quantities at each temporal lag h are the spatio-temporal asymmetry, denoted by $a^\ell(\mathbf{s}_i,\mathbf{s}_j,h,\bar{\theta})$, where $\bar{\theta}$ is the average yearly wind direction at the wind farm. For our wind data, $\bar{\theta} = 264.24°$ (where $0°$ represents north), meaning that the average dominant wind on the wind farm is from west to east, as shown in the right panel of Figure 2.

Denote the collection of asymmetry values at each temporal lag by $A(\mathbf{s},h) = \{a^\ell(\mathbf{s}_i,\mathbf{s}_j,h)\}_{\ell=1}^L$, where L is the total number of turbine pairs, and its 50^{th} percentile as $\bar{A}(\mathbf{s},h)$. Figure 3 shows the 25^{th}, 50^{th} and 75^{th} percentiles of $A(\mathbf{s},h)$ for $h \in \{0,\ldots,24\}$ with a 3-hour increment. All median asymmetry values in Figure 3 are slightly positive, indicating a potential tendency towards asymmetric behavior. The largest median occurs at h^* = 12 and is approximately 0.024 on the correlation scale. To put this value in the context, we note that the work of [4] reports an asymmetry value of 0.12 for asymmetric large-scale wind flow over Ireland. Similar figures are reported in [7], ranging between 0.04 to 0.14 and averaged at 0.11. As such, an asymmetry of 0.024 appears to be rather weak to justify the existence of asymmetry in the local wind field. From a modeling perspective, one would understandably trade such weak asymmetry for computational efficiency and simplicity gained by the symmetry assumption. This conclusion explains the broad use of separable models in the wind application literature.

Our hypothesis, however, is that the weak asymmetry in Figure 3 is due to the non-optimal handling of wind farm data. In general, the common practice is to decompose data using regular calendar periods, like a week, a month, or a year. The wind data is grouped for the whole year when producing Figure 3. In a large-scale atmospheric process, a dominant wind can persist for a sustained period of time and travel a substantial distance; these patterns can be pre-identified through climatological expertise over a region of interest, and as such, regular calendar decompositions appear to be a reasonable choice. For a local wind field, however, observational data suggest that alternations in local winds occur at a relatively high rate, resulting in several distinct

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.2018.2789685, IEEE Transactions on Sustainable Energy

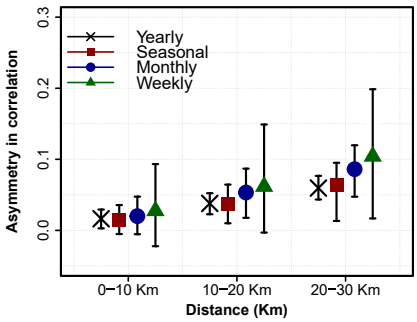


Fig. 4. The 25^{th}, 50^{th} and 75^{th} percentiles of asymmetry values of temporal decompositions versus separating distance in kilometers.

wind characteristics at each wind alternation. In such setting, regular calendar periods rarely contain a single dominant wind scenario; rather, they contain various dominant winds that create multiple asymmetries having distinct directions and magnitudes. Consequently, aggregating the heterogeneous, and perhaps opposite, asymmetries leads to an underestimation of the true asymmetry level.

The physical differences between local wind fields and large-scale atmospheric processes require special adjustments to spatio-temporal resolution used to analyze wind measurements, in order to reveal the underlying asymmetry pattern. We show next that upon implementing a pair of spatio-temporal "lens" that accounts for the fine-scale variations in local wind fields, strong degrees of asymmetry are detected.

B. Spatio-temporal lens for asymmetry quantification

We propose a spatio-temporal lens that allows us to unearth the underlying asymmetric behavior of local wind fields. The devised lens comprises of two components: a temporal adjustment and a spatial adjustment.

1) Temporal adjustment: We start off with performing a series of arbitrary, calendar period-based temporal decompositions. Specifically, we perform seasonal, monthly and weekly decompositions of the wind farm data and compute the asymmetry level in each sub-interval. Take the seasonal decomposition as an example, the average wind direction is computed for the fall season and is denoted by $\bar{\theta}_1$ and the asymmetry level between all pairs of turbines within that season is computed at the time lag h_1^* that maximizes the median asymmetry level for the fall season, resulting in the vector $\{a^\ell(\mathbf{s}_1, \mathbf{s}_2, h_1^*, \bar{\theta}_1)\}_{\ell=1}^{n_1}$. We repeat this process for the remaining three seasons, resulting in the wind directions $\bar{\mathbf{\Theta}} = \{\bar{\theta}_1, \bar{\theta}_2, \bar{\theta}_3, \bar{\theta}_4\}$, the optimal time lags $\mathbf{h}^* = \{h_1^*, h_2^*, h_3^*, h_4^*\}$ and the corresponding asymmetry vectors. We then group the asymmetry values into three subgroups, corresponding to three distance ranges: 0-10 kilometers (km), 10-20 km, and 20-30 km. The $25^{th}, 50^{th}$ and 75^{th} percentiles of asymmetry values in each subgroup are computed and displayed in Figure 4. We then repeat this process for all the temporal decompositions and display the $25^{th}, 50^{th}$ and 75^{th} percentiles of asymmetry values under each scenario for the three distance subgroups.

It is apparent that adjusting the temporal lens by which we search for the asymmetry to capture the fine temporal details leads to estimates of stronger asymmetries. However, the adjustment based purely on calendar periods seems not effective enough, because the decomposition intervals are created arbitrarily. If we can identify the time points when the dominant wind changes its direction, say, from northerly to westerly, we can then isolate the time intervals in which a unique dominant wind persists and consequently detect the underlying asymmetry in such intervals. We call such intervals as the "prevailing periods." This calls for the use of a change point detection procedure.

In order to identify the change points of the wind direction over the year, we implement a rolling binary segmentation of a circular change point detection algorithm, which is a modified version of the method described in [21]. Given hourly wind direction observations for one year, denoted by $\mathbf{\Theta} = \{\theta_1, \theta_2, ..., \theta_q\}$, where q is the total number of wind direction observations, we are interested in detecting the intermediate points at which changes have occurred in the wind direction. We assume that the wind direction variable θ follows a von Mises distribution with parameters $0 \leq \mu < 2\pi$ and $\kappa \geq 0$ as the mean and concentration parameters, respectively. The choice of the distributional form is motivated by the fact that wind direction is a circular variable and the von Mises distribution can characterize its directionality [22].

The parameter vector for the change detection test would be $\{\omega, \mu_1, \mu_2, \kappa\}$, where ω is the index of the change point to be detected, μ_1 and μ_2 are the means of the before and after subsequences and κ is the concentration parameter. Specifically, the null hypothesis for this change point test is that $H_0 : \omega = q$ indicating no change versus $H_1 : 1 \leq \omega \leq q-1$ indicating a change occurring at the ω^{th} observation. The generalized likelihood ratio method is used to conduct the test, where H_0 is rejected whenever $\lambda_\omega > c$ such that $\lambda_\omega = \sup_{j \in \{1,...,q\}} (Q_{1j} + Q_{2j}) - Q$ where Q, Q_{1j} and Q_{2j} denote the resultant lengths for the sequence and the resulting subsequences, respectively, and c is a critical cut-off value determined by the prescribed significance level; for more details, please refer to [21]. Because the test has no simple known distributional form, we resort to Monte Carlo simulation, as guided by [21] in order to determine the 95^{th} cut-off value which will be used in this test. In the Monte Carlo simulation, we initially tried a run size of 10,000 runs, but it turned out that the results are only trivially different from that obtained at a run size of 1,000 runs.

Since we are searching for an unknown number of potential change points, a common approach is to implement a binary segmentation algorithm, where the most significant change point in the whole dataset is detected, then the next significant change point is searched for in the resulting subsequences before and after the already-detected change point. For more details about binary segmentation, please refer to [23]. Since we are expecting change points to occur on a fine temporal resolution in local wind fields, we implement a one-month rolling binary segmentation, in which the first month of data is analyzed separately and then the period of interest is shifted by one month from the last change point detected in the first month for the next round of detection.

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.2018.2789685, IEEE Transactions on Sustainable Energy

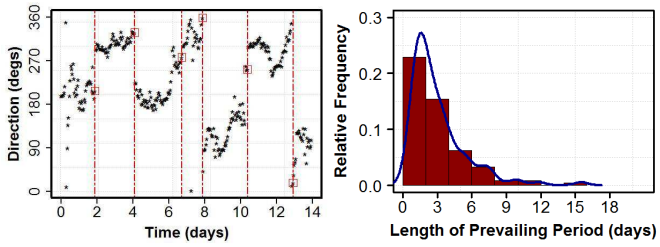
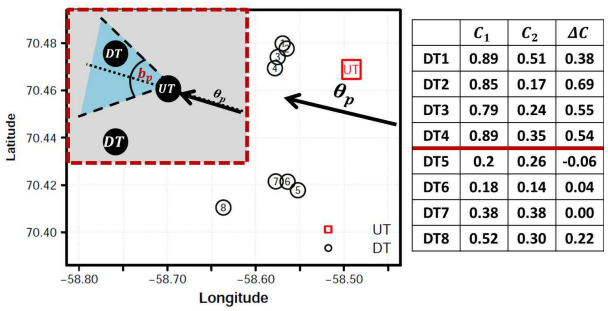


Fig. 5. Left panel: Change detection for first two weeks of wind direction data. Right panel: Distribution of the prevailing periods' length in days.



	C_1	C_2	ΔC
DT1	0.89	0.51	0.38
DT2	0.85	0.17	0.69
DT3	0.79	0.24	0.55
DT4	0.89	0.35	0.54
DT5	0.2	0.26	-0.06
DT6	0.18	0.14	0.04
DT7	0.38	0.38	0.00
DT8	0.52	0.30	0.22

Fig. 6. Effect of spatial position on asymmetry quantification. "UT" and "DT" denote an upstream and a downstream turbine, respectively.

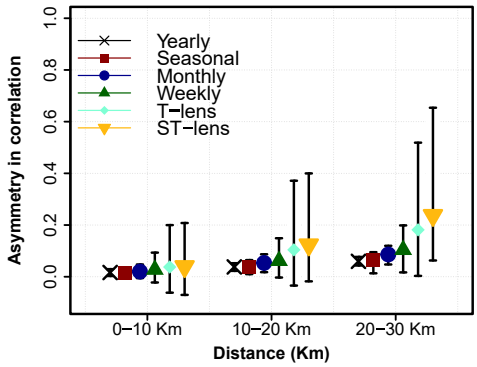


Fig. 7. The 25^{th}, 50^{th} and 75^{th} percentiles of asymmetry values of different scenarios versus separating distance in kilometers.

The left panel of Figure 5 shows the detected change points for the first two weeks of data. The right panel of Figure 5 shows the distribution of the length of the prevailing periods. The results show that, on average, a dominant wind direction lasts for 3.04 days with a standard deviation of 2.46 days. For 50% of the prevailing periods, the wind direction alternated in less than 2.27 days. The maximum interval of time in which a dominant wind direction is found to be persistent is 15.5 days, while the shortest prevailing period's length is found to be 6 hours. A total of 119 change points are detected in the year long wind data, leading to 120 prevailing periods identified over the year. These statistics indicate the fine temporal scale at which wind dynamics take place, resulting in a dynamic alternation of dominant winds over time. Quantifying asymmetry for individual prevailing periods would capture the lack of symmetry associated with each distinct dominant wind, and hence, provide reliable asymmetry estimates.

2) Spatial adjustment: On the spatial level, the relative position of the turbines on a wind farm is a factor that affects the asymmetry level at any given time. Physically, asymmetry exists when wind propagates from an upstream turbine to a downstream one, implying that the latter is in the along-wind direction with respect to the former. To show the effect of the spatial configuration of turbines on asymmetry, we pick an 8-hour prevailing period, where the wind is easterly, as illustrated in Figure 6. We compute the Pearson correlations between the wind speeds at the upstream turbine (UT) and those in a subset of 8 arbitrarily-selected downstream turbines (DTs 1-8) after a time lag of $h = 3$ hours; we denote this correlation by C_1. It is apparent that the correlations in the along-wind direction (between UT and DTs 1-4) is larger than those in the span-wind direction (between UT and DTs 5-8). Also, we compute the pairwise correlations at a time lag of $-h$ (when the DTs are leading the UT) and denote it by C_2. The difference between C_1 and C_2, denoted by ΔC, is largely positive for the along-wind DTs, indicating the existence of a strong asymmetry pattern. These observations prompt the need to select only the along-wind turbines in asymmetry quantification.

For a given prevailing period p, the dominant wind direction is denoted as θ_p and a distinct spatial bandwidth b_p will be selected. Our spatial bandwidth selection procedure, as illustrated in the upper left corner of Figure 6, executes as follows: we vary the bandwidth in the range $[2.5°, 45°]$ in increments of $2.5°$ and then select the bandwidth that maximizes the median asymmetry as the optimal bandwidth b_p^*. With the spatial adjustment, the asymmetry metric $a(\cdot)$ takes in one more input and is now denoted as $a(\mathbf{s}_1, \mathbf{s}_2, h_p, \theta_p, b_p)$.

3) Asymmetry quantification: In light of the spatio-temporal lens adjustments described above, we determine an optimal time lag h_p^* and bandwidth b_p^* that constitute our final pair of spatial and temporal lens, maximizing the median asymmetry level in each prevailing period $\bar{A}(\mathbf{s}_1, \mathbf{s}_2, h_p^*, \theta_p, b_p^*)$. Figure 7 shows the 25^{th}, 50^{th} and 75^{th} percentiles of the asymmetry level versus the separating distance subgroups for the different scenarios thus considered: yearly, seasonal, monthly, weekly, temporal-only lens scenario and spatio-temporal lens scenario. It is apparent that applying the spatio-temporal lens detects much higher asymmetry levels. For instance, at separating distances greater than 20 km, all of the turbine pairs exhibit positive asymmetry and 50% of them exhibit an asymmetry level higher than 0.2 on the correlation scale, a level considered significant in the past study [4] and nearly an order of magnitude greater than the median asymmetry of 0.024 detected earlier on the yearly data. The median asymmetry values of all distance subgroups are classified in Table I, where 93% of the prevailing periods exhibit positive median asymmetry, nearly a quarter of them exhibit a greater than 0.2 median asymmetry, and more than 41% of them exhibit a median asymmetry larger than 0.1, the level of asymmetry previously reported in [4], [7] for signaling the existence of appreciable asymmetric behavior in the large-scale atmospheric processes.

The findings made above indicate that not only does strong

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.2018.2789685, IEEE Transactions on Sustainable Energy

TABLE I
CLASSIFICATION OF PREVAILING PERIODS ACCORDING TO THE MEDIAN ASYMMETRY LEVEL.

Group	Range	Percentage
1.	$\bar{A}(\mathbf{s}_1, \mathbf{s}_2, h_p^*, \theta_p, b_p^*) \leq 0$	7%
2.	$0 < \bar{A}(\mathbf{s}_1, \mathbf{s}_2, h_p^*, \theta_p, b_p^*) < 0.05$	27%
3.	$0.05 \leq \bar{A}(\mathbf{s}_1, \mathbf{s}_2, h_p^*, \theta_p, b_p^*) < 0.1$	25%
4.	$0.1 \leq \bar{A}(\mathbf{s}_1, \mathbf{s}_2, h_p^*, \theta_p, b_p^*) < 0.2$	20%
5.	$0.2 \leq \bar{A}(\mathbf{s}_1, \mathbf{s}_2, h_p^*, \theta_p, b_p^*)$	21%

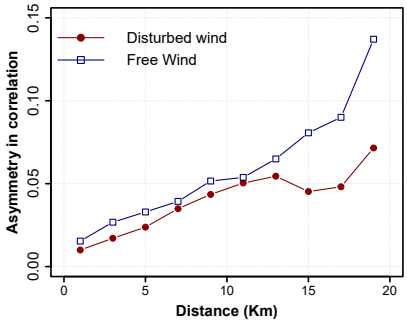


Fig. 8. Wake effect and its implications on spatio-temporal asymmetry.

asymmetry exist in local wind fields, but also the discovered asymmetry appears to fluctuate spatially and temporally, and in both magnitude and direction. Each prevailing period appears to have unique asymmetry pattern, creating a temporal fluctuation of asymmetry throughout the year. The asymmetry appears to exhibit a spatial variation as well, by taking on high values in the along-wind and low values in the span-wind directions. As such, separable spatio-temporal models appear overly-simplified in modeling farm-level local wind dynamics.

4) Asymmetry and wake effect: The implications of capturing the asymmetry in a local wind field can enrich our understanding of complex physical phenomena on a wind farm such as the wake effect. The spatio-temporal dynamics within a wind farm are affected by the wake effect because the rotating turbine blades cause changes in the speed, direction and turbulence intensity of the propagating wind [24]. In Figure 8, we divide the wind farm for each prevailing period based on the wind direction into two regions having approximately the same number of turbines. The first region is considered as the set of "free-stream" wind turbines that receive a relatively wake-free, less turbulent wind and the second region is considered as the set of "wake" wind turbines which are in the wake of other turbines and receive the disturbed, turbulent wind. We plot the medians of the asymmetry for each region. The free-stream region appears to exhibit higher asymmetry, which is consistent with the physical understanding since the less-turbulent wind is the driving force creating the asymmetry. This analysis indicates that asymmetry level spatially varies on a wind farm due to wake effect. Incorporating such pattern in a spatio-temporal model could benefit model fitting and prediction, as well as aid researchers in wake characterization.

IV. ENHANCED SHORT-TERM WIND FORECASTING BASED ON NON-SEPARABLE ASYMMETRIC MODELING

The modeling of asymmetry in local wind fields can have a vital impact on wind farm operations. The most obvious application is in wind forecasting. Accurate wind speed forecasting, for instance, can lead to tremendous improvements in power production estimation, which is pivotal to risk and cost reduction in grid integration [25]. Wind direction and speed forecasts can greatly benefit active turbine control and economic dispatch operations [26]. For short-term forecasts, like a couple of hours ahead, data-driven models are known to be more powerful than the physics-based numerical weather prediction models [14], [15]. In a spatio-temporal context, the performance of the data-driven forecasts hinges upon the quality of an adequate probabilistic model that can capture the spatio-temporal dynamics for it to perform well in the forward-kriging prediction mechanism [20], [27], [28].

In light of the findings of Section III, we devise a framework for short-term wind forecasting that takes into account the asymmetric nature of local wind dynamics. There is a rich body of literature on short-term wind forecast [15], [16], but we set ourselves apart in terms of the following three aspects:

- **Spatial correlation**: Our focus is to take advantage of the spatial correlation among the neighborhood turbines on a wind farm while making turbine-specific wind speed forecasts, as opposed to obtaining a single time-series prediction for one turbine or a single *aggregated* time series for the whole farm [29], [30].
- **Spatio-temporal resolution**: The within-farm local wind field provides us a spatially and temporally dense wind dataset, as opposed to the situations where measurements come from a small number of locations spread over large areas, as in [20], [25], [31], i.e., the between-farm forecast. As mentioned before, the spatial and temporal resolutions of our wind data are one mile and one hour, respectively, which are at a finer scale than those in the between-farm settings.
- **Lack of spatio-temporal symmetry**: As shown in Section III, local wind dynamics are strongly asymmetric. Our framework takes into account this lack of symmetry. To our best knowledge, there exists no methodology for short-term wind forecasting on a wind farm that considers this physical phenomenon.

The flowchart in Figure 9 presents the steps of the proposed forecasting procedure. We want to note that to perform h-hour ahead forecasts and following the terminology presented in Section III, only the data in the preceding prevailing period that share similar wind and asymmetry characteristics is used for model training. This implies that a small subset of data relevant to the current prevailing period is used for model training. Doing so, the benefit is two-fold. First, it eliminates the computational burden of fitting non-separable spatio-temporal models. Second, it makes use of a local informative spatio-temporal neighborhood that is most relevant to short-term horizons. We note that Pourhabib *et al.* [14] used informative spatial neighborhoods for short-term speed forecast and Yan

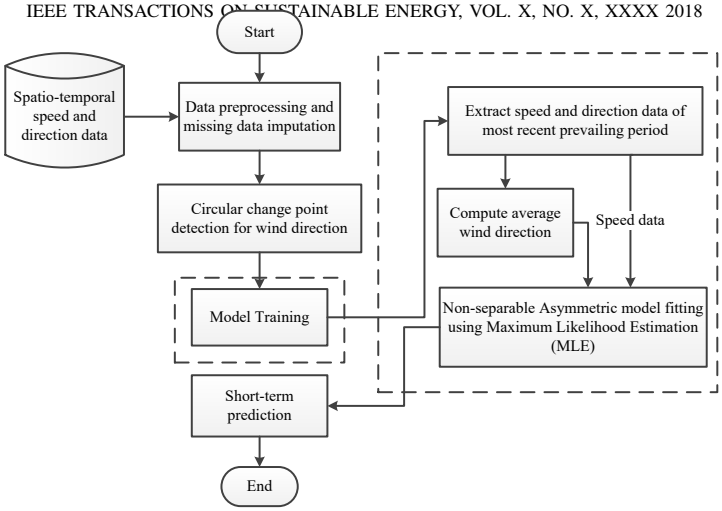


Fig. 9. Flowchart of short-term forecast using asymmetric modeling.

et al. [30] used local temporal windows for short-term time-series power predictions.

A. Models for short-term wind forecast

1) Asymmetric non-separable spatio-temporal model: In Figure 9, one major component is the fitting of an asymmetric non-separable spatio-temporal model. Such models have been proposed in the past, although not used for local wind fields. In this research, we use a modified version of the asymmetric non-separable spatio-temporal model proposed in [7]. We briefly explain the model for this paper to be self contained.

In general, a spatio-temporal model is defined as the sum of a mean term and a zero-mean normally distributed error $\boldsymbol{\epsilon}$ with covariance $\boldsymbol{\Sigma}$ such that $\mathbf{Y} = \mathbf{M}\mathbf{z} + \boldsymbol{\epsilon}$, where $\mathbf{Y}$ is the $N \times 1$ vector of wind speeds, $\mathbf{M}$ is a pre-specified design matrix and $\mathbf{z}$ is a vector of coefficients. The key issue here is to define $\boldsymbol{\Sigma}$ using a parametric form. For the model in [7], its parametric covariance function can be expressed as follows:

$$K_a(\mathbf{u}, h) = \sigma^2 \Big\{ (1-\lambda) K_{NS}(\mathbf{u}, h) + \lambda K_T(\mathbf{u}, h) \Big\} + \eta \mathbb{1}_{\{||\mathbf{u}||=|h|=0\}}, \tag{2}$$

where K_T is an asymmetric correlation function to be given below and K_{NS} is a non-separable symmetric correlation function such that:

$$K_{NS}(\mathbf{u}, h) = \frac{1-\delta}{1+\alpha|h|^2} \left(\exp\left[-\frac{c||\mathbf{u}||}{(1+\alpha|h|^2)^{\frac{\beta}{2}}} \right] + \frac{\delta}{1-\delta} \mathbb{1}_{\{||\mathbf{u}||=0\}} \right). \tag{3}$$

In (2) and (3), $\mathbb{1}_{\{\cdot\}}$ is an indicator function, $\mathbf{u} = (u_1, u_2)^T$ is the spatial lag consisting of longitudinal and latitudinal components u_1 and u_2 and $||\cdot||$ is the Euclidean norm. The parameters α and c determine the temporal and spatial ranges, whereas $0 \leq \delta < 1$ and η are the spatial and spatio-temporal nugget effects, respectively, such that α, c and η are all non-negative, and $\sigma^2 > 0$ is the spatio-temporal variance. The non-separability parameter β represents the strength of the spatio-temporal interaction and the asymmetry parameter λ represents the lack of symmetry; $\beta, \lambda \in [0, 1]$.

There are different approaches to define K_T. A simpler one is defined in [7] as a Lagrangian compactly supported function:

$$K_T(\mathbf{u}, h) = \left(1 - \frac{1}{2||\mathbf{V}||} ||\mathbf{u} - \mathbf{V}h|| \right)_+, \tag{4}$$

where $(\cdot)_+ = \max(\cdot, 0)$ and $\mathbf{V} = (V_1, V_2)^T$ is a two-dimensional vector having a longitudinal and latitudinal component, respectively, and is defined based on the knowledge of the weather system. For example, if the dominant wind is known to be strictly westerly, then $\mathbf{V}$ is chosen to be $(V_1, 0)^T$, i.e., a non-zero longitudinal wind velocity reflecting the traveling of the wind along the longitudinal axis. A generalized version of K_T is proposed in [32], which is, instead of using a constant vector, defines $\mathbf{V}$ as a random variable that follows a multivariate normal distribution, i.e., $\mathbf{V} \sim \mathcal{N}(\boldsymbol{\mu}, \frac{\mathbf{D}}{2})$. As such, K_T is defined as follows:

$$K_T(\mathbf{u}, h) = \frac{1}{\sqrt{|\mathbf{1}_{2\times 2} + h^2\mathbf{D}|}} \exp\big[-(\mathbf{u} - h\boldsymbol{\mu})^T (\mathbf{1}_{2\times 2} + h^2\mathbf{D})^{-1} (\mathbf{u} - h\boldsymbol{\mu}) \big], \tag{5}$$

where $|\cdot|$ in (5) denotes the matrix determinant. Our research shows that using the generalized K_T boosts the forecast quality, and for this reason, our forecast model uses the K_T in (5). This does not come as a surprise since local wind dynamics are highly random, and as such, $\mathbf{V}$ is best described as random rather than constant. We refer to this asymmetric non-separable model hereinafter as ASYM.

2) Seperable spatio-temporal models: By setting $\beta = \lambda = 0$ in ASYM, we get a symmetric separable model. In the results section, we use two variants of this separable model: SEP1 and SEP2. The model SEP1 borrows the estimated parameters of ASYM but sets $\beta = \lambda = 0$. As for SEP2, we first set $\beta = \lambda = 0$ before parameter estimation, and then freely estimate the remaining parameters from the data.

3) Prediction using spatio-temporal models: A method of prediction is to establish a functional relationship between the output, e.g. the wind speed at a future time, and the inputs, e.g. the wind speed measurements obtained in the past. The functional relationship is characterized by a model structure specified through the chosen spatio-temporal models. The model's parameters are estimated, or learned, using the historical training data. Specifically, for ASYM, SEP1, and SEP2, a Maximum Likelihood Estimation (MLE) is used to estimate the covariance parameters as well as the vector $\mathbf{z}$ in the mean structure.

Once all parameters are estimated (or learned), a future wind speed value can be calculated using the wind speed observations collected up to the current time; this calculated value is known as the prediction or forecast. The specific prediction mechanism used in this paper is called "kriging" and it is a standard approach to obtain predictions based on spatio-temporal models in the spatial statistics and geo-statistic literature [33]. The kriging mechanism computes the spatio-temporal predictions $\hat{Y}$ at a location $\mathbf{s}_o$ and a time t_o as a linear combination of the vector of observed spatio-temporal data points $\mathbf{Y}$, such that $\hat{Y}(\mathbf{s}_o, t_o) = \boldsymbol{\Phi}\mathbf{Y}$. The vector of the linear combination coefficients, $\boldsymbol{\Phi}$, is computed as in (6).

$$\mathbf{\Phi} = \{\mathbf{\Sigma}^{-1} - \mathbf{\Sigma}^{-1}\mathbf{M}(\mathbf{M}^T\mathbf{\Sigma}^{-1}\mathbf{M})^{-1}\mathbf{M}^T\mathbf{\Sigma}^{-1}\}\mathbf{k} + \mathbf{\Sigma}^{-1}\mathbf{M}(\mathbf{M}^T\mathbf{\Sigma}^{-1}\mathbf{M})^{-1}m_o, \quad (6)$$

where $\mathbf{\Sigma}$ is the covariance matrix that is computed depending on the model choices mentioned above, be it ASYM, SEP1 or SEP2, the vector $\mathbf{k}$ contains the covariances between $\mathbf{Y}$ and $Y(\mathbf{s}_o, t_o)$. For a constant mean spatio-temporal model, as used in this research, $\mathbf{M}$ reduces to $\mathbf{1}_{N\times 1}$ and $m_o = 1$.

4) Persistence model: Forecasts based on persistence models are widely used in the renewable industry and are simply obtained by extending the current wind charactersitics for the look-ahead horizons such that $\hat{Y}_{i,(j+h)} = Y_{ij}, \ \forall h$. We refer to this model as PER.

5) Autoregressive moving average (ARMA) model: We would like to compare ASYM against a classical time series method which does not account for spatial correlation on the wind farm (and thus no spatio-temporal asymmetry). Please note that spatio-temporal models (ASYM, SEP1 and SEP2) can be looked at as generalizations of time-series models, since they account for both spatial and temporal correlations.

A common approach for time-series modeling and forecasting is the autoregressive moving average (ARMA) model, which is defined by the orders of the autoregressive and the moving average parts, p and q respectively, and a set of parameters $\{a, b_1, .., b_p, c_1, ..., c_q\}$. An ARMA$(p, q)$ model for a spatial location i is expressed as in Equation (7).

$$Y_i(t) = a + \sum_{l=1}^{p} b_l Y_i(t-l) + \sum_{l=1}^{q} c_l \epsilon(t-l) + \epsilon(t), \quad (7)$$

where $\epsilon \sim N(0, \sigma^2_{ts})$. We fit an ARMA model using the command `arima` in the `R` package `stats`. Our analysis suggests that a better prediction is achieved by using low-order, rather than high-order ARMA models. This aligns with the recommendations of [17] and [14] who suggest that low-order ARMA models are more suitable for short-term forecasts. Therefore, we set $p = q = 1$ and fit an ARMA$(1, 1)$ model for each turbine out of the 200 locations at the wind farm. As such, we generate 200 independent forecasts and then compute the overall aggregate forecast error.

6) Support vector machines (SVMs): Machine learning approaches have been reported to benefit wind forecasting. Amongst those, support vector machines (SVMs) are a popular choice and have been previously implemented for time-series wind speed and power forecasting [19], [34], [35]. Simply speaking, SVM performs a nonlinear mapping of the data into a high-dimensional feature space, as shown in Equation (8):

$$Y_i(t) = \sum_{j=1}^{d} \beta_j \phi_j(x_i(t)) + \beta_0, \quad (8)$$

where $x_i(t)$ is an explanatory input to be defined later, $\{\phi_j\}_{j=1}^d$ are the so-called feature basis functions and $\{\beta_j\}_{j=0}^d$ are the parameters to be estimated from the data [19]. For an SVM model, instead of explicitly defining the feature basis functions, a kernel function is specified that defines the inner product in the feature space. This is referred to in the machine learning literature as the "kernel trick" [36]. Following the same logic made above regarding the suitability of most recent lagged values with respect to short-term forecasts, we set $x_i(t) = Y_i(t-1)$. We fit an SVM model for the time series training data of each turbine of the 200 locations at the wind farm. Specifically, we use the command `svm` in the `R` package `e1071` and a radial basis kernel function, which is a common choice in the SVM literature [36]. We refer to this model as SVM.

7) Hybrid forecast of ASYM & SVM: A growing trend in the wind forecasting literature is to hybridize multiple data-driven methods to achieve better prediction accuracy [37], [38]. We follow an approach similar to [37], which was originally proposed for integrating classical time-series models with neural networks for wind speed forecasting. Specifically, we first fit ASYM to the spatio-temporal training data. The choice of ASYM as the base model is motivated by our goal to capture the spatio-temporal dynamics and asymmetries on the wind farm. We then fit an SVM to the residuals obtained by ASYM to capture any non-linearities that are not covered by the base model. The final hybrid model has an additive form as in Equation (9):

$$Y_i(\mathbf{s}, t) = Y_i^A(\mathbf{s}, t) + \mathcal{E}_i^S(\mathbf{s}, t) + e_i, \quad (9)$$

where $Y_i^A(\mathbf{s}, t)$ is the ASYM model fit, $\mathcal{E}_i^A(\mathbf{s}, t)$ represents the SVM model fit to the spatio-temporal residuals after the ASYM fit, and e_i is the final residual term. We refer to the hybrid model in Equation (9) as HYB.

B. Case study

We demonstrate the merit of our procedure on four prevailing periods from different times of the year. For all the periods, we select 6 hours for model training. We will forecast up to 4-hours ahead, i.e., $h = \{1, 2, 3, 4\}$. Our choice for the training period is motivated by observing that the shortest prevailing period length, as shown in Section III, is about 6 hours. As such, a training period of 6 hours more or less guarantees temporal homogeneity and stationarity over the training data, allowing for a reliable model estimation. Furthermore, for short-term wind prediction, previous studies [14], [39], [40] showed that using a longer history of wind measurements is not necessarily helpful, as evident by the low time lag order used in the time series models. For the testing horizon, our collaborating industry partner (who is also the data provider) is interested in 2-hour ahead forecasts for various purposes. At time t, it is preferable to have the forecast at $t+1$ and $t+2$, so that when it is $t+1$, the old forecast at $t+2$ can still be used as the 1-hour ahead forecast, while the model starts running an update with a new forecast for the next two hours. This updating process is conducted in a rolling forward fashion on an hourly basis. For research purposes, we do not stop at the 2-hour ahead, but rather evaluate the forecast accuracy for longer horizons (up to 4 hours), albeit not needed by our industry collaborator.

As mentioned previously, in ASYM, $\mathbf{V} \sim \mathcal{N}(\boldsymbol{\mu}, \frac{\mathbf{D}}{2})$. To specify $\boldsymbol{\mu}$ and $\mathbf{D}$, we make use of the average wind speed $\bar{Y}$ and average wind direction $\bar{\theta}$ of the training data as follows. First, we set $||\boldsymbol{\mu}|| = \bar{Y}$. Then, given $\bar{\theta}$, we compute the wind velocity vector in degrees, by decomposing $||\boldsymbol{\mu}||$ into longitudinal and

TABLE II
LOG-LIKELIHOODS OF THE ASYMMETRIC VERSUS THE SEPARABLE MODELS. BOLD-FACED VALUES INDICATE THE BEST PERFORMANCE.

Period	Month	ASYM	SEP1	SEP2
1.	October, 2010	**-2087.84**	-2091.69	-2089.19
2.	December, 2010	**-1980.29**	-2263.43	-2030.04
3.	January, 2011	**-1796.45**	-1826.90	-1799.82
4.	June, 2011	**-2181.68**	-2463.77	-2185.45

latitudinal components, denoted by $\hat{v}_1$ and $\hat{v}_2$, respectively. These components will be used as the estimate for $\boldsymbol{\mu}$ and their signs would represent the wind directionality information. In our current study, we found that when the 2×2 matrix $\mathbf{D}$ is chosen as a diagonal matrix with its diagonal entries set to $||\boldsymbol{\mu}||$, the parameter estimation procedure produces higher MLE values, suggesting a better fit. This revelation surprises us a bit as the traditional wisdom would estimate $\mathbf{D}$ from the training data. One of our ongoing pursuits is to understand and then devise the optimal way for setting these parameters in the non-separable models.

The MLE is implemented in `R` using the routine `nlm`. For instance, for a prevailing period in January, 2011, $\bar{\theta} = 122.28°$, $\bar{Y} = 20.15$ km/hr, $(\hat{v}_1, \hat{v}_2) = (-0.20°, 0.10°)$ and the MLE's are $\{\hat{\sigma}^2, \hat{\nu}, \hat{\alpha}, \hat{c}, \hat{\eta}, \hat{\beta}, \hat{\lambda}\}$ = $\{19.22, 0.18, 0.46, 0.35, 3.35, 0.99, 0.47\}$. The estimate for the mean term is 21.98 km/hr. The values for $\hat{\beta}$ and $\hat{\lambda}$ indicate the need to consider spatio-temporal interaction and asymmetry.

The log-likelihoods for ASYM, SEP1 and SEP2 are presented in Table II. The log-likelihood values for ASYM are higher than those of SEP1 and SEP2 for all the periods, suggesting a higher explanatory power in favor of the asymmetric model and the importance of incorporating this physical phenomenon in spatio-temporal modeling.

Next, we make forecasts based on the models presented in Section IV-A and evaluate them in terms of Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). The results are illustrated in Table III. Please note that for Period 2, the testing data for $h = 4$ had a large amount of missing values when our industrial collaborator provided the data, and for this reason, we were not able to assess the forecast quality for that particular forecast horizon. We, therefore, compute the aggregate error based on the first 3-hour ahead forecasts for Period 2. For Periods 1, 3 and 4, the aggregate measure reported is the average over all 4-hour ahead forecasts.

It is also known that an important goal for power engineers is to accurately predict power production, in addition to wind speed forecasts. In practice, turbine-specific power curves are provided by the manufacturer and are used to assess the prediction accuracy of competing models. Since we do not readily have these power curves at hand, we use the binning method over the available one-year worth of data to estimate the turbine-specific power curves. The binning method is a nonparametric method commonly used in the wind industry [41] and is based on discretizing the wind speed domain into a number of bins and then outputting the average power value for each bin as an estimate of power generation. Using the estimated power curves, we predict the power generated at each turbine given the wind speed forecasts. We compare the competing models in terms of RMSE in Table IV. Please note that for confidentiality reasons, we have normalized the power output in the range of $[0, 1]$, so that the values reported in Table IV are the ratio relative to the maximum turbine power.

The results presented in Tables III and IV show that the forecasts based on asymmetric non-separable model outperform those based on separable model (SEP1 and SEP2), persistence model (PER), time-series model (ARMA), and machine learning model (SVM), in terms of both wind speed and wind power forecasting. The improvement of ASYM over the separable models is explained by capturing the strong asymmetries discovered in Section III, whereas the improvement over ARMA and SVM is mostly due to the characterization of spatial correlations, and subsequently the asymmetry, both of which the time-series models fail to capture.

Furthermore, hybridizing ASYM with SVM (the HYB model) appears to achieve a further enhancement in prediction accuracy over the ASYM only approach, demonstrating the additional benefit brought by the machine learning method. The improvements of HYB over ASYM range from 1% to 7%, and on average 3.5%, for wind speed forecast, and up to 8%, but on average 4.0%, for wind power forecast.

Combining the strength of the asymmetrical modeling and machine learning, in terms of wind speed forecast, HYB improves, on the average of the four periods, 22.5% in RMSE (24.3% in MAE, same below) over SEP1, 8.0% (9.7%) over SEP2, 19.5% (19.0%) over PER, 21.2% (24.5%) over ARMA(1,1) and 22.9% (25.1%) over SVM. In terms of wind power forecast, HYB on average improves in terms of reduction in RMSE 23.5% over SEP1, 9.4% over SEP2, 18.2% over PER, 22.8% over ARMA(1,1), 24.3% over SVM. These results are aligned with the findings made in Section III that local wind fields are strongly asymmetric at the fine-scale spatio-temporal resolutions and as such, spatio-temporal models that capture such physical phenomenon are expected to perform well in short-term forecasts.

V. CONCLUSION

Spatio-temporal wind farm analytics can provide insights to minimize uncertainty while using the wind resource. We demonstrate that, contrary to common practice, local wind fields are strongly asymmetric. This asymmetry is detected upon implementing a set of spatio-temporal adjustments that unearth the fine-scale spatio-temporal dynamics. The asymmetry pattern changes spatially and temporally, in both magnitude and direction. As such, the traditional separable spatio-temporal models appear to be overly-simplified in modeling farm-level wind dynamics. Given these findings, an enhanced procedure for short-term wind forecasting was devised and shown to outperform the commonly used forecast methods based on persistence, time-series, machine learning, and separable spatio-temporal models.

ACKNOWLEDGMENT

This research was partially supported by NSF grants no. CMMI-1300560, IIS-1741173, and DMS-1613003.

TABLE III
RMSE AND MAE OF WIND SPEED FORECASTS. MISSED THE DATA FOR $h = 4$ AND PERIOD 2 IN THE ORIGINAL DATASET. BOLD-FACED VALUES INDICATE THE BEST PERFORMANCE. THE PERCENTAGE IMPROVEMENTS ARE THE ERROR INFLATION RATE RELATIVE TO HYB.

		RMSE						MAE					
Period	Forecast	$h=1$	$h=2$	$h=3$	$h=4$	Average	% Imp.	$h=1$	$h=2$	$h=3$	$h=4$	Average	% Imp.
1	ASYM	**0.95**	1.37	2.70	2.86	2.14	1%	0.80	1.18	2.48	2.64	1.77	1%
	SEP1	1.23	1.63	2.92	3.03	2.34	9%	1.06	1.45	2.72	2.81	2.01	12%
	SEP2	1.01	1.50	2.85	2.99	2.26	6%	0.86	1.31	2.63	2.77	1.89	7%
	PER	1.29	1.72	2.98	3.16	2.42	12%	1.05	1.49	2.80	2.90	2.06	15%
	ARMA(1,1)	1.63	2.06	3.48	3.62	2.83	25%	1.31	1.83	3.29	3.35	2.47	29%
	SVM	1.61	1.91	3.33	3.44	2.70	21%	1.40	1.69	3.14	3.18	2.35	25%
	HYB	**0.95**	**1.36**	**2.69**	**2.85**	**2.13**		**0.79**	**1.17**	**2.46**	**2.62**	**1.76**	
2	ASYM	1.42	2.53	2.31	–	2.14	4%	1.04	2.28	2.09	–	1.80	5%
	SEP1	2.77	3.25	2.65	–	2.90	29%	2.19	2.80	2.30	–	2.43	30%
	SEP2	1.64	2.62	2.30	–	2.22	7%	1.29	2.36	2.06	–	1.90	10%
	PER	1.83	2.88	2.57	–	2.47	17%	1.49	2.55	2.27	–	2.10	19%
	ARMA(1,1)	1.99	3.05	2.78	–	2.64	22%	1.62	2.81	2.52	–	2.31	26%
	SVM	2.42	3.68	3.41	–	3.21	36%	2.19	3.41	3.10	–	2.87	39%
	HYB	**1.25**	**2.45**	**2.27**	–	**2.06**		**0.89**	**2.19**	**2.04**	–	**1.71**	
3	ASYM	**0.87**	0.94	1.02	1.35	1.06	2%	**0.72**	0.77	0.85	1.17	0.88	2%
	SEP1	1.21	1.25	1.21	1.52	1.31	20%	0.99	1.04	1.01	1.30	1.08	21%
	SEP2	0.88	1.09	1.14	1.50	1.17	11%	0.73	0.92	0.97	1.31	0.98	13%
	PER	1.01	1.07	1.36	1.51	1.25	17%	0.81	0.84	1.05	1.20	0.98	13%
	ARMA(1,1)	1.11	1.32	1.30	1.65	1.36	24%	0.93	1.15	1.14	1.43	1.16	26%
	SVM	1.03	1.16	1.34	1.68	1.33	21%	0.84	0.94	1.07	1.40	1.06	19%
	HYB	0.89	**0.91**	**1.00**	**1.31**	**1.04**		**0.72**	**0.74**	**0.83**	**1.13**	**0.85**	
4	ASYM	1.27	1.45	**1.87**	3.94	2.37	7%	1.00	1.20	**1.48**	3.73	1.85	6%
	SEP1	3.38	2.45	2.66	4.19	3.24	32%	2.68	1.99	2.13	3.75	2.64	34%
	SEP2	1.41	1.59	1.99	3.80	2.40	8%	1.13	1.33	1.61	3.60	1.91	9%
	PER	1.88	2.10	2.53	5.28	3.25	32%	1.49	1.71	2.06	4.78	2.51	31%
	ARMA(1,1)	2.07	1.77	2.14	3.81	2.58	14%	1.67	1.44	1.76	3.51	2.09	17%
	SVM	1.71	1.74	2.15	3.89	2.54	13%	1.43	1.44	1.71	3.60	2.05	15%
	HYB	**1.26**	**1.43**	1.92	**3.50**	**2.21**		**0.99**	**1.18**	1.54	**3.26**	**1.74**	

TABLE IV
RMSE OF WIND POWER PREDICTIONS. BOLD-FACED VALUES INDICATE THE BEST PERFORMANCE. THE PERCENTAGE IMPROVEMENTS ARE THE ERROR INFLATION RATE RELATIVE TO HYB.

P	h	ASYM	SEP1	SEP2	PER	ARMA	SVM	HYB
1	1	**0.10**	0.13	0.11	0.13	0.17	0.16	**0.10**
	2	**0.17**	0.21	0.19	0.20	0.24	0.23	**0.17**
	3	0.37	0.39	0.38	0.38	0.44	0.43	**0.36**
	4	**0.40**	0.42	0.42	0.42	0.48	0.45	**0.40**
	Av.	0.29	0.31	0.30	0.31	0.36	0.34	**0.28**
	%	1%	7%	5%	6%	19%	16%	
2	1	0.13	0.31	0.17	0.19	0.22	0.29	**0.12**
	2	0.25	0.36	0.27	0.30	0.33	0.43	**0.24**
	3	**0.24**	0.30	0.25	0.28	0.32	0.41	**0.24**
	4	–	–	–	–	–	–	–
	Av.	0.22	0.33	0.23	0.26	0.29	0.38	**0.20**
	%	4%	36%	11%	21%	29%	45%	
3	1	**0.10**	0.14	0.11	0.12	0.13	0.13	**0.10**
	2	0.12	0.15	0.14	0.12	0.16	0.14	**0.11**
	3	0.12	0.14	0.14	0.14	0.16	0.15	**0.11**
	4	0.17	0.19	0.19	0.18	0.21	0.20	**0.16**
	Av.	0.13	0.16	0.15	0.14	0.17	0.16	**0.12**
	%	3%	17%	13%	9%	22%	18%	
4	1	**0.17**	0.34	0.18	0.25	0.27	0.22	**0.17**
	2	0.19	0.31	0.20	0.28	0.23	0.23	**0.18**
	3	**0.24**	0.33	0.26	0.33	0.28	0.28	**0.24**
	4	0.42	0.46	0.39	0.62	0.41	0.42	**0.35**
	Av.	0.25	0.36	0.26	0.38	0.30	0.29	**0.23**
	%	8%	34%	9%	37%	20%	17%	

REFERENCES

[1] US Department of Energy, Washington D.C., "Wind Vision: A new era for wind power in the United States," 2015.

[2] C. E. Rasmussen and C. K. Williams, *Gaussian Processes for Machine Learning*. MIT press Cambridge, 2006.

[3] N. Cressie and C. Wikle, *Statistics for Spatio-Temporal Data*. John Wiley & Sons, 2011.

[4] T. Gneiting, "Nonseparable, stationary covariance functions for space-time data," *Journal of the American Statistical Association*, vol. 97, no. 458, pp. 590–600, 2002.

[5] Y. Wan, M. Milligan, and B. Parsons, "Output power correlation between adjacent wind power plants," *Journal of Solar Energy Engineering*, vol. 125, pp. 551–555, 2003.

[6] M. Stein, "Space-time covariance functions," *Journal of the American Statistical Association*, vol. 100, no. 469, pp. 310–321, 2005.

[7] T. Gneiting, M. Genton, and P. Guttorp, "Geostatistical space-time models, stationarity, separability and full symmetry," in *Statistical Methods for Spatio-Temporal Systems*, B. Finkenstadt, L. Held, and V. Isham, Eds. Chapman & Hall/CRC, 2007, ch. 4.

[8] M. Jun and M. Stein, "An approach to producing space-time covariance functions on spheres," *Technometrics*, vol. 49, no. 4, pp. 468–479, 2007.

[9] N. Cressie, S. Burden, W. Davis, P. Krivitsky, P. Mokhtarian, T. Suesse, and A. Zammit-Mangion, "Capturing multivariate spatial dependence: model, estimate and then predict," *Statistical Science*, vol. 30, no. 2, pp. 170–175, 2015.

[10] N. Cressie and H. Huang, "Classes of nonseparable, spatio-temporal stationary covariance functions," *Journal of the American Statistical Association*, vol. 94, no. 448, pp. 1330–1340, 1999.

[11] M. Park and M. Fuentes, "New classes of asymmetric spatial-temporal covariance models," North Carolina State University, Department of Statistics, Tech. Rep. 2584, 2006.

[12] A. Kusiak and W. Li, "Estimation of wind speed: A data-driven approach," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 98, pp. 559–567, 2010.

[13] H. Miao, L. Yang, J. Zhang, and V. Vittal, "A spatio-temporal analysis approach for short-term forecast of wind farm generation," *IEEE Transactions on Power Systems*, vol. 29, no. 4, pp. 1611–1622, 2014.
[14] A. Pourhabib, J. Huang, and Y. Ding, "Short-term wind speed forecast using measurements from multiple turbines in a wind farm," *Technometrics*, vol. 58, no. 1, pp. 138–147, 2016.
[15] X. Zhu and M. G. Genton, "Short-term wind speed forecasting for power system operations," *International Statistical Review*, vol. 80, no. 1, pp. 2–23, 2012.
[16] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl, "The state-of-the-art in short-term prediction of wind power: A literature overview, 2nd edition," ANEMOS.plus, Tech. Rep., 2011.
[17] Z. Huang and Z. Chalabi, "Use of time-series analysis to model and forecast wind speed," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 56, no. 2, pp. 311 – 322, 1995.
[18] J. Torres, A. Garca, M. D. Blas, and A. D. Francisco, "Forecast of hourly average wind speed with arma models in navarre (spain)," *Solar Energy*, vol. 79, no. 1, pp. 65 – 77, 2005.
[19] M. A. Mohandes, T. O. Halawani, S. Rehman, and A. A. Hussain, "Support vector machines for wind speed prediction," *Renewable Energy*, vol. 29, no. 6, pp. 939–947, 2004.
[20] T. Gneiting, K. Larson, K. Westrick, M. Genton, and E. Aldrich, "Calibrated probabilistic forecasting at the stateline wind energy center," *Journal of the American Statistical Association*, vol. 101, no. 475, pp. 968–979, 2006.
[21] S. Jammalamadaka and A. SenGupta, *Topics in Circular Statistics*. World scientific, 2001.
[22] K. V. Mardia, *Statistics for Directional Data*. Academic Press, 1972.
[23] P. Fryzlewicz, "Wild binary segmentation for multiple change-point detection," *The Annals of Statistics*, vol. 42, no. 6, pp. 2243–2281, 2014.
[24] A. Crespo, J. Hernndez, and S. Frandsen, "Survey of modelling methods for wind turbine wakes and wind farms," *Wind Energy*, vol. 2, no. 1, pp. 1–24, 1999.
[25] A. Hering and M. Genton, "Powering up with space-time wind forecasting," *Journal of the American Statistical Association*, vol. 105, pp. 92–104, 2010.
[26] L. Xie, Y. Gu, X. Zhu, and M. Genton, "Short-term spatio-temporal wind power forecast in robust look-ahead power system dispatch," *IEEE Transactions on Smart Grids*, vol. 5, no. 1, pp. 511–520, 2014.
[27] P. Pinson, "Wind energy: forecasting challenges for its operational management," *Statistical Science*, vol. 28, no. 4, pp. 564–585, 2013.
[28] T. Gneiting and M. Katzfuss, "Probabilistic forecasting," *The Annual Review of Statistics and its Application*, vol. 1, pp. 125–151, 2014.
[29] P. Ailliot and V. Monbet, "Markov-switching autoregressive models for wind time series," *Environmental Modelling & Software*, vol. 30, pp. 92–101, 2012.
[30] J. Yan, K. Li, E. Bai, J. Deng, and A. Foley, "Hybrid probabilistic wind power forecasting using temporally local gaussian process," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 1, pp. 87–95, 2016.
[31] J. Tatsu, P. Pinson, P. Trombe, and H. Madsen, "Probabilistic forecasts of wind power generation accounting for geographically dispersed information," *IEEE Transactions on Smart Grid*, vol. 5, no. 1, pp. 480–489, 2014.
[32] M. Schlater, "Some covariance models based on normal scale mixtures," *Bernoulli*, vol. 16, no. 3, pp. 780–797, 2010.
[33] N. Cressie, "The origins of kriging," *Mathematical Geology*, vol. 22, no. 3, pp. 239–252, 1990.
[34] L. Yang, M. He, J. Zhang, and V. Vittal, "Support-vector-machine-enhanced markov model for short-term wind power forecast," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 3, pp. 791–799, 2015.
[35] G. Santamaría-Bonfil, A. Reyes-Ballesteros, and C. Gershenson, "Wind speed forecasting for wind farms: A method based on support vector regression," *Renewable Energy*, vol. 85, pp. 790–809, 2016.
[36] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning*. Springer series in statistics New York, 2001, vol. 1.
[37] E. Cadenas and W. Rivera, "Wind speed forecasting in three different regions of mexico, using a hybrid arimaann model," *Renewable Energy*, vol. 35, no. 12, pp. 2732 – 2738, 2010.
[38] J. Hu, J. Wang, and G. Zeng, "A hybrid forecasting approach applied to wind speed time series," *Renewable Energy*, vol. 60, no. Supplement C, pp. 185 – 194, 2013.
[39] B. G. Brown, R. W. Katz, , and A. H. Murphy, "Time series models to simulate and forecast wind speed and wind power," *Journal of Climate and Applied Meteorology*, vol. 23, pp. 1184–1195, 1984.
[40] E. Erdem and J. Shi, "ARMA based approaches for forecasting the tuple of wind speed and direction," *Applied Energy*, vol. 88, no. 4, pp. 1405–1414, 2011.
[41] IEC-International Electrotechnical Commission, IEC-61400-12, Geneva, Switzerland, "Wind turbines-part 12-1: Power performance measurements of electricity producing wind turbines," 2005.

Ahmed Aziz Ezzat received the B.Sc. degree in Industrial & Management Engineering from the Arab Academy of Science & Technology in Alexandria, Egypt in 2013. He is currently pursuing his Ph.D. degree in Industrial and Systems Engineering at Texas A&M University. His research focuses on the spatio-temporal statistical modeling and control in engineering systems with particular emphasis on wind energy applications.

Mikyoung Jun received her Ph.D. in Statistics from University of Chicago (2005). She is currently an Associate Professor of Statistics at Texas A&M University. One of her main research areas is spatio-temporal modeling for environmental and climate applications, especially development of parametric non-stationary and non-separable covariance functions for univariate, as well as multivariate processes.

Yu Ding (M'01, SM'11) received B.S. from University of Science & Technology of China (1993); M.S. from Tsinghua University, China (1996); M.S. from Penn State University (1998); received Ph.D. in Mechanical Engineering from University of Michigan (2001). He is currently the Mike and Sugar Barnes Professor of Industrial & Systems Engineering and a Professor of Electrical & Computer Engineering at Texas A&M University. His research interests are in system informatics and quality and reliability engineering. Dr. Ding is a fellow of IIE, a fellow of ASME, a senior member of IEEE, and a member of INFORMS.