An Integro-Difference Equation Model for Spatio-Temporal Offshore Wind Forecasting

Feng Ye
Department of Industrial & Systems Engineering
Rutgers University
Piscataway, USA
feng.ye@rutgers.edu

Ahmed Aziz Ezzat†
Department of Industrial & Systems Engineering
Rutgers University
Piscataway, USA
aziz.ezzat@rutgers.edu

***Abstract*—Accurate short-term wind forecasts are instrumental to the optimal operation and management of offshore wind farms. While there have been significant advancements in data-science-based modeling for wind forecasting applications, little attention has been devoted to approaches that adopt the integro-difference equation (IDE) framework, wherein the time evolution of a stochastic spatial process (e.g., the hub-height wind field) is conditioned on its past history through the specification of an appropriate redistribution kernel. We propose a hierarchical model, rooted in IDE, wherein key kernel hyperparameters that bear physical relevance to the wind field dynamics are modeled as the output of a latent spatio-temporal regression model, thereby allowing them to change over time in line with the dynamic nature of local weather patterns. Using hub-height observations from the offshore wind energy areas in the U.S. Mid-Atlantic where several large-scale projects are in-development, we demonstrate that the proposed approach can result in considerable improvements, in terms of wind resource and power forecasting accuracy, relative to a representative set of time series and spatio-temporal methods that are prevalent in short-term wind forecasting applications.**

***Index Terms*—Integro-Difference Equation, Offshore Wind, Spatio-temporal Forecasting, Wind Energy.**

I. INTRODUCTION

The reliable operation of offshore wind farms largely hinges on accurate short-term forecasts of the offshore wind resource and generation issued at forecast horizons ranging from few minutes up to several hours ahead [1]. Those forecasts are regularly used by grid operators and energy producers to inform several of their short-term operations, including economic dispatch and production scheduling [2], [3], asset operations and management [4], [5], and control [6].

Data-science-based forecasting models—often referred to collectively as data-driven models—have been widely recognized in the wind forecasting literature and practice to possess higher skill at shorter forecast horizons (i.e., few minutes up to few hours ahead) [7]. Those models primarily rely on local time series observations, potentially recorded at multiple sites, to learn and extrapolate historical trends, patterns, and correlations into the near future. The last couple of decades have seen tremendous advancements in the development and adoption of data-science-based models in short-term wind forecasting applications. Examples include but are not limited to autoregressive-based models [8], [9], kernel- and tree-based approaches [10], [11], spatio-temporal statistical methods [12]–[14], and most recently, deep-learning-based models [15], [16].

This paper is primarily concerned with *statistical spatio-temporal models* for short-term wind forecasting, i.e., approaches that leverage local observations in the form of time series data recorded simultaneously at multiple proximate locations, in order to make probabilistic short-term wind resource or power forecasts. Understandably, the spatial proximity of those time series observations provide a strong case to a joint modeling approach that can characterize and leverage the inherent spatial and temporal dependencies for enhanced learning and forecasting skill [17]. While there has been an emerging literature on designing effective deep learning models for spatio-temporal wind forecasting [18], [19], statistical spatio-temporal methods—which are the focus of this work—remain a viable and attractive alternative by virtue of their natural probabilistic characterization, model interpretability, and fairly low data requirements. Prevalent statistical approaches for spatio-temporal wind forecasting often invoke adding exogenous spatial regressors in time series models [20], spatio-temporal and vector autoregressive modeling [9], [21], and geostatistical approaches with appropriate covariance structures [22], [23].

An alternative—albeit fairly under-utilized—approach for spatio-temporal statistical forecasting is the so-called "Integro-Difference Equation" (IDE) framework. IDEs describe the evolution of a spatial process over time using a Markovian-like dependence structure [17], making them well-suited to model highly dynamical processes like local wind conditions over a fairly short time horizon. Let $Y_t(\mathbf{s})$ be the spatio-temporal process denoting the hub-height wind speed such that $\mathbf{s} \in D_s$ and $t \in \mathbb{Z}^+$ are the spatial location and time index, respectively, then the key defining aspect in IDE models is the specification of a "redistribution kernel," denoted as $m(\mathbf{s}, t, \boldsymbol{\theta})$, which broadly governs the evolution of the spatial wind field

This work has been supported in part by the National Science Foundation (Award #: ECCS-2114422) and in part by the National Offshore Wind Research & Development Consortium (Project # 192900-133).

†Corresponding author; contact e-mail: aziz.ezzat@rutgers.edu

over time, as expressed in (1).

$$Y_t(\mathbf{s}) = \int_{D_s} m(\mathbf{s}, \mathbf{x}; \boldsymbol{\theta}) Y_{t-1}(\mathbf{x}) d\mathbf{x} + \eta_t(\mathbf{s}), \tag{1}$$

where $\boldsymbol{\theta} \in \mathbb{R}^d$ are the kernel parameters and $\eta_t(\mathbf{s})$ is a spatio-temporal error term.

Gaussian, squared exponential, Matérn and other off-the-shelf kernel functions have been popular choices for $m(\mathbf{s}, t, \boldsymbol{\theta})$. Those often invoke a "best-fitting" constant value for $\boldsymbol{\theta}$ that is either time-invariant, or, at best, updated over time in a rolling window fashion [24], [25]. This assumption, albeit significantly simplifying the inference of the IDE model may often be unrealistic for complex physical processes like wind fields where the dynamics governing wind evolution are known to change over time and space. Understandably, directly making $\boldsymbol{\theta}$ temporally and/or spatially varying can significantly inflate the parameter space for the IDE model. Instead, a viable alternative is to parameterize $\boldsymbol{\theta}$ as function of time, space, and/or spatio-temporal covariates [26], [27].

Along those lines, we propose a hierarchical IDE-based model with a physically meaningful kernel, for which key parameters in $\boldsymbol{\theta}$ are modeled as the output of a latent spatio-temporal regression model, enabling them to smoothly change over time as function of the local wind conditions, and hence are denoted as $\boldsymbol{\theta}_t(\mathbf{s})$. We show, through real-world experiments, that the proposed IDE model with a state-dependent kernel leads to improved forecasting accuracy, across both the wind speed and power domains. Our central focus is on a test case from the rising U.S. offshore wind energy sector in the U.S. Mid-Atlantic where several Gigawatt-scale wind farms are either planned or under-development. We make use of recently collected hub-height measurements from the offshore wind energy areas in this geographical region to evaluate the forecasting performance of our proposed approach against a representative set of time series and spatio-temporal methods that are prevalent in short-term wind forecasting applications.

The remainder of this paper is organized as follows. Section II describes the real-world data and test case that motivates this work. In Section III, we introduce our proposed IDE-based forecasting approach, which is then followed by Section IV where the experimental results are presented and discussed. Finally, Section V concludes the paper.

II. Data Description and Processing

This work is motivated by the ongoing large-scale offshore wind developments in the U.S. Mid -Atlantic and Northeastern U.S., and in particular, the New York/New Jersey (NY/NJ) Bight, which is set to host multiple Gigawatt-scale offshore wind projects [28]. In order to collect fine-resolution, hub-height wind data that are relevant for wind energy assessment and forecasting in this geographical region, the New York State Energy Research and Development Authority (NYSERDA) has deployed two floating Lidar buoys, referred to hereinafter as E05 and E06, respectively [29], The buoys are approximately 77 Km apart and their exact coordinates, co-located with the active offshore wind energy lease areas shown in Figure 1(a), are 39°58'10"N and 72°43'00"W for E05, and 39°32'50"N and 73°25'45"W for E06 [30].

Of relevance to this work are the 10-minute wind speed observations, recorded at a height of 140 meters, which is the closest altitude to the hub height of several modern-day offshore wind turbine designs [31]. The dataset used in this study span the winter months of November and December 2019, which experience strong wind speeds, with an average of 10.86 m/s. To pre-process the data and better align it with downstream modeling assumptions, we transform the raw data using a Box-Cox transformation with a time-varying parameter, λ_t, that is updated over time for each forecasting roll. Forecasts are then transformed back to their original domain using an inverse Box-Cox transformation. Figure 1(b) shows the time series of the hub-height wind speeds at E05 and E06 (shown in blue and red colors, respectively) which clearly show strong signs of spatial dependencies due to their geographical proximity, motivating a joint spatio-temporal modeling and forecasting approach, as proposed in this paper.

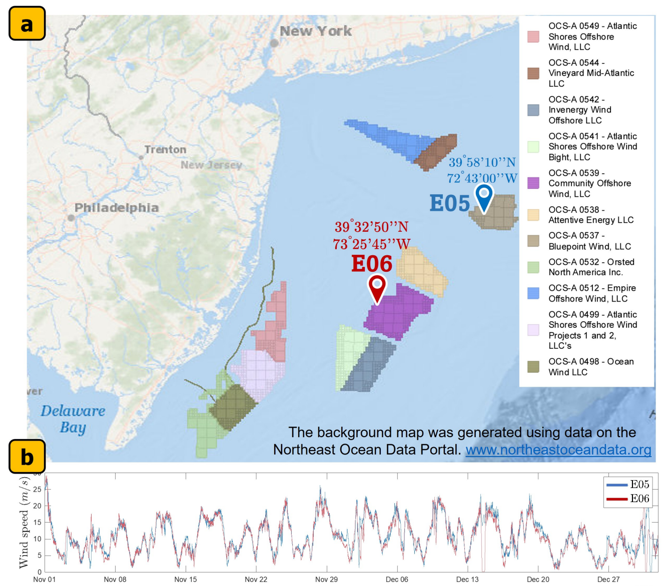

Fig. 1. (a) Locations of NYSERDA's E05 and E06 floating Lidar buoys on top of the active offshore wind energy lease areas in the U.S. Mid-Atlantic. The background map is generated using the Northeast Ocean Data Portal [30]. (b) Time series of the hub-height wind speeds (at 140-m altitude) at E05 (blue) and E06 (red), showing strong signs of spatial dependence.

III. An IDE-Based Approach with a Latent Spatio-Temporal Regression Model

Integro-difference equation (IDE) models describe the evolution of a spatial process over time by modeling the conditional dependence of the spatial process at the present time on its recent history. This is often modeled within a hierarchical framework comprising a *data model* (often referred to as the observation model) and a *process model* [17]. In this work, we assume the following structure for the *data model*:

$$\mathbf{Z}_t = \mathbf{H}_t \mathbf{Y}_t + \boldsymbol{\epsilon}_t, \tag{2}$$

where $\mathbf{Z}_t = [Z_t(\mathbf{s}_1), ..., Z_t(\mathbf{s}_m)]^T$ is an observation vector, such that $Z_t(\mathbf{s})$ is the wind speed measurement at location $\mathbf{s}$ and time t. Similarly, we have $\mathbf{Y}_t = [Y_t(\mathbf{s}_1), ..., Y_t(\mathbf{s}_n)]^T$, such that $Y_t(\mathbf{s})$ is the underlying state process of interest at location $\mathbf{s}$ and time t. The matrix $\mathbf{H}_t$ is an $m \times n$ measurement operator matrix, whereas $\boldsymbol{\epsilon}_t$ is the zero-mean Gaussian random variable denoting measurement error with variance $\sigma^2_{\boldsymbol{\epsilon}}$.

Considering the spatio-temporal process in continuous space and discrete time, $\{Y_t(\cdot)\}$, the correspondent IDE *process model* with time-varying kernel parameters can be written as:

$$Y_t(\mathbf{s}) = \int_{D_s} m(\mathbf{s}, \mathbf{x}; \boldsymbol{\theta}_t(\mathbf{s})) Y_{t-1}(\mathbf{x}) d\mathbf{x} + \eta_t(\mathbf{s}), \quad (3)$$

where $m(\mathbf{s}, \mathbf{x}; \boldsymbol{\theta}_t(\mathbf{s}))$ is the so-called "redistribution kernel" defining the weighted contribution of the process at location $\mathbf{x} \in D_s$ at time $t-1$ in determining $Y_t(\mathbf{s})$. The kernel is described by the parameter vector $\boldsymbol{\theta}_t(\mathbf{s})$, which are treated in this work as spatio-temporal functional parameters (more on that in the sequel). We further assume $\eta_t(\mathbf{s}) \sim \mathcal{N}(0, \sigma^2_\eta)$.

Defining $m(\mathbf{s}, \mathbf{x}; \boldsymbol{\theta}_t(\mathbf{s}))$ is essential for IDE models. In this work, we adopt a Lagrangian kernel that coincides with the physical feature of advection (often known as the transport effect) of wind fields [32], [33], as in (4).

$$m(\mathbf{s}, \mathbf{x}; \boldsymbol{\theta}_t(\mathbf{s})) = \sigma^2_m \exp(-||\mathbf{s} - \mathbf{x} - \boldsymbol{\phi}_t(\mathbf{s})||), \quad (4)$$

where $\boldsymbol{\theta}_t(\mathbf{s}) := \{\sigma^2_m, \boldsymbol{\phi}_t(\mathbf{s})\}$ such that σ^2_m is the marginal variance of the kernel, while $\boldsymbol{\phi}_t(\mathbf{s}) \in \mathbb{R}^2$ is an advection parameter that captures the transport effect of the wind field, allowing along-wind dependence to be stronger than opposite-wind dependence. In contrast to directly learning a best-fitting constant value for $\phi^1_t(\mathbf{s})$ and $\phi^2_t(\mathbf{s})$, we model each of them as the output of a latent spatio-temporal regression model that depends on the local wind conditions at previous time steps and multiple locations, as in (5).

$$\phi^i_t(\mathbf{s}) = \alpha^i_0 + \sum_{k=1}^{K} \sum_{j=1}^{\tau} \alpha^i_{k,j} Y_{t-j}(\mathbf{s}_k), \quad i \in \{1, 2\}, \quad (5)$$

where α^i_0 and $\{\alpha^i_{k,j}\}^{K,\tau}_{k=1,j=1}$ are unknown sets of regression coefficients. By making $\boldsymbol{\theta}_t(\mathbf{s})$ depend on $Y_t(\mathbf{s})$, we inherently make the IDE kernel state-dependent.

Given $\boldsymbol{\Theta} = \{\boldsymbol{\theta}_t(\cdot)\}^T_{t=1}$, $\mathbf{Z}_{1:T} \equiv [\mathbf{Z}_1, ..., \mathbf{Z}_T]$, where T is the present time after which forecasts are to be made, then the likelihood function $\mathcal{L}(\mathbf{Z}_{1:T}|\boldsymbol{\Theta})$ can be generally written as in (6). The estimation can then proceed using standard Kalman Filter machinery in order to obtain the forecast and filtering distributions [17]. Figure 2 illustrates how the proposed IDE model is trained and used to make forecasts once the parameters have been learned.

$$\mathcal{L}(\mathbf{Z}_{1:T}|\boldsymbol{\Theta}) = \prod_{t=1}^{T} \mathcal{P}(\mathbf{Z}_t|\boldsymbol{\theta}_t(\cdot)). \quad (6)$$

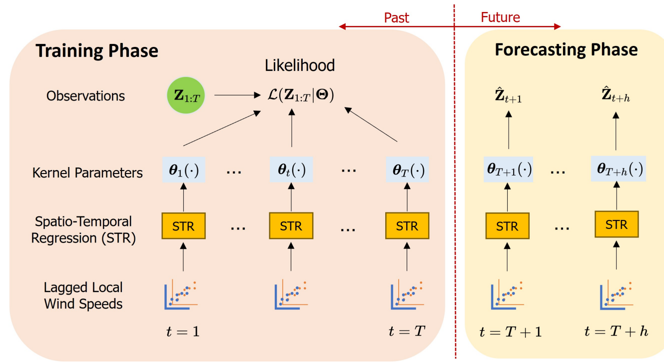

Fig. 2. Illustration of the proposed hierarchical IDE-based framework, including the training (left) and forecasting (right) phases.

IV. Real-world Case Study

We evaluate the proposed IDE-based approach using a rolling forecasting scheme, for forecast horizons, $h \in \{1, ..., 6\}$ hours in 10-minute increments, that is, we have a total of 36 forecasts per spatial location. The rolling window is six hours. For each forecasting roll, we re-train the model by updating its parameters, make future spatio-temporal forecasts, then roll by another six hours, and repeat the whole training/forecasting process. This leads to a total of 220 rolls. For the training process, we find that five days of historical data is a reasonable training data size to balance goodness of fit and computational efficiency.

A. Wind Speed Forecasting Results

We dub our approach as IDE-STR which stands for Integro-Difference Equation with Latent Spatio-Temporal Regression. We compare the wind speed forecasts obtained from IDE-STR against the following set of representative methods that are commonly used for spatio-temporal or time series modeling in short-term wind forecasting:

- GP: A spatio-temporal Gaussian Process model with a squared exponential covariance function for which the hyperparameters are learnt and updated at each forecasting roll using maximum likelihood estimation (MLE).
- ARIMA(p,q,d): An autoregressive integrated moving average model trained separately for each location. Bayesian information criterion (BIC) and MLE are used to dynamically update the model order parameters, and model coefficients, respectively, for each forecasting roll.
- IDE-CNT: This is the base version of the IDE model with the best-fitting constant value for the kernel parameters, which we update for each forecasting roll. The key difference between IDE-STR and IDE-CNT is that the latter does not invoke the latent spatio-temporal regression (STR) model in (5).

Table I shows the mean absolute error (MAE) values for the wind speed forecasts of all competing models at various forecast horizons. Looking at Table I, we can draw few key insights. First, it is clear that, on average, IDE-STR achieves the lowest average MAE with percentage improvements ranging from 1.93% relative to spatio-temporal GPs and up to

2.56% relative to IDE-CNT (the base version of IDE with time-invariant kernel parameters). Second, we note that the improvement of IDE-based models (IDE-STR and IDE-CNT) increases as the forecast horizon gets longer ($h \geq 3$). For example, the improvements of IDE-STR over ARIMA reaches up to 3.6% at $h = 6$ hours ahead. In contrast, the ARIMA model appears to perform fairly well for ultra short-term forecasting ($h \leq 2$ hours ahead). The trend is reversed for longer horizons, where the spatial neighborhood effect appears to play a noticeable role as evident by how spatio-temporal models (IDE-STR, IDE-CNT, and GP) significantly outperform the ARIMA model at $h > 2$ hours ahead, with increasing percentage improvements as h is longer. Third, we find that the improvement of IDE-STR over its base version, IDE-CNT, appears to be more amplified at longer horizons ($h \geq 2$). This showcases that, albeit a constant value for $\boldsymbol{\theta}_t$ might be acceptable for shorter time scales, capturing the temporally- and spatially-varying nature of the advection parameters through our proposed STR approach brings in significant value when the wind dynamics in the forecast horizon start departing from those at the time of forecast.

TABLE I
WIND SPEED FORECASTING RESULTS. GREY-COLORED CELLS REPRESENT RESULTS AT E06. AV. AND % DENOTE AVERAGE PERFORMANCE AND PERCENTAGE IMPROVEMENT OF IDE-STR OVER ALL MODELS, RESPECTIVELY. BOLD-FACED VALUES DENOTE BEST PERFORMANCE.

	IDE-STR		IDE-CNT		ARIMA(p,q,d)		GP	
	E05	E06	E05	E06	E05	E06	E05	E06
1	**0.76**	0.69	**0.76**	0.71	0.77	**0.66**	0.78	0.69
2	**1.31**	**1.18**	**1.31**	1.26	1.32	**1.18**	1.33	**1.18**
3	**1.72**	**1.66**	**1.72**	1.76	1.74	1.67	1.75	1.69
4	**2.07**	**2.09**	2.09	2.21	2.13	2.11	2.12	2.14
5	**2.31**	**2.40**	2.32	2.51	2.41	2.45	2.37	2.46
6	**2.61**	**2.85**	2.62	2.95	2.74	2.92	2.67	2.91
Av.	**1.80**		1.85		1.84		1.84	
%	-		2.56%		1.97%		1.93%	

Figure 3 shows the normalized values of the estimates of $\boldsymbol{\theta}_t$ from IDE-STR versus that from IDE-CNT, showing how the latent STR model adequately captures the spatio-temporal variation of the advection parameters that may be motivated by evolving wind dynamics. In contrast, a rolling window update of a single value for $\boldsymbol{\theta}_t$ as in IDE-CNT appears to be rigid and offers far less modeling flexibility. Figure 4 shows probabilistic forecasts issued using IDE-STR which suggest faithful agreement with the actual measurements. This inherently probabilistic nature of IDE-based models make them well-suited to inform subsequent operational decision-making under uncertainty [34], [35].

B. Wind Power Forecasting Results

To further demonstrate the value of IDE-STR to offshore wind operations, we transform its wind speed forecasts into wind power predictions using statistically constructed power curves obtained using the standard method of bins [36]–[38] applied to actual SCADA data from an operational wind farm in the US [39]. We scale the power output to the $[0, 1]$ interval, wherein a value of 1 denotes the rated capacity. The power

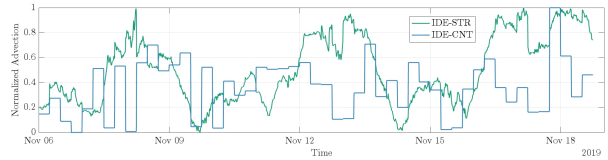

Fig. 3. Normalized values for $\boldsymbol{\theta}_t$ for IDE-STR versus those from IDE-CNT (updated on a rolling window fashion), clearly showing the modeling flexibility of the former in capturing the spatially- and temporally- varying advection dynamics relative to the latter.

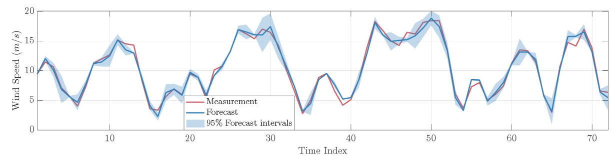

Fig. 4. Probabilistic forecasts from IDE-STR on top of the actual wind speed measurements (95% forecast intervals shown).

curves are then used as "look-up tables" to convert both the true wind speeds and their correspondent forecasts for all competing methods into wind power predictions. We then evaluate the power predictions from all methods using the power curve error (PCE) loss [40], which is defined as in (7).

$$PCE(P, \widehat{P}) = \begin{cases} g\left[P_{T+h}(\mathbf{s}) - \widehat{P}_{T+h}(\mathbf{s})\right] \\ \quad \text{if } \hat{Z}_{T+h}(\mathbf{s}) \leq Z_{t+h}(\mathbf{s}), \\ (1-g)\left[\widehat{P}_{T+h}(\mathbf{s}) - P_{T+h}(\mathbf{s})\right] \\ \quad \text{if } \hat{Z}_{T+h}(\mathbf{s}) > Z_{T+h}(\mathbf{s}), \end{cases} \tag{7}$$

where $P_{T+h}(\mathbf{s})$ and $\widehat{P}_{T+h}(\mathbf{s})$ are the normalized power observations and forecasts at location s at time $T + h$, whereas g is the under-estimation weight, which is typically set at values higher than 0.5 [40]. Table II, shows the average PCE values across all horizons for values of g ranging between 0.5 and 0.8 with 0.1 increment. Again, IDE-STR is shown to outperform all of its competitors, further confirming the merit of the proposed IDE-based framework in offshore wind power forecasting applications.

V. CONCLUSIONS

In this work, we proposed a hierarchical approach for short-term wind speed and power forecasting based on the integro-difference equation modeling framework wherein the kernel parameters are modeled as the output of a latent

TABLE II
WIND POWER FORECASTING RESULTS FOR ALL BENCHAMRKS. BOLD-FACED VALUES DENOTE THE BEST PERFORMANCE.

g	IDE-TSR	IDE-CNT	ARIMA	GP
0.5	**0.0593**	0.0613	0.0597	0.0599
0.6	**0.0567**	0.0584	0.0570	0.0575
0.7	**0.0542**	0.0555	0.0544	0.0551
0.8	**0.0516**	0.0526	0.0518	0.0528

spatio-temporal regression model. Tested on actual hub-height measurements, the proposed approach is shown to provide considerable improvements relative to time series and spatio-temporal modeling approaches that are prevalent in the short-term wind forecasting literature. Future research will explore more advanced modeling structures for the latent model within the IDE framework, as well as the integration of exogenous weather information in order to extend the forecasting skill beyond few hours ahead.

References

[1] C. Sweeney, R. J. Bessa, J. Browell, and P. Pinson, "The future of forecasting for renewable energy," *Wiley Interdisciplinary Reviews: Energy and Environment*, vol. 9, no. 2, e365, 2020.

[2] A. Lorca and X. A. Sun, "Adaptive robust optimization with dynamic uncertainty sets for multi-period economic dispatch under significant wind," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 1702–1713, 2014.

[3] M. S. Modarresi, L. Xie, M. C. Campi, *et al.*, "Scenario-based economic dispatch with tunable risk levels in high-renewable power systems," *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 5103–5114, 2018.

[4] P. Papadopoulos, D. W. Coit, and A. A. Ezzat, "Seizing opportunity: Maintenance optimization in offshore wind farms considering accessibility, production, and crew dispatch," *IEEE Transactions on Sustainable Energy*, vol. 13, no. 1, pp. 111–121, 2022. DOI: 10.1109/TSTE.2021.3104982.

[5] P. Papadopoulos, F. Fallahi, M. Yildirim, and A. A. Ezzat, "Joint optimization of maintenance and production in offshore wind farms: Balancing the short-and long-term needs of wind energy operation," *IEEE Transactions on Sustainable Energy*, 2023, accepted.

[6] M. F. Howland and J. O. Dabiri, "Wind farm modeling with interpretable physics-informed machine learning," *Energies*, vol. 12, no. 14, 2019.

[7] G. Giebel and G. Kariniotakis, "Wind power forecasting—a review of the state of the art," *Renewable energy forecasting*, pp. 59–109, 2017.

[8] P. Pinson and H. Madsen, "Adaptive modelling and forecasting of offshore wind power fluctuations with markov-switching autoregressive models," *Journal of Forecasting*, vol. 31, no. 4, pp. 281–313, 2012.

[9] J. Dowell and P. Pinson, "Very-short-term probabilistic wind power forecasts by sparse vector autoregression," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 763–770, 2015.

[10] J. Zeng and W. Qiao, "Support vector machine-based short-term wind power forecasting," in *2011 IEEE/PES power systems conference and exposition*, IEEE, 2011, pp. 1–8.

[11] G. I. Nagy, G. Barta, S. Kazi, G. Borbély, and G. Simon, "GEFCom2014: Probabilistic solar and wind power forecasting using a generalized additive tree ensemble approach," *International Journal of Forecasting*, vol. 32, no. 3, pp. 1087–1093, 2016.

[12] A. A. Ezzat, "Turbine-specific short-term wind speed forecasting considering within-farm wind field dependencies and fluctuations," *Applied Energy*, vol. 269, p. 115 034, 2020, ISSN: 0306-2619.

[13] F. Ye, J. Brodie, T. Miles, and A. A. Ezzat, "AIRU-WRF: A physics-guided spatio-temporal wind forecasting model and its application to the US Mid-Atlantic offshore wind energy areas," *arXiv preprint arXiv:2303.02246*, 2023.

[14] F. Ye, J. Brodie, T. Miles, and A. A. Ezzat, "Ultra-short-term probabilistic wind forecasting: Can numerical weather predictions help?" In *2023 IEEE Power Energy Society General Meeting (PESGM)*, 2023, pp. 1–5.

[15] Y. Ju, G. Sun, Q. Chen, M. Zhang, H. Zhu, and M. U. Rehman, "A model combining convolutional neural network and lightgbm algorithm for ultra-short-term wind power forecasting," *IEEE Access*, vol. 7, pp. 28 309–28 318, 2019.

[16] M.-S. Ko, K. Lee, J.-K. Kim, C. W. Hong, Z. Y. Dong, and K. Hur, "Deep concatenated residual network with bidirectional LSTM for one-hour-ahead wind power forecasting," *IEEE Transactions on Sustainable Energy*, vol. 12, no. 2, pp. 1321–1335, 2020.

[17] N. Cressie and C. K. Wikle, *Statistics for Spatio-Temporal Data*. John Wiley & Sons, 2015.

[18] M. Khodayar, O. Kaynak, and M. E. Khodayar, "Rough deep neural architecture for short-term wind speed forecasting," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 6, pp. 2770–2779, 2017.

[19] A. Ghaderi, B. M. Sanandaji, and F. Ghaderi, "Deep forecast: Deep learning-based spatio-temporal forecasting," in *The 34th International Conference on Machine Learning (ICML), Time series Workshop*, 2017.

[20] J. Tastu, P. Pinson, P.-J. 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, 2013.

[21] A. Pourhabib, J. Z. 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.

[22] N. Chen, Z. Qian, I. T. Nabney, and X. Meng, "Wind power forecasts using gaussian processes and numerical weather prediction," *IEEE Transactions on Power Systems*, vol. 29, no. 2, pp. 656–665, 2013.

[23] A. A. Ezzat, M. Jun, and Y. Ding, "Spatio-temporal short-term wind forecast: A calibrated regime-switching method," *The Annals of Applied Statistics*, vol. 13, no. 3, pp. 1484–1510, 2019.

[24] C. K. Wikle, "A kernel-based spectral model for non-gaussian spatio-temporal processes," *Statistical Modelling*, vol. 2, no. 4, pp. 299–314, 2002.

[25] C. K. W. Ke Xu and N. I. Fox, "A kernel-based spatio-temporal dynamical model for nowcasting weather radar reflectivities," *Journal of the American Statistical Association*, vol. 100, no. 472, pp. 1133–1144, 2005.

[26] E. De Bézenac, A. Pajot, and P. Gallinari, "Deep learning for physical processes: Incorporating prior scientific knowledge," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2019, no. 12, p. 124 009, 2019.

[27] A. Zammit-Mangion and C. K. Wikle, "Deep integro-difference equation models for spatio-temporal forecasting," *Spatial Statistics*, vol. 37, p. 100 408, 2020.

[28] NYSERDA, "Research and development roadmap version 2.0," New York State Energy Research and Development Authority, Tech. Rep., 2019.

[29] NYSERDA, *NYSERDA floating lidar buoy data*, https://oswbuoysny.resourcepanorama.dnvgl.com/, 2019.

[30] NROC, *2021 offshore wind energy lease areas in the u.s. mid-atlantic*, https://www.northeastoceandata.org/, 2023.

[31] DOE, *Wind turbines: The bigger, the better*, https://www.energy.gov/eere/articles/wind-turbines-bigger-better, 2023.

[32] M. L. O. Salvaña and M. G. Genton, "Nonstationary cross-covariance functions for multivariate spatio-temporal random fields," *Spatial Statistics*, vol. 37, p. 100 411, 2020, ISSN: 2211-6753.

[33] A. A. Ezzat, M. Jun, and Y. Ding, "Spatio-temporal asymmetry of local wind fields and its impact on short-term wind forecasting," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 3, pp. 1437–1447, 2018.

[34] P. Pinson, "Wind Energy: Forecasting Challenges for Its Operational Management," *Statistical Science*, vol. 28, no. 4, pp. 564–585, 2013.

[35] P. Papadopoulos, D. W. Coit, and A. A. Ezzat, "STOCHOS: Stochastic opportunistic maintenance scheduling for offshore wind farms," *IISE Transactions*, pp. 1–15, 2022.

[36] "Wind Energy Generation Systems - Part 12-1: Power Performance Measurements of Electricity Producing Wind Turbines," *IEC 61400-12-1*, 2017, International Electrotechnical Commission.

[37] B. Golparvar, P. Papadopoulos, A. Ezzat, and R.-Q. Wang, "A surrogate-model-based approach for estimating the first and second-order moments of offshore wind power," *Applied Energy*, vol. 299, p. 117 286, 2021.

[38] P. Nasery and A. Ezzat, "Yaw-adjusted wind power curve modeling: A local regression approach," *Renewable Energy*, vol. 202, pp. 1368–1376, 2023.

[39] Y. Ding, *Data Science for Wind Energy*. CRC Press, 2019.

[40] A. S. Hering and M. G. Genton, "Powering up with space-time wind forecasting," *Journal of the American Statistical Association*, vol. 105, no. 489, pp. 92–104, 2010.