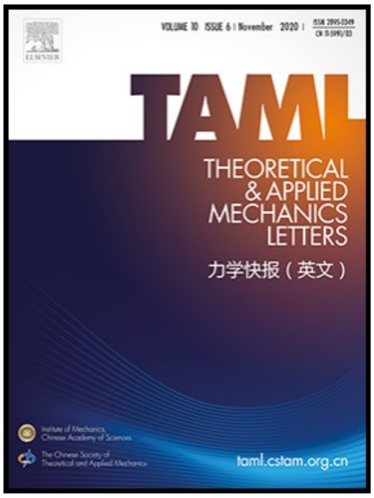
Journal Pre-proof



A Call for Enhanced Data-Driven Insights into Wind Energy Flow Physics

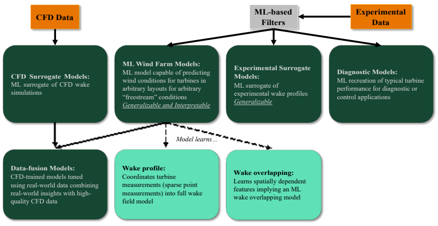
Coleman Moss, Romit Maulik, Giacomo Valerio Iungo

PII: S2095-0349(23)00059-4
DOI: https://doi.org/10.1016/j.taml.2023.100488
Reference: TAML 100488

To appear in: *Theoretical and Applied Mechanics Letters*

Received date: 11 October 2023
Revised date: 11 December 2023
Accepted date: 12 December 2023

Please cite this article as: Coleman Moss, Romit Maulik, Giacomo Valerio Iungo, A Call for Enhanced Data-Driven Insights into Wind Energy Flow Physics, *Theoretical and Applied Mechanics Letters* (2023), doi: https://doi.org/10.1016/j.taml.2023.100488


CFD Data
ML-based Filters
Experimental Data
CFD Surrogate Models: ML surrogate of CFD wake simulations
ML Wind Farm Models: ML model capable of predicting wind conditions for turbines in arbitrary layouts for arbitrary "freestream" conditions Generalizable and Interpretable
Experimental Surrogate Models: ML surrogate of experimental wake profiles Generalizable
Diagnostic Models: ML recreation of typical turbine performance for diagnostic or control applications
Model learns…
Data-fusion Models: CFD-trained models tuned using real-world data combining real-world insights with high-quality CFD data
Wake profile: Coordinates turbine measurements (sparse point measurements) into full wake field model
Wake overlapping: Learns spatially dependent features implying an ML wake overlapping model

Highlights

- Machine learning models of wind turbine wakes are typically used to generate surrogates of CFD-simulated data rather than enabling new physics insights from real-world data.
- Use of real-world data is essential to gain new insights into turbine-wake processes, yet implies challenges related to data filtering, statistical stationarity, and accuracy, as well as spatio-temporal resolution.
- Three strategies are identified to incorporate real-world data into machine-learning models of wind turbine wakes for enhanced data-driven modeling.

A Call for Enhanced Data-Driven Insights into Wind Energy Flow Physics

Coleman Moss[a], Romit Maulik[b], Giacomo Valerio Iungo[a,*]

[a] *organization=Wind Fluids and Experiments (WindFluX) Laboratory, Mechanical Engineering Department, The University of Texas at Dallas,addressline=800 W. Campbell Rd., city=Richardson, postcode=75080, state=TX, country=U.S.A.*

[b] *organization=College of Information Sciences and Technology, The Pennsylvania State University, city=State College, state=PA, postcode=16801, country=U.S.A.*

Abstract

With the increased availability of experimental measurements aiming at probing wind resources and wind turbine operations, machine learning (ML) models are poised to advance our understanding of the physics underpinning the interaction between the atmospheric boundary layer and wind turbine arrays, the generated wakes and their interactions, and wind energy harvesting. However, the majority of the existing ML models for predicting wind turbine wakes merely recreate CFD-simulated data with analogous accuracy but reduced computational costs, thus providing surrogate models rather than enhanced data-enabled physics insights. Although ML-based surrogate models are useful to overcome current limitations associated with the high computational costs of CFD models, using ML to unveil processes from experimental data or enhance modeling capabilities is deemed a potential research direction to pursue. In this letter, we discuss recent achievements in the realm of ML modeling of wind turbine wakes and operations, along with new promising research strategies.

Keywords: Machine Learning, Wake, Wind Turbine, Wind Farm, SCADA, Uncertainty

*Corresponding Author
Email address: valerio.iungo@utdallas (Giacomo Valerio Iungo)

Wind energy is a promising avenue for big-data analytics with applications including power forecasting, energy management, wind turbine condition monitoring and diagnostics, and maintenance of wind power plants [1]. Supervisory control and data acquisition (SCADA) enables monitoring of wind turbine operations by recording statistics of power capture, turbine settings, wind conditions, and fault parameters streamed continuously to the control room of a power station, then typically recorded as statistics over ten-minute periods [2, 3, 4, 5, 6]. Considering that in 2021 the global installed wind capacity was 906 GW [7] and average onshore wind turbines are rated about 3 MW, it is easy to picture the large scale of existing data collection. Further, wind turbine operations are often assisted by monitoring wind resources at different spatio-temporal scales, not only through anemometers installed on the turbine nacelle, but also from nearby meteorological towers [8] and various remote sensing techniques, such as light detection and ranging (LiDAR) [9, 10, 11], and, for research projects, even radars [12], unmanned aerial vehicles [13], aircraft [14], and satellites [15].

In wind energy, machine learning (ML) methods have been widely applied for data filtering [16, 17] and modeling power curves [18, 19, 20, 21], wakes [22, 23, 24, 25], and wind farms [26, 27, 28, 29, 30], as well as forecasting power and wind speed over various time horizons [31, 32, 33, 34]. ML provides analysis of observations without bias of preconceived notions (a threat for classical statistical approaches using user-selected parameters and thresholds), identification of unknown features and processes (a potential limitation for existing first-principle models), and enhancement of data quality through noise removal [35, 36] and super-resolution models [37, 31, 38]. Generalizability, interpretability, and explainability are key properties to maximize the scientific impact of ML models [26, 39, 40]. ML models should provide predictive capabilities even for new parameter values within the training ranges, avoiding overfitting the training data (generalizability) [41]. Further, the user should be able to interpret the predictions generated and the variability of the "weights" of the various ML nodes based on the analysis of the input variability [27], or interpret input parameter impact via other methods [42], as well as explain what (physical) processes might have led to the predictions obtained as a result of the input variability [43]. In this letter, we stress that the majority of the current applications of ML for wind turbine wake modeling fail to accomplish all these goals simultaneously, which sets this work apart from previous reviews on data-driven wind farm modeling [26], while underlying the research potential in working with real-world data.

The flow region past an operating wind turbine, i.e., the wind turbine wake, is characterized by lower wind speed than the freestream condition, which is the result of extracting kinetic energy to generate electricity [9, 10, 44, 45], and by enhanced turbulence intensity (*TI*) due to the vorticity structures generated from the turbine blade rotation, their instabilities [46, 47, 48], and the mechanically-generated turbulence associated with the wake velocity shear [5, 49, 50, 51]. Since wind turbine power is related to the cube of the incoming wind speed, reductions in speed entail power losses for downwind wind turbines. However, it is noteworthy that uncertainty or error in wind speed predictions is magnified when predicting power capture, thus they can affect estimates of power losses [52]. The design of the farm layout should be optimized to minimize wake interactions among neighboring turbines and maximize farm power capture. This task requires simulating wakes generated by multiple wind turbines for many inflow wind conditions, which quickly add to a large number of simulations to be performed and, thus, the need to use models with very low computational costs and, if possible, adequate resolution [53, 54].

Current wake models typically fall into one of two categories: analytical and computational. Analytical (engineering) wake models are based on reduced-order physics formulations and are computationally cheap at the cost of accuracy [55]. Computational fluid dynamics (CFD) methods, such as large eddy simulation (LES) or Reynolds-averaged Navier Stokes (RANS), are computationally expensive but provide much higher accuracy [56, 57, 58, 59, 60]. The ML wake models described below attempt to combine the high accuracy of the computational methods with the low computational cost of the analytical methods by functioning as surrogate models for computational simulations. While successful, they may fail to be generalizable, interpretable, explainable, or built on real-world data.

Computational Surrogate Wake Models - This letter provides a concise overview of the main existing strategies for ML wind farm modeling; however, it is not meant to be a comprehensive review of the state-of-the-art on this research topic. The work by Ti *et al.* [22] demonstrates the standard approach for works seeking to generate surrogate wake models from simulation data using ML. The authors used a RANS solver coupled with an actuator disk model with rotation to simulate the three-dimensional velocity deficit and *TI* on a structured grid in the wake of a single turbine. This grid was then separated into 2000 partitions and a unique ML model was trained to predict the velocity deficit or turbulence intensity of a partition as a function of the hub height inflow velocity and *TI*. By using linear or sum-of-squares wake overlapping approaches [61], they simulated an offshore wind farm using the proposed ML method and compared it against SCADA and LES data, finding a good agreement. Purohit and coworkers [62] used high-fidelity CFD data to benchmark XGBoost, Support Vector Regression, and artificial neural network (ANN) algorithms. Each model predicted the velocity at a given point in the wake of a single turbine as a function of the incoming wind speed, the turbines thrust force, and the downstream position relative to the turbine.

Two works used a convolution neural network (CNN) in interesting ways. First, Zhang and Zhao [24] combined a CNN with a generative adversarial network (GAN), which was comprised of a generator network and a discriminator network. In that work, the generator network produced snapshots of a wake flow using the upstream wind profile and turbine yaw misalignment as inputs. In contrast, the discriminator network determined whether a given snapshot was real or generated. The two networks were trained adversarially, such that as the discriminator improved in detecting generated snapshots, the generator improved at generating realistic snapshots. Second, Li and coworkers [63] split the wake field prediction into two CNN problems: the foreground and background. The background problem attempts to predict the future wake velocity field from past wake snapshots. The foreground problem tries to predict future snapshots as a function of past impacting wind conditions and turbine yaw misalignment. Other works use graph neural networks (GNN), which learn data on unstructured grids, to learn the wake profile without the need for a uniform CFD grid [64].

When an ML model is trained from simulation data to predict wind or wind turbine parameters from some relevant information, usually the incoming wind conditions, the resulting models are typically assessed against CFD data [24, 62, 63], while in other cases the models are also compared against real-world measurements [22, 64]. Considering the main properties of a desirable ML model, it is not specified to what extent these models may be generalizable to other turbine models or site climatological conditions than those encompassed within the train-ing datasets. Nonetheless, they are generally difficult to interpret. It is not immediately clear how the input wind conditions affect the weights of the various nodes of the ML models and out-put wake properties. More generally, it is unclear which inputs mainly govern variability in the generated outputs or which input sets are optimal to enhance prediction accuracy. Finally, since these models are not trained on real-world data, a systemic problem results wherein the models can only be as accurate and informative as the data on which they were trained. The ML models described above make CFD data quickly reproducible, which is a valuable feature, but they do not expand upon it. By using CFD data to train ML models, several challenges inherent to using real-world data, such as instrument noise and bias, marginal statistical convergence, and limited spatial resolution, among others, are removed. This allows the development of ML models to be the focus. While the result is a better understanding of the modeling process and features, it may fail to enable scientific insights, which can be encompassed in real-world data.

Potential Advancements for Machine Learning Modeling of Wind Turbine Wakes - Given the above considerations, how can ML be used to generate accurate predictions of wind turbine wakes and advance the scientific understanding of wake processes by leveraging the availability of big data from wind turbines and the wind field? Returning to the main properties of improved ML models (i.e., being built on real-world data, generalizable, interpretable, and explainable), we propose three approaches. The first approach consists of combining ML models with historical wind turbine and wind velocity data to generate diagnostic models. The second provides ML wake models trained from real-world wake data and used for general wind turbine and wind farm applications. The third approach encompasses the evolution and overlapping of multiple wakes for generic wind farm layouts and site climatological conditions. In all cases, the key to obtaining new insights is to train the models on real-world data rather than simulated data. While this may yield scientific insights unavailable to CFD, specific drawbacks of real-world data should be kept in mind. Importantly, any bias or large uncertainty in the training data will impact model accuracy. It is often difficult to determine the uncertainty in ML predictions but this becomes more critical when the training data carries inherent statistical uncertainty.

Approach 1: Diagnostic ML Wake Models - Taking a simple approach first, ML can pull new insights from historical data through specific models. For example, the power produced by a given wind turbine is a function of the incoming wind speed, which is a function of the ambient wind condition and any upstream wind turbines that can affect turbine performance by generating wakes. A model can be trained on historical data to predict wind speed at a given turbine as a function of the ambient wind conditions. If wind direction is considered as an input, then the resulting ML diagnostic model will implicitly contain information on the location of neighboring wind turbines. Therefore, the resultant model is case-specific since it cannot be applied to any other turbines because the relative positions of neighboring turbines would change and the obtained predictions would be incorrect.

Although non-generalizable, this kind of model is still useful for diagnostic purposes or control applications. By carefully probing the model under specific ambient conditions, the wake effect can be identified, as well as any other effects contained in the historical data. It should be pointed out that a similar result could also be obtained using a statistical binning analysis. However, a statistical approach will suffer potential drawbacks associated with the definition of bin centroids and width, which are determined by preconceived, and potentially inaccurate, notions. In contrast, the ML approach lets any features contained in the data surface naturally, without any arbitrarily imposed constraints. This modeling approach may hold advantages over model-based approaches by predicting off-design conditions, such as strong yaw misaligned operations or unconventional weather conditions, when those events occur with sufficient frequency in the training data.

Recent work has already demonstrated how ML methods can extract wake information unavailable via typical statistical methods [65] while others have applied this approach to an onshore wind farm [66]. Considering the above-mentioned key properties of ML models, this approach is not general but specific to individual turbines. Furthermore, this approach is not very interpretable, simply treating the ML models as black boxes. Nonetheless, these models are trained on real-world data and, thus, may expand our physical knowledge. Figure 1 summarizes

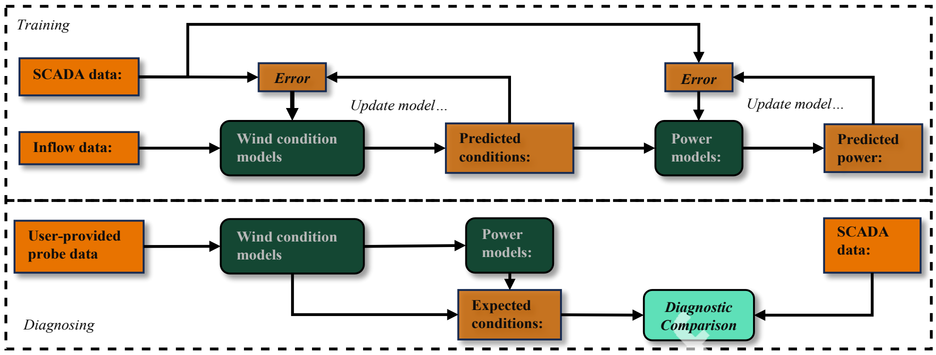


Fig. 1: The training and testing of a diagnostic ML model. Using inflow data as input, wind condition models are trained to predict wind conditions at a specific turbine location. These predictions are then used as inputs for models that predict turbine power. Combined, the result is turbine power predicted as a function of inflow, including wake interactions and other turbine-array effects. To use these models diagnostically, the user provides inflow data and compares the predicted performance against SCADA data to detect off-design performance.

training and using diagnostic ML models.

Approach 2: Data-driven ML wake models trained on real-world measurements - Considering the second approach for training ML wake models on real-world data, the intrinsic challenges associated with collecting and post-processing real-world data, such as data from SCADA, met-towers, and other remote sensing instruments [67, 68], should be noted. The main idea to develop ML wake models trained on real-world data is to replicate results obtained with ML wake models trained on CFD data but using real-world data. A recent work by Renganathan *et al.* [25] demonstrated how two-dimensional wake velocity fields can be predicted using scanning Doppler LiDAR data, providing as inputs the incoming atmospheric conditions. Indeed, that work could provide much new insight into wake dynamics, though the authors focused mainly on validating the models rather than probing them. If the data is appropriately non-dimensionalized, this approach should be generalizable to different turbine models than the training turbine. Since it is also trained on real-world data, it meets two of the above-mentioned ideal properties for ML models. Nonetheless, the ML model involved is still mostly a black box, so this approach struggles to be interpretable, though it may be explainable.

Approach 3: Data-driven ML wind farm modeling - Finally, an ML wind farm model, which we will argue meets all desirable ML model properties and can provide new insights, such as wake overlapping leading to wind farm wakes, should be capable of predicting wind conditions at constituent turbines for arbitrary inflow conditions and farm layouts while being trained on SCADA data. By requiring the model to be trained on SCADA data while also being able to handle arbitrary layouts, the model is forced to learn spatially dependent effects. This equates to coordinating sparse velocity point measurements, treating individual turbines as point measurements, into a full wake profile. Since this occurs at the farm level, wake overlapping effects are also learned. Figure 2 illustrates an ML wind farm model that learns wake profiles and wake overlapping.

Some researchers have developed such ML wind farm models achieving different levels of performance. Howland and Dabiri modeled turbine power capture as a regression problem utilizing power data of upstream turbines [27]. While the result is easily interpretable since neighboring turbines are given weights corresponding to their impact on the turbine of interest, it seems that each turbine requires a different model. The model is therefore not generalizable. Sun *et al.* trained an ANN model to predict the total power of an array of five turbines [69]. Notably, they avoided using wind direction as an input, which often results in non-generalizable models. Instead, they calculate the wake factor of each turbine for each wind direction using the 2D Jensen model and provide that as an input instead of the wind direction. The other inputs include turbine wind speeds and yaw misalignments. Given that no turbine positional information is provided, all wake interactions must be implicit in the data, and the model is likely specific to the training layout. Thus, the model would need to be retrained on any other layout, and the result is not generalizable.

Recent works offer two solutions to the problem of training generalizable ML wind farm models. Some utilize GNNs since GNNs learn spatial dependencies in non-uniform grids, which is suitable for studying different wind farm layouts [28, 29]. In these works, one uses physical guidance to assist the GNN [28], while the other uses attention mechanisms to enhance the interpretability of the network [29]. Yet none of these works use real-world SCADA data, but data generated from analytical wake models. Other works use SCADA data and graph structures but abandon the GNN for a simpler graph network, and thus fail to learn spatial dependencies [70].

The second approach to generalizing ML wind farm models is to encapsulate important layout information in a few parameters. Ghaisas and Archer [71] defined several such parameters. Of these parameters, additional works have used blockage ratio - the fraction of a wind turbine rotor shadowed by upstream rotors - and blockage distance - the distance of a given turbine to upstream turbines weighted by the amount of the rotor they shadow - for general wind farm models [30, 42]. For a given wind direction, these two parameters can be averaged across the entire farm, producing farm-averaged blockage ratio and distance. By replacing wind direction with these parameters and training an ANN to predict farm power from the wind speed and blockage parameters, Yan *et al.* [30] built a generalizable wind farm model. Other researchers even ex-

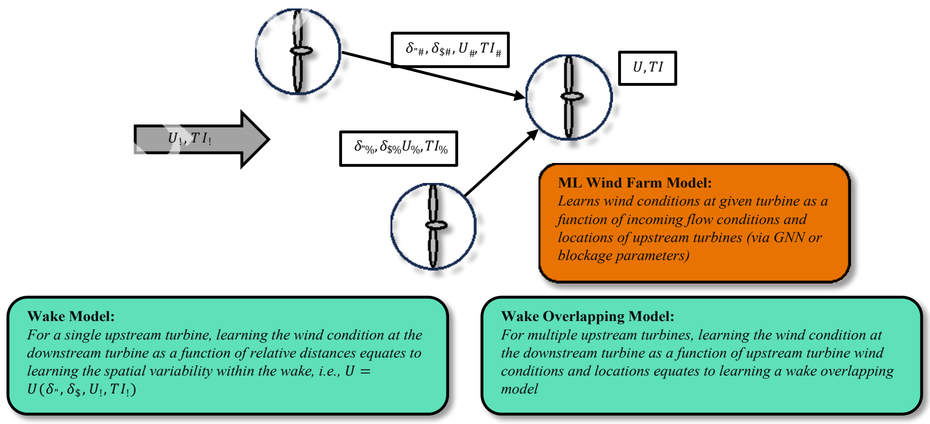


Fig. 2: The wind farm model, when able to generalize to arbitrary layouts, must learn the spatial dependency of velocity on upstream turbine locations, implying that the model learns the single turbine wake profile as well as wake overlapping.

tended this work to individual turbines and also introduced physics-informed models guided by basic analytic wake models [42]. Though it is not clear whether the resultant models are learning wake phenomena or if the blockage parameters encapsulate all the important wake information, this remains an interesting and promising research path. Additionally, encapsulating wake effects only through pure geometric parameters might be a pragmatic yet too simplistic approach. For instance, the defined parameters fail to encompass the modulation of wake velocity deficit and wake interactions with the atmospheric stability regime. Nonetheless, the geometric parameters can be considered an initial condition for ML wind-farm wake modeling, then further refined through an ML model to include additional layout and atmospheric effects.

Whether an ML wind farm model is developed based on GNN or blockage parameters, it should be generalizable, interpretable, explainable, and trained on real-world data. A promising approach might consist of blending training from simulation data, with ad-hoc refinement by re-training an ML model on real-world data, though such data fusion should be done carefully, lest the simulated data dominate the model performance [42]. The potential ML wind-farm wake model described should be capable of reproducing operations with arbitrary layouts and for different inflow conditions, thus it is generalizable. The resultant models should be interpretable as well. In the case of GNNs, this means describing the learned weights for interacting turbines as a function of turbine spacing or inflow conditions. For blockage parameter models, this can mean understanding the effect of changing farm layout on the blockage parameters and, thus, on the resultant predictions. In any case, the ML wind-farm wake model envisioned attempts to meet all the properties of a desirable ML model.

Outlook - We have reviewed several existing approaches for developing machine learning (ML) wake models for wind energy applications and scrutinized promising new research directions that can enhance the potential embedded in big data collected from wind power plants and measuring systems monitoring wind resources. It has been emphasized that an improved ML model of wind turbine wakes should be: 1) trained on real-world data to unveil physical processes governing wake evolution, interaction, and power capture; 2) generalizable to new values of the input parameters beyond those utilized to train ML models while still within the training ranges, avoiding extrapolation; 3) interpretable by identifying the variability in the “weights” of the ML models associated with the variability in the input parameters; 4) and explainable to identify the physical processes responsible for the input-output inter-dependencies. By fulfilling these properties, ML wake models will advance scientific understanding of the underlying phenomena governing wake evolution, interaction, and power capture, rather than simply recreate results achievable with current first-principle models, yet significantly reducing the required computational costs. Though using real-world data might seem a straightforward approach, it is by no means trivial. Real-world data are often noisy, statistically non-stationary, and collected with a relatively low spatio-temporal resolution. All these challenges need to be overcome when working with real-world data, and, even for these tasks, ML can be an invaluable resource.

We have discussed three main different research approaches to developing ML wake models trained on real-world data. For the first approach, we have considered diagnostic ML models, which can identify and predict different operative conditions of wind turbines based on the input parameters characterizing the freestream wind conditions. While providing good accuracy for practical applications, such as quantification of power wake losses, these models are not generalizable to other turbines than those considered for the ML model training. The second approach is to develop an ML wake model analogous to the surrogate wake models developed from numerical data. The main limitation is represented by the need for models for wake overlapping to reproduce wind farm flows. Finally, an ML wind-farm wake model might be generated using ad-hoc geometric parameters or graph neural networks (GNN). These models might encapsulate both the effects associated with the variability of the incoming wind/atmospheric conditions and the interactions with wakes generated by neighboring wind turbines. The latter seems to be the most comprehensive and promising approach for modeling wind farm flows. However, there are still important challenges to overcome, such as pre-processing of the available real-world data, accuracy, and generalizability of the ML models, and refinement of the models based on specific datasets.

Declaration of Competing Interests. The authors declare that they have no known compet-ing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements. This material is based upon work supported by the National Science Foundation (NSF) CBET, Fluid Dynamics CAREER program, award no. 2046160, program manager Ron Joslin.

References

[1] National Academy of Engineering, Frontiers of Engineering: Reports on Leading-Edge Engineering from the 2017 Symposium, National Academies Press, 2018.

[2] W. Yang, R. Court, J. Jiang, Wind turbine condition monitoring by the approach of SCADA data analysis, Renew. Energy 53 (2013) 365376.

[3] J. Tautz-Weinert, S. J. Watson, Using SCADA data for wind turbine condition monitoring a review, IET Renew. Power Generation 11 (2017) 382394.

[4] A. Zaher, S.D.J. McArthur, D.G. Infield, et al., Online wind turbine fault detection through automated SCADA data analysis, Wind Energy 12 (2009) 574593.

[5] S. El-Asha, L. Zhan, G.V. Iungo, Quantification of power losses due to wind turbine wake interactions through SCADA, meteorological and wind LiDAR data, Wind Energy 20 (2017) 18231839.

[6] K. Panthi, G.V. Iungo, Quantification of wind turbine energy loss due to leading-edge erosion through infrared-camera imaging, numerical simulations, and assessment against SCADA and meteorological data, Wind Energy 26 (2023) 266282.

[7] M. Hutchinson, F. Zhao, B. Backwell, et al., Global Wind Report 2023, Global Wind Energy Council (2023).

[8] J.K. Lundquist, J.M. Wilczak, R. Ashton, et al., Assessing State-of-the-Art Capabilities for Probing the Atmospheric Boundary Layer: The XPIA Field Campaign, Bull. of the American Meteorological Society 98 (2017) 289314.

[9] M. Puccioni, G.V. Iungo, C. Moss, et al., LiDAR Measurements to Investigate Farm-to-Farm Interactions at the AWAKEN Experiment, J. Phys.: Conf. Ser. 2505 (2023) 012045.

[10] G.V. Iungo, F. Port-Agel, Volumetric Lidar Scanning of Wind Turbine Wakes under Convective and Neutral Atmospheric Stability Regimes, J. of Atmospheric and Ocean. Technology 31 (2014) 20352048.

[11] M. Nafisifard, J.B. Jakobsen, J.T. Snbjrnsson, et al., Lidar measurements of wake around a bridge deck, J. of Wind Engineering and Industrial Aerodynamics 240 (2023) 105491.

[12] B.D. Hirth, J.L. Schroeder, W.S. Gunter, et al., Measuring a Utility-Scale Turbine Wake Using the TTUKa Mobile Research Radars, J. of Atmospheric and Ocean. Technology, 29 (2012) 765771

[13] W.H. Al-Sabban, L.F. Gonzalez, R.N. Smith, Wind-energy based path planning for Unmanned Aerial Vehicles using Markov Decision Processes, IEEE International Conference on Robotics and Automation (2013) 784789.

[14] A. Platis, S.K. Siedersleben, J. Bange, et al., First in situ evidence of wakes in the far field behind offshore wind farms, Sci Rep 8 (2018).

[15] M.B. Christiansen, C.B. Hasager, Wake effects of large offshore wind farms identified from satellite SAR, Remote Sens. of Environ. 98 (2005) 251268.

[16] B. Manobel, F. Sehnke, J.A. Lazzs, et al., Wind turbine power curve modeling based on Gaussian Processes and Artificial Neural Networks, Renew. Energy 125 (2018) 10151020.

[17] M. Yesilbudak, Implementation of novel hybrid approaches for power curve modeling of wind turbines, Energy Convers. and Management 171 (2018) 156169.

[18] Y. Wang, Q. Hu, L. Li, et al., Approaches to wind power curve modeling: A review and discussion, Renew. and Sustainable Energy Reviews 116 (2019) 109422.

[19] M. Lydia, A.I. Selvakumar, S.S. Kumar, et al., Advanced Algorithms for Wind Turbine Power Curve Modeling, IEEE Transactions on Sustainable Energy 4 (2013) 827835

[20] J. Zhou, P. Guo, and X.R. Wang, Modeling of wind turbine power curve based on Gaussian process, International Conference on Machine Learning and Cybernetics (2014) 7176.

[21] S. Pei and Y. Li, Wind Turbine Power Curve Modeling with a Hybrid Machine Learning Technique, Appl. Sciences 9 (2019).

[22] Z. Ti, X.W. Deng, H. Yang, Wake modeling of wind turbines using machine learning, Appl. Energy 257 (2020) 114025.

[23] Z. Ti, X.W. Deng, M. Zhang, Artificial Neural Networks based wake model for power prediction of wind farm, Renew. Energy 172 (2021) 618631.

[24] J. Zhang, X. Zhao, Wind farm wake modeling based on deep convolutional conditional generative adversarial network, Energy 238 (2022) 121747

[25] S.A. Renganathan, R. Maulik, S. Letizia, et al., Data-driven wind turbine wake modeling via probabilistic machine learning, Neural Comput. & Applic. 34 (2022) 61716186.

[26] N. Zehtabiyan-Rezaie, A. Iosifidis, M. Abkar, Data-driven fluid mechanics of wind farms: A review, J. of Renew. and Sustainable Energy 14 (2022) 032703.

[27] M.F. Howland, J.O. Dabiri, Wind Farm Modeling with Interpretable Physics-Informed Machine Learning, Energies 12 (2019).

[28] J. Park, J. Park, Physics-induced graph neural network: An application to wind-farm power estimation, Energy 187 (2019) 115883.

[29] L.Ø. Bentsen, N.D. Warakagoda, R. Stenbro, et al., Wind Park Power Prediction: Attention-Based Graph Networks and Deep Learning to Capture Wake Losses, J. Phys.: Conf. Ser. 2265 (2022) 022035.

[30] C. Yan, Y. Pan, C.L. Archer, A general method to estimate wind farm power using artificial neural networks, Wind Energy 22 (2019) 14211432.

[31] G. Liu, S. Zhang, H. Zhao, et al., Super-resolution perception for wind power forecasting by enhancing historical data, Frontiers in Energy Res. 10 (2022).

[32] J. Duan, H. Zuo, Y. Bai, et al., Short-term wind speed forecasting using recurrent neural networks with error correction, Energy 217 (2021) 119397.

[33] A.A. Ezzat, Turbine-specific short-term wind speed forecasting considering within-farm wind field dependencies and fluctuations, Appl. Energy 269 (2020) 115034.

[34] J. Zhang, X. Zhao, Digital twin of wind farms via physics-informed deep learning, Energy Convers. and Management 293 (2023) 117507.

[35] T.H. Foken, B. Wichura, Tools for quality assessment of surface-based flux measurements, Agricultural and For. Meteorology 78 (1996) 83105.

[36] T. Foken, M. Gockede, M. Mauder, et al., Post-Field Data Quality Control, in: X. Lee, W. Massman, B. Law (Eds.), Handbook of Micrometeorology: A Guide for Surface Flux Measurement and Analysis, Dordrecht Springer, Netherlands, 2005, pp. 181208.

[37] D.T. Tran, H. Robinson, A. Rasheed, et al., GANs enabled super-resolution reconstruction of wind field, J. Phys.: Conf. Ser. 1669 (2020) 012029.

[38] S. Maebius, N. French, P. Mason, Offshore Wind Field Super-Resolution Using Deep Learning Methods, SSRN (2023).

[39] S.L. Brunton, J.N. Kutz, K. Manohar, et al., Data-Driven Aerospace Engineering: Reframing the Industry with Machine Learning, AIAA J. 59 (2021) 28202847.

[40] R. Vinuesa, S.L. Brunton, Enhancing computational fluid dynamics with machine learning, Nat. Comput. Sci. 2 (2022).

[41] S.J. Andersen, J.P.M. Leon, Predictive and stochastic reduced-order modeling of wind turbine wake dynamics, Wind Energy Science 7 (2022) 21172133.

[42] N. Zehtabiyan-Rezaie, A. Iosifidis, M. Abkar, Physics-Guided Machine Learning for Wind-Farm Power Prediction: Toward Interpretability and Generalizability, PRX Energy 2 (2023) 013009.

[43] G. De Cillis, S. Cherubini, O. Semeraro, et al., Stability and optimal forcing analysis of a wind turbine wake: Comparison with POD, Renew. Energy 181 (2022) 765785.

[44] G.V. Iungo, Experimental characterization of wind turbine wakes: Wind tunnel tests and wind LiDAR measurements, J. of Wind Engineering and Industrial Aerodynamics 149 (2016) 3539.

[45] L. Zhan, S. Letizia, G.V. Iungo, LiDAR measurements for an onshore wind farm: Wake variability for different incoming wind speeds and atmospheric stability regimes, Wind Energy 23 (2020) 501527.

[46] G.V. Iungo, F. Viola, S. Camarri, et al., Linear stability analysis of wind turbine wakes performed on wind tunnel measurements, J. of Fluid Mechanics 737 (2013) 499526.

[47] F. Viola, G.V. Iungo, S. Camarri, et al., Prediction of the hub vortex instability in a wind turbine wake: stability analysis with eddy-viscosity models calibrated on wind tunnel data, J. of Fluid Mechanics 750 (2014) R1.

[48] S. Sarmast, R. Dadfar, R.F. Mikkelsen, et al., Mutual inductance instability of the tip vortices behind a wind turbine, J. of Fluid Mechanics 755 (2014) 705731.
[49] K.S. Hansen, R.J. Barthelmie, L.E. Jensen, et al., The impact of turbulence intensity and atmospheric stability on power deficits due to wind turbine wakes at Horns Rev wind farm, Wind Energy, 15 (2012) 183196.
[50] M. Magnusson, A.S. Smedman, Influence of Atmospheric Stability on Wind Turbine Wakes, Wind Engineering 18 (1994) 139152.
[51] M. Abkar, F. Port-Agel, Influence of atmospheric stability on wind-turbine wakes: A large-eddy simulation study, Physics of Fluids 27 (2015) 035104.
[52] T. Burton, N. Jenkins, D. Sharpe, et al., Wind Energy Handbook, first ed., Wiley, 2011.
[53] V. Santhanagopalan, S. Letizia, L. Zhan, et al., Profitability optimization of a wind power plant performed through different optimization algorithms and a data-driven RANS solver, in: Wind Energy Symposium, American Institute of Aeronautics and Astronautics (2018).
[54] R. Shakoor, M.Y. Hassan, A. Raheem, et al., Wake effect modeling: A review of wind farm layout optimization using Jensen's model, Renew. and Sustainable Energy Reviews 58 (2016) 10481059.
[55] J.K. Kaldellis, P. Triantafyllou, P. Stinis, Critical evaluation of Wind Turbines analytical wake models, Renew. and Sustainable Energy Reviews 144 (2021) 110991.
[56] K. Nilsson, S. Ivanell, K.S. Hansen, et al., Large-eddy simulations of the Lillgrund wind farm, Wind Energy 18 (2015) 449467.
[57] C. Santoni, E.J. Garca-Cartagena, U. Ciri, et al., One-way mesoscale-microscale coupling for simulating a wind farm in North Texas: Assessment against SCADA and LiDAR data, Wind Energy 23 (2020) 691710.
[58] M. Tabib, A. Rasheed, T. Kvamsdal, LES and RANS simulation of onshore Bessaker wind farm: analysing terrain and wake effects on wind farm performance, J. Phys.: Conf. Ser. 625 (2015) 012032.
[59] G.V. Iungo, V. Santhanagopalan, U. Ciri, et al., Parabolic RANS solver for low-computational-cost simulations of wind turbine wakes, Wind Energy 21 (2018) 184197.
[60] S. Letizia, G.V. Iungo, Pseudo-2D RANS: A LiDAR-driven mid-fidelity model for simulations of wind farm flows, J. of Renew. and Sustainable Energy 14 (2022) 023301.
[61] Z. Shao, Y. Wu, L. Li, et al., Multiple Wind Turbine Wakes Modeling Considering the Faster Wake Recovery in Overlapped Wakes, Energies 12 (2019).
[62] S. Purohit, E.Y.K. Ng, I.F.S.A. Kabir, Evaluation of three potential machine learning algorithms for predicting the velocity and turbulence intensity of a wind turbine wake, Renew. Energy 184 (2022) 405420.
[63] R. Li, J. Zhang, X. Zhao, Dynamic wind farm wake modeling based on a Bilateral Convolutional Neural Network and high-fidelity LES data, Energy 258 (2022) 124845.
[64] S. Li, M. Zhang, M.D. Piggot, End-to-end wind turbine wake modelling with deep graph representation learning, Appl. Energy 339 (2023) 120928.
[65] G.V. Iungo, R. Maulik, S.A. Renganathan, et al., Machine-learning identification of the variability of mean velocity and turbulence intensity for wakes generated by onshore wind turbines: Cluster analysis of wind LiDAR measurements, J. of Renew. and Sustainable Energy 14 (2022) 023307.
[66] C. Moss, R. Maulik, G.V. Iungo, Modeling Wind Turbine Performance and Wake Interactions with Machine Learning, arXiv (2022).
[67] R. Morrison, X. Liu, Z. Lin, Anomaly detection in wind turbine SCADA data for power curve cleaning, Renew. Energy 184 (2022) 473486.
[68] H. Beck, M. Khn, Dynamic Data Filtering of Long-Range Doppler LiDAR Wind Speed Measurements, Remote Sens. 9 (2017).
[69] H. Sun, C. Qiu, L. Lu, et al., Wind turbine power modelling and optimization using artificial neural network with wind field experimental data, Appl. Energy 280 (2020) 115880.
[70] F. Hammer, N. Helbig, T. Losinger, et al., Graph machine learning for predicting wake interaction losses based on SCADA data, J. Phys.: Conf. Ser. 2505 (2023) 012047.
[71] N.S. Ghaisas, C.L. Archer, Geometry-Based Models for Studying the Effects of Wind Farm Layout, J. of Atmospheric and Ocean. Technology 33 (2016) 481501.

Declaration of interests

☐The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Giacomo Valerio Iungo reports financial support was provided by National Science Foundation. Giacomo Valerio Iungo reports a relationship with National Science Foundation that includes: funding grants.