Spatio-Temporal Graph Neural Networks for Power Prediction in Offshore Wind Farms Using SCADA Data

Simon Daenens[1, 2, 3], Timothy Verstraeten[1, 2, 3], Pieter-Jan Daems[1, 2], Ann Nowé[3], and Jan Helsen[1, 2]

[1]OWI-Lab, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Elsene, Belgium
[2]Acoustics & Vibration Research Group, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Elsene, Belgium
[3]Artificial Intelligence Lab Brussels, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Elsene, Belgium

Correspondence: Simon Daenens (simon.daenens@vub.be)

Abstract. This paper introduces a novel model for predicting wind turbine power output within a wind farm at a high temporal resolution of 30 seconds. The wind farm is represented as a graph, with Graph Neural Networks (GNNs) used to aggregate selected input features from neighboring turbines. A temporal component is ~~added~~ introduced by feeding a ~~timeseries~~ time series of input features into the graph ~~and utilizing a hybrid GNN-LSTM model architecture. This approach sequentially extracts temporal features with the LSTM component and spatial features with the GNN component~~, processed through a Long Short-Term Memory (LSTM) network before being passed to the GNN. Our model is integrated into a Normal Behavior Model (NBM) framework for analyzing power loss events in wind farms. The results show that both the Spatial and Spatio-Temporal GNN models outperform traditional data-driven power curve methods, ~~with~~ achieving reductions in Mean Absolute Error (MAE) of approximately 22.6% and 30.3%, respectively, and in Mean Absolute Percentage Error (MAPE) of around 20.7% and 30.5%. Notably, the Spatio-Temporal GNN ~~demonstrating superior performance due~~ demonstrates superior performance, attributed to its ability to effectively capture both spatial and temporal dynamics. Additionally, the model achieves remarkable agreement with SCADA-derived energy ratios across the full range of wind directions, with a weighted average error of 0.0373; an improvement of approximately 57.4% compared to the power curve binning method. This advantage is especially pronounced under waked conditions, where traditional methods such as the power curve and Multilayer Perceptron (MLP) models exhibit significantly higher error rates. Beyond power prediction, we illustrate the model's effectiveness in detecting and analyzing instances of reduced performance and its ability to identify various types of abnormal events beyond what is recorded in standard status logs. Compared to the power curve method, the Spatio-Temporal GNN reduces the rate of undetected power loss events from 12.6% to just 0.02%, demonstrating a substantial improvement in capturing abnormal events.

1 Introduction

Wind energy plays a vital role in addressing the escalating global energy demand while aligning with sustainability goals to combat climate change. As countries worldwide strive to reduce greenhouse gas emissions and transition towards renewable energy sources, wind energy emerges as a clean and abundant resource capable of meeting a significant portion of our energy needs (Yousefi et al., 2019). Offshore wind farms offer promising opportunities to harness stronger and more consistent wind speeds compared to onshore locations, thereby contributing to a more resilient and sustainable energy infrastructure

25 (International Energy Agency, 2023). As technology advances and costs continue to decline, offshore wind projects are becoming more economically viable and attractive investments for governments, developers, and energy consumers alike ~~. Moreover, offshore wind farms have the potential to generate large-scale energy outputs, making them crucial contributors to achieving renewable energy targets and reducing dependence on fossil fuels.~~ (European Commission, 2023).

Accurate estimation of potential power allows wind farm operators to optimize the operation of individual turbines and the
30 entire wind farm. By understanding the potential power each turbine could generate under different wind conditions, operators can make informed decisions about operation settings and maintenance schedules to maximize production while minimizing costs. As wind conditions and turbine performance can fluctuate significantly, developing precise power prediction models becomes essential to navigating these complexities.

Building on this need for accurate power forecasts, it is equally important to utilize these predictions to identify power loss
35 events - instances of reduced performance where the generated power falls below the potential output - within the wind farm. By comparing predicted energy output with actual performance, we can detect potential inefficiencies and operational issues.

In this context, this research addresses two key questions:

1. Can we leverage temporal ~~and spatial patterns within a~~ trends in wind flow, turbine positioning, and integrated measurements from different turbines within the wind farm to improve power prediction accuracy?

40 2. Can we detect power loss events by comparing precise power predictions with actual turbine performance?

Potential power prediction methods can be integrated into a broader performance monitoring framework through the implementation of normal behavior modeling (NBM). The essence of NBM lies in training models to recognize normal behavior, enabling the identification of abnormal observations based on low model support. For instance, in the context of a regression-based NBM model, training on a dataset representing purely normal behavior enables the detection of abnormalities
45 by identifying deviations from an expected residual of zero when presented with new test set observations.

Additionally, quantifying losses during power loss events enables stakeholders to assess the economic impact of turbine downtime or grid curtailments. High levels of curtailment can lead to increased wear and tear on turbines, reduced operational lifespan, and potential reliability issues (Robbelein et al., 2023). With curtailment strategies evolving to be more dynamic, the need to refine these strategies increases. By ~~understanding~~ quantifying the magnitude and frequency of ~~the~~ power losses
50 during curtailments, operators can ~~implement measures to minimize~~ identify patterns and evaluate the economic impact of curtailments. This data enables them to implement strategies to reduce curtailment frequency or duration and make informed decisions about investments in grid upgrades or energy storage solutions that could mitigate future losses.

Finally, the development of state-of-the-art wind farm control methods typically relies on models that capture the wake effect between wind turbines ~~at a high temporal resolution~~ (Verstraeten et al., 2021). Wake effects are highly non-linear and difficult
55 to capture, and traditionally modeled using physics-based wake models that ~~are often~~ can be calibrated using real-world data (Van Binsbergen et al., 2024a, b, c). In this work, we propose ~~an alternative to physics-based models for capturing wake effects and predicting~~ a model to predict the potential power of wind turbines operating in wakes, specifically designed for wind farm control applications.

~~The main contributions of this paper are the following:~~

~~We developed a model to predict the potential power of all turbines within a wind farm at~~ The model utilizes integrated measurements from the supervisory control and data acquisition (SCADA) system of multiple turbines at a high temporal resolution (~~of~~ 30 ~~second) by modeling the wind farm as a graph and using Graph Neural Networks to aggregate selected input features from neighboring turbines and predict the power output across the wind farm. A temporal component was added by feeding a timeseries of the input features to the graph and using a hybrid GNN-LSTM modelarchitecture to sequentially extract the temporal features using the LSTM and the spatial features with the GNN component. We integrate our power prediction model into a NBM framework and showcase the practical implications of turning power prediction methods into a methodology for analyzing power loss events.~~ seconds.

The remainder of this paper is structured as follows: In Sect. 2, we position our research within the context of existing literature. Sect. 3 details our approach, including the data processing pipeline, the GNN framework, and the training and hyperparameter tuning process of our models. Sect. 4 presents a comprehensive analysis of the model's performance, including its predictive accuracy and ability to detect abnormal events, compared against baseline methods. In Sect. 5, the most important results are further discussed, and finally, in Sect. 6, we summarize the key findings of our study and outline potential directions for further research.

2   Related Work

An example of a normal behavior model ~~, consisting of~~ for performance analysis, based on an artificial neural network ~~variant~~ trained on an abnormality-filtered ~~dataset~~ SCADA dataset for power prediction, is shown by Lyons and Gocmen (2021). The authors ~~show its effectiveness in accomplishing the~~ demonstrate the effectiveness of the developed NBM for power performance analysis ~~objective, and~~ by qualitatively discussing instances of over- and underperformance ~~captured by the developed NBM network are presented and discussed~~identified by the model. However, the NBM seemed to struggle to model power production at higher wind speeds and expressed a tendency to underestimate the wake effect at play within the farm. In another study by Bilendo et al. (2022), a different normal behavior model is introduced for condition monitoring of wind turbines, leveraging a heterogeneous stacked regressor (HET-SR) algorithm. This algorithm learns from optimal power curve data to serve as a predictive model within their NBM framework. While ~~their findings provide evidence on the effectiveness of the proposed method for accurate fault detection~~a qualitative analysis of a variety of faults that can be detected by the model is shown, employing more advanced prediction models that account for additional variables beyond wind speed could offer opportunities for a more comprehensive performance analysis beyond general fault detection.

Traditionally, ~~wind power production is governed by the~~ the manufacturer's power curve, ~~a simple relationship that gives~~ which describes the theoretical behavior of a wind turbine in constant and low turbulence wind conditions, is used to predict the expected power output at different wind speeds. However, ~~it cannot provide a precise forecast of wind power as it fails to accurately model the non-linear~~ this method fails to account for the specific atmospheric conditions in a particular wind farm. Therefore, most wind farm operators use the method of binning (International Electrotechnical Commission (IEC), 2017) to

estimate the power curve based on measurement data from their own wind turbines. With this method, the range of measured wind speeds is partitioned into separate bins of 0.5 m/s, and the power response is calculated by averaging the power data falling in each bin. While this method is easy to implement and can capture the primary nonlinear relationship between wind

95   speed and power output, especially it does not account for inter-turbine interactions and varying environmental conditions such as turbulence intensity.

These limitations are particularly important in large wind farms, where wake losses and other complex dynamics become more prevalent. Modeling these wake losses and the flow patterns within the farm has been an active area of research and various approaches have been devised to predict expected power in wake-affected wind farms. These methods can be categorized as

100   either physics-based or data-driven (or a combination of the two).

Physics-based models aim to model the wind and wake flows within the wind farm based on prior knowledge about the physical behavior of the system. These models range from low to high fidelity, depending on the amount of detail they capture. Low-fidelity models are relatively fast, but neglect lots of details in modeling the wind flow. An example is the FLOw Redirection and Induction in Steady State (FLORIS) model developed at NREL, a control-focused wind farm simulation software

105   incorporating steady-state engineering wake models into a Python framework (NREL, 2024). On the other hand, high-fidelity models describe the flow in detail based on the 3D Navier-Stokes equations and use large-eddy simulations (LES) to accurately resolve the turbulent flow structures, but they are limited by their high computational cost. Examples of high-fidelity models are SOWFA (NREL, 2012) and PALM (Raasch and Schröter, 2001).

Data-driven models construct relationships between the inputs and outputs based on statistical or machine learning mod-

110   els, without prior knowledge about the physical behavior of the process. Recent advancements in deep learning and big data resulted in a growing interest towards wind farm power prediction using deep learning methods. For instance, Lin and Liu (2020) provide a comprehensive overview of prior studies employing deep learning methods and introduces a predictive model using deep learning in conjunction with high-frequency SCADA data. Similarly, Lyons and Gocmen (2021) delve into power prediction tasks using deep learning, utilizing high-frequency SCADA data and integrating local information from neighboring

115   turbines to enhance predictive accuracy. Moreover, the significance of spatiotemporal factors influencing wind power generation is addressed in Zhang et al. (2021) and Yin et al. (2021), employing specialized model architectures such as Convolutional Neural Networks (CNNs) and Long Short- Term Memory networks (LSTMs) to extract spatial and temporal feature information. Finally, Daenens et al. (2024) developed a turbine-level power prediction model, incorporating high-frequency SCADA data from neighboring turbines into a prediction model for potential power based on a hybrid CNN-LSTM model architecture,

120   by organizing the input data in a grid structure based on the wind farm layout.

A different approach to represent the spatial correlations between wind turbines in a wind farm is the use of graphs. Graphs can be used to represent complex data that is not necessarily ordered inherently structured in a grid-like structure, but instead can be modeled by arbitrary graphs consisting of nodes and edges. With this approach , manner, allowing for arbitrary connections between nodes through edges. This approach has proven beneficial for wind farms in previous studies.

125   For example Verstraeten et al. (2021) utilized coordination graphs to develop scalable wind farm control strategies, effectively

decomposing large-scale multi-agent optimization problems. In another study, Hammer et al. (2023) represented the wind farm as a graph and employed an extreme gradient boosting model to predict the wake interaction losses between turbine pairs.

Graph Neural Networks (GNNs) ~~can be employed for the prediction of wind farm power output.~~ (Scarselli et al., 2009) have emerged as powerful tools for learning on graph-structured data. GNNs leverage the structure of the graph to learn rich node representations by iteratively aggregating information from neighboring nodes. They are particularly well-suited for tasks that involve node-level predictions, such as power prediction in wind farms.

For example, Bleeg (2020) introduced GNNs as surrogate models for steady-state Reynolds-Averaged Navier-Stokes (RANS) simulations, providing a more computationally efficient alternative to traditional RANS models while maintaining reasonable accuracy. More complex GNNs are explored in the work of Park and Park (2019) and Bentsen et al. (2022), where the former proposed a physics-induced GNN (PGNN) and the latter an attention-based GNN for the power prediction of individual wind turbines in a wind farm. Both these methods use synthetic wind farm data simulated using the FLORIS model.

In a different application, de N Santos et al. (2024) presented a GNN framework for layout-agnostic modeling of fatigue load effects. Using graph representations derived from PyWake (Pedersen et al., 2023) simulations, with inflow conditions as inputs and geometrical properties encoded on edges, the framework models wind farm dynamics across various layouts and conditions.

Combining spatial and temporal aspects of wind farm dynamics, Yu et al. (2020) developed the Superposition Graph Neural Network, a spatio-temporal model that processes a timeseries of graphs to predict the power output of the wind turbines in a wind farm. This model was trained on data for four offshore wind farms sampled at 10-minute intervals and showcases the capability of GNNs to capture the evolution of wind flow through the farm and its impact on power generation.

3 Methodology

The data-driven methodology proposed in this work leverages high-resolution SCADA data to predict the potential power for each wind turbine within a wind farm. In this section, we ~~will~~ describe the key aspects of our modeling framework and we ~~will~~ show the design choices that were made to develop a unified framework that can be applied to any wind farm. The data as well as the preprocessing pipeline are described in detail in Sect. 3.1~~, and the~~. In Sect. 3.2, the graph representation of the wind farm is introduced. The different model architectures and characteristics are elaborated on in Sect. 3.3~~.~~, and finally, the hyperparameter tuning process is explained in Sect. 3.4.

3.1 Data collection and preprocessing

We considered a wind farm in the Dutch-Belgian offshore zone for this study, consisting of ~~over~~ more than 40 turbines with >8MW rated power. Specifically, the signals obtained by the ~~supervisory control and data acquisition (SCADA)~~ SCADA system were used, ~~spanning~~ covering a two-year period from 2021 to 2022. The SCADA system records real-time information at a resolution of 1 Hz collected from sensors and control systems for monitoring and optimizing turbine performance and overall wind farm operations and is widely used in the offshore wind industry. As most contemporary wind farms have ~~a~~

~~SCADA system, a SCADA-based methodology is easily adaptable~~ the used input signals readily available (e.g., through their SCADA system), the model is easily accessible and applicable to other wind farms.

160 From the SCADA system, the data signals shown in table 1 were retained and used in our analysis. The input features to the prediction model are `wind_speed`, `wind_direction_sin`, `wind_direction_cos`, and `turbulence_intensity`. ~~Wind speed is directly taken from the~~ They were chosen because they directly capture the key physical factors influencing wind farm power generation, such as wind resource characteristics, directional influences, and flow variability. Other features were excluded due to data availability limitations, potential redundancy, and a focus on simplifying the model while retaining

165 interpretability and accuracy. Wind speed was taken directly from SCADA data and represents the wind speed at turbine hub height as measured by the anemometer located on the nacelle. ~~Wind directionin our model is represented using~~ The absolute wind direction, measured by a wind vane mounted on the nacelle, was represented by sine and cosine transformations to account for its circular nature. This approach ensures a smooth representation of wind direction throughout its entire range and prevents issues where a model might incorrectly interpret a large difference between wind directions close to 0 and 360 degrees.

170 Turbulence intensity (TI) refers to the measure of fluctuations in wind speed over time, indicating the variability or instability of the airflow. The turbulence intensity ~~is~~ was calculated using Eq. (1).

$$TI = \frac{\sigma_{ws}}{\mu_{ws}} \tag{1}$$

where σ_{ws} is the standard deviation of wind speed (ws), and μ_{ws} is the mean wind speed, both computed over a 10-minute interval centered around the 1-second data point. In practical terms, this method requires wind speed data from 5 minutes before

175 and 5 minutes after each point to compute the turbulence intensity. As a result, any real-time application of this calculation would inherently introduce a 5-minute lag, since the full 10-minute window is not complete until 5 minutes after the moment being analyzed. However, in our case, this method is used for historical data analysis rather than real-time predictions, so this lag does not impact the accuracy or usability of our power predictions.

The target of our prediction model is the `active_power` signal extracted from the SCADA data. Additionally, we ~~include~~

180 included two more SCADA signals: `rotor_speed` and `pitch_angle`, which ~~are~~ were utilized for filtering normal operational behavior across our training, validation, and test datasets.

The steps to prepare the raw data for the training, validation, and testing of our prediction model are outlined in Table 2.

First, the raw SCADA data with a sampling frequency of 1 Hz was resampled to 30-second averages as a trade-off between high temporal resolution, acceptable noise levels, and manageable computational costs. This approach smooths out rapid

185 fluctuations while preserving key temporal dynamics. Next, the data was annotated with control conditions ~~to be used when filtering for~~ for filtering normal behavior. As mentioned previously, to create ~~an NBMthat can accurately detect~~ a Normal Behavior Model (NBM) capable of accurately detecting abnormal behavior and ~~predict~~ predicting potential power output under any conditions, the model must be trained ~~on data representing~~ exclusively on data that represents normal system behavior. ~~The filtration method employed in thisstudy is~~

Table 1. Overview of Data Signals and Input Features

Signal Name	Description	Data source	Usage
wind_speed	The wind speed as measured by the anemometer located on the turbine nacelle [m/s]	SCADA	Model Input
wind_direction	The wind direction as measured by the wind vane located on the turbine nacelle [°]	SCADA	Calculation of input features
wind_direction_sin	The sine of the wind direction [-]	Calculated	Model Input
wind_direction_cos	The cosine of the wind direction [-]	Calculated	Model Input
turbulence_intensity	The turbulence intensity as calculated by formula 1 [-]	Calculated	Model Input
active_power	The active power produced by the wind turbine [kW]	SCADA	Model Target
rotor_speed	The rotational speed of the wind turbine blades [rpm]	SCADA	Data Filtering
pitch_angle	The angle ~~of attack of the wind turbine blades relative to the incoming wind~~ between the plane of rotation and the chord line of the blade [°]	SCADA	Data Filtering

190   To achieve this, we used a physics-based ~~and uses~~ filtration method, which relies on the properties of the power curve ~~based on the IEC standard~~ as per IEC standards (International Electrotechnical Commission (IEC), 2017) to annotate steady-state control conditions. Using this approach, ~~different control regimes can be identified, and outliers and abnormal behavior can be removed~~ we classified the data into different operating regions (i.e., torque control, pitch control), flagged data points falling outside these regions as abnormal, and removed them from the dataset. ~~Additionally, to leverage the~~ We removed the entire

195   timestamp across all turbines when any individual turbine exhibited abnormal operation. While this reduces the number of data points available for model training, we found that this approach yields the best results in creating a reliable NBM.

To leverage temporal patterns in the wind flow ~~through~~ throughout the wind farm, we incorporated lagged values of the input features ~~are incorporated~~ into the model. Specifically, a time series of input features between the prediction time t and $t - T$ (with T = 5 minutes) was included. Given the time-series nature of the data, it is ~~important~~ essential to maintain temporal

200   continuity ~~when lagged values are used as inputs. Consequently,~~ , so a final filtering step was implemented to ensure that ~~the lagged values utilized~~ lagged values used by the model are consecutive and ~~there are no discontinuities due to~~ free from discontinuities caused by missing data. To ensure continuity in the time series, we required that no timestamps were missing across all turbines for the current and lagged timestamps.

Finally, the resulting dataset ~~is~~ was partitioned into distinct subsets for model training, validation, and testing. The first year

205   of data was used as training data and was used to fit the model parameters. The second year of data was split into equal parts as the validation and test set. Validation data is utilized to tune hyperparameters and prevent overfitting, and testing data is used to assess the model's performance on unseen data, ensuring its generalizability.

Table 2. Preprocessing Steps and Remaining Data Points

Preprocessing Step	Initial Data Points	Remaining Data Points	Final Data Split
Raw SCADA Data	39,118,140	-	-
Resampling to 30s Averages	-	1,303,938	-
Retaining Normal Behavior	-	262,614	-
Consecutive Timestamps	-	164,124	-
Train, Validation, Test Split			
Training Set	-	-	116,410
Validation Set	-	-	23,857
Test Set	-	-	23,857

3.2 Graph representation of the wind farm

With the datasets for training, validation, and testing of the prediction model defined, they ~~have been~~ were transformed into a format suited for our model. Specifically, the wind farm ~~is~~ was represented as a graph, and the selected input features from the SCADA data ~~are~~ were converted into graph-structured data.

A graph G is usually defined as a tuple of two sets $G = (V, E)$, where $V = \{v_1, v_2, \ldots, v_N\}$ and $E \subseteq V \times V$ are the sets of nodes and edges. To model a wind farm as a graph using this structure, each turbine in the wind farm is represented as a node, and edges connect neighboring turbines. An edge $e_{ij} = (v_i, v_j) \in E$ between nodes v_i and v_j exists if these nodes have line-of-sight visibility and there are no obstacles (other turbines) blocking the direct view. This visibility check is performed using Algorithm 1.

Algorithm 1 Visibility Check

1: **Input:** node1, node2, obstacles, tolerance

2: **Output:** True if node1 and node2 are in line-of-sight, False otherwise

3: Initialize a flag, *visible*, to True.

4: **for** each obstacle **in** obstacles **do**

5:     **if** obstacle $\neq$ node1 **and** obstacle $\neq$ node2 **then**

6:         Calculate the distance from node1 to the obstacle, $d1$.

7:         Calculate the distance from node2 to the obstacle, $d2$.

8:         Calculate the direct distance between node1 and node2, d_direct.

9:         **if** $|d1 + d2 - d\_direct| <$ tolerance (250m) **then**

10:             Set *visible* to False and **break** out of the loop.

11:         **end if**

12:     **end if**

13: **end for**

14: **return** *visible*, indicating whether node1 and node2 have line-of-sight visibility.

Each node $v \in V$ can be associated with a vector of features $x_v \in X$, comprising the input features as defined previously. Similarly, a vector of edge attributes can be defined for each edge $e \in E$, containing a set of geometric features, ~~such as~~ specifically the length and direction of each edge. Similar to our representation of wind direction, we represented edge direction using its sine and cosine components to account for its circular nature.

3.3 GNN Models

~~Graph Neural Networks (GNNs) (Scarselli et al., 2009) have emerged as powerful tools for learning on graph-structured data. They are particularly well-suited for tasks that involve node-level predictions, such as power prediction in wind farms. GNNs leverage the structure of the graph to learn rich node representations by iteratively aggregating information from neighboring nodes.~~

At the heart of GNNs is the concept of message passing, a process through which nodes in a graph communicate with their neighbors to update their representations (Gilmer et al., 2017). This iterative process can be described in two main steps: message aggregation and node update.

1. **Message Aggregation**: Each node aggregates messages from its neighbors. The nature of this aggregation can vary, but common methods include summation, mean, and max pooling of neighbor features.

2. **Node Update**: After aggregation, each node updates its own feature based on the aggregated message and its previous state. This update is typically performed using a neural network, such as a multilayer perceptron (MLP).

~~Mathematically, the k-th layer of a GNN can be expressed as:~~

$$\mathbf{h}_i^{(k)} = \text{UPDATE}\left(\mathbf{h}_i^{(k-1)}, \text{AGGREGATE}\left(\{\mathbf{h}_j^{(k-1)} : j \in \mathcal{N}(i)\}\right)\right)$$

~~where UPDATE and AGGREGATE are arbitrary differentiable functions (i.e., neural networks), $\mathbf{h}_i^{(k)}$ is the representation of node i at the k-th layer, and $\mathcal{N}(i)$ denotes the set of neighbors of node i.~~

For the task of power prediction, we developed two models: a Spatial GNN and a Spatio-Temporal GNN. The Spatial GNN considers the node features only at the current timestep, whereas the Spatio-Temporal GNN also incorporates lagged values of the input features. Both models utilize the GENeralized Graph Convolution (GENConv) model proposed by Li et al. (2020) due to its ability to incorporate edge features into the message-passing process. Its message construction at layer k can be expressed as:

$$\mathbf{h}_i^{k+1} = \text{MLP}\left(\mathbf{h}_i^k + \text{AGG}\left(\{\text{ReLU}\left(\mathbf{h}_j^k + \mathbf{e}_{ji}\right) + \epsilon : j \in \mathcal{N}(i)\}\right)\right)$$

where $\mathbf{h}_i^{(k)}$ is the representation of node i at the k-th layer, $\mathbf{e}_{ji}$ is the feature vector associated with the edge connecting nodes i and j, ϵ is a small positive constant set as 10^{-7}, and $\mathcal{N}(i)$ denotes the set of neighbors of node i. Finally, AGG represents the aggregation function, with softmax aggregation used in this work. ~~This model employs generalized message aggregators, pre-activation residual connections, and message normalization layers to ensure robustness in deep networks. Additionally, it can integrate edge features, which we believe contain crucial information about intra-farm wake behavior.~~

The Spatial GNN model consists of a node and edge feature encoder, followed by a series of ~~graph convolutional layers(GCN layers)~~ message-passing layers, and concludes with a decoder: a dense layer with sigmoid activation to normalize the predictions

250 between 0 and 1. Both the node and edge encoders are two-layer multi-layer perceptrons (MLPs)~~with ReLU activations. The GCN layers are~~. The message-passing layers were implemented using PyTorch Geometric (Fey and Lenssen, 2019) ~~, and comprise DeepGCNLayer objects that incorporate pre-activation residual connections. Each DeepGCNLayer employs layer normalization, ReLU activation, and the GENeralized Graph Convolution (GENConv) as the convolution operator~~ and were constructed according to the GENConv model by Li et al. (2020).

255 The Spatio-Temporal GNN model first processes the time series of node features with an LSTM network. Similar to the Spatial GNN, this block serves as the node feature encoder; however, in this case, the output of the LSTM network is passed through the GNN model.

An overview of the methodology for power prediction is shown in Figure 1.

3.4 Hyperparameter tuning

260 Hyperparameter tuning is crucial for optimizing a machine learning model's performance and ensuring it generalizes well to unseen data. In this study, we used the hyperparameter optimization framework Optuna (Akiba et al., 2019). Optuna allows users to dynamically construct the parameter search space, benefits from efficient sampling and pruning algorithms, and is easy to set up.

Table 3 lists the hyperparameters that were tuned, along with their ~~suggested ranges. After defining these ranges,~~ proposed

265 ranges, as well as the optimized value for both models. Optuna automatically searched for the optimal hyperparameters ~~based on~~ by minimizing the objective function. ~~Within the objective function, models were trained with the selected hyperparameters and scored based on mean squared error~~, defined as the Mean Squared Error (MSE) ~~.~~

between the model's predictions and the target values. Hyperparameter combinations that resulted in the lowest MSE for the validation dataset were retained and used to build the final models.

Table 3. Hyperparameters~~and Their Considered~~, Proposed Ranges, and Optimal Values for Each Model

Hyperparameter	~~Suggested Range~~Description	Proposed Range	Best Hyperparameters	
			Spatial GNN	Spatio-Temporal GNN
Encoding Channels	Number of units in the encoder MLP	{8, 16, 32, 64}	64	16
Hidden Channels	Number of units in the MLP for message passing (hidden state)	{8, 16, 32, 64}	64	64
Number of Layers	Number of message passing layers	{1, 2, 4, 8}	8	8
Number of LSTM Layers	Number of LSTM layers in the Spatio-Temporal GNN encoder	{1, 2, 4, 8}	n/a	1
GNN Dropout	Dropout percentage in message passing layers	[0, 0.1]	0.0367	0.0971
Encoding Dropout	Dropout percentage in the encoder MLP	[0, 0.1]	0.0236	0.0199
Number of Epochs	Total number of training epochs	[20, 100]	100	100
Batch Size	Number of samples per batch during training	{64, 128, 256, 512}	64	256
Learning Rate (lr)	Step size for optimization (log scale)	[1e-5, 0.5] ~~(log scale)~~	0.00217	0.0134
Learning Rate Decay	Factor by which the learning rate is reduced over time	[0.9, 1]	0.998	0.997
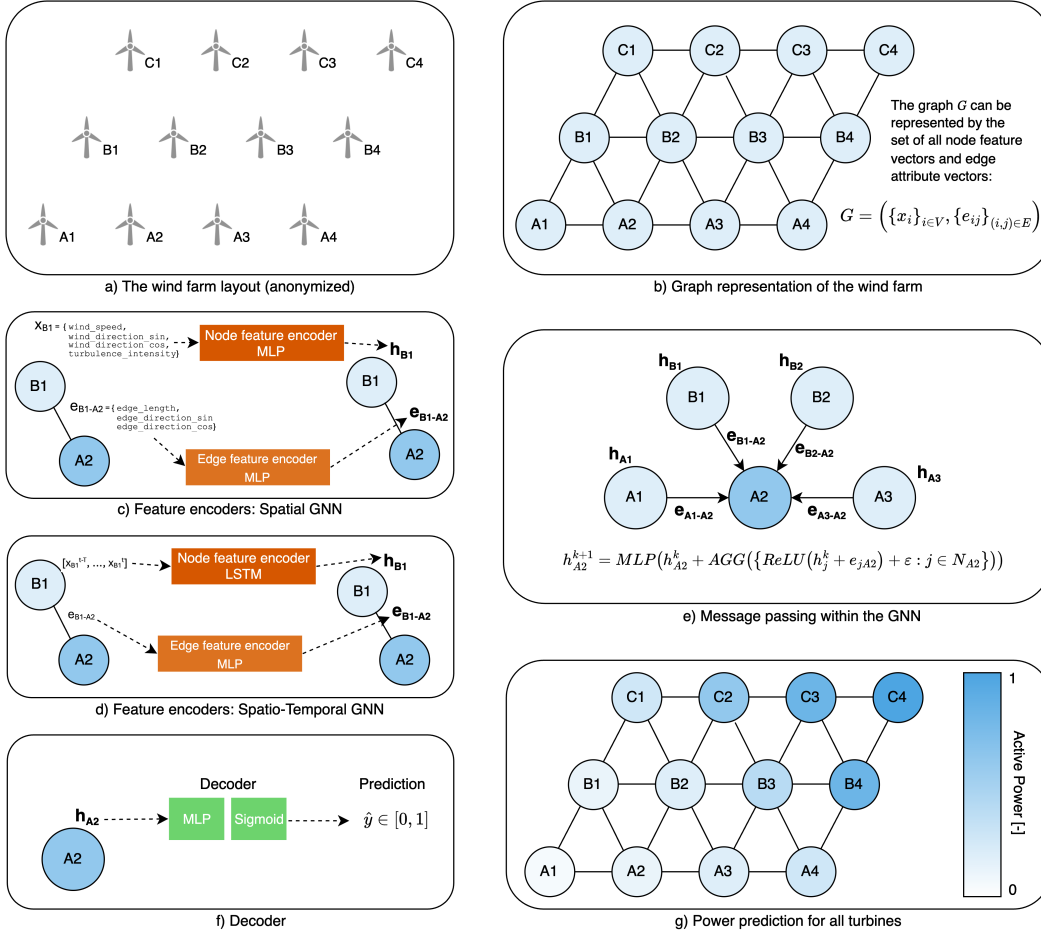
Figure 1. Overview of the proposed methodology for power prediction. The wind farm (a) is represented as a graph (b). The input features are encoded by the encoder blocks. In the Spatial GNN, these are MLPs (c); in the Spatio-Temporal GNN, the timeseries of node features is encoded using an LSTM network. The edge feature encoder remains the same (d). The GNN learns rich representations for each node in the graph using message passing (ee). Each node representation is decoded using an MLP with sigmoid activation, to normalize the predictions between 0 and 1 (f). Finally, a power prediction is made for each node within the graph (dg).

4   Results& Discussion

In this section, the performance of the different models concerning the specified objectives is discussed. As mentioned in Sect. 3.1, SCADA data of an offshore wind farm comprising over 40 turbines has been used to train and evaluate the Spatial and Spatio-Temporal GNNs described in Sect. 3.3.

4.1 Model performance during normal operation

275 In Table 4, the Mean Absolute Error (MAE) and the Mean Absolute Percentage Error (MAPE) are reported for each model. The models are compared with a ~~data-driven power curve method~~ Power Curve model based on the power curve binning method (International Electrotechnical Commission (IEC), 2017) as a baseline. ~~With the power curve method, the turbine's active power is estimated from the power curve, created by analyzing collected data on wind speed and the turbine's actual power production. This curve illustrates how the turbine responds to different wind speeds, enabling predictions of power generation~~

280 ~~for varying wind conditions within the observed range~~ Furthermore, we evaluated our models compared to a standard Multilayer Perceptron (MLP), a commonly used data-driven approach in wind power forecasting. The MLP configuration includes two hidden layers with 64 and 32 units, respectively. The input features for this MLP are `wind_speed` and `turbulence_intensity`; wind direction is excluded as it does not directly affect the power output of an individual turbine.

Table 4. Performance Metrics. The MAE values have been normalized by dividing them by the turbine's rated power for confidentiality reasons. The lowest errors are highlighted to showcase the model with the best predictive performance on each dataset.

		Power Curve	MLP	Spatial GNN	Spatio-Temporal GNN
MAE	Train	0.0476	0.0450	0.0287	**0.0258**
	Validation	0.0481	0.0473	0.0348	**0.0311**
	Test	0.0478	0.0468	0.0370	**0.0333**
MAPE	Train	12.717 %	12.451 %	7.859%	**6.839 %**
	Validation	13.265 %	13.306 %	9.916%	**8.516 %**
	Test	12.442 %	12.503 %	9.872%	**8.647 %**

The results indicate that both the Spatial GNN and the Spatio-Temporal GNN models significantly outperform the ~~data-driven~~

285 ~~power curve method~~ power curve binning method and the MLP across all metrics and datasets. ~~The superior performance of the GNN models emphasizes the advantage of representing the wind farm as a graph. In this approach, each turbine not only considers its local measurements but also integrates information from its neighbors through message passing, enabling the model to capture complex spatial dependencies within the wind farm.~~

Specifically, the Spatio-Temporal GNN model exhibits the best performance, as evidenced by the lowest Mean Abso-

290 lute Error (MAE) and Mean Absolute Percentage Error (MAPE) values in the training, validation, and test datasets. ~~The Spatio-Temporal GNN's ability to incorporate lagged values of the input features provides it with an additional temporal dimension, allowing it to detect and leverage trends over time. This capability is particularly beneficial for capturing dynamic changes in wind conditions, such as fluctuations in wind speed and direction, or for smoothing out the inherent noise in high-frequency SCADA data.~~

295 Comparatively, while the Spatial GNN also outperforms the power curve ~~method~~ binning method and the MLP, it falls short of the Spatio-Temporal GNN. The Spatial GNN only utilizes input features at the time of prediction, missing out on the

temporal trends that the Spatio-Temporal GNN can exploit. Therefore, while spatial awareness, integrated measurements, and the ability to model interactions between turbines are crucial, the inclusion of temporal dynamics further enhances predictive accuracy.
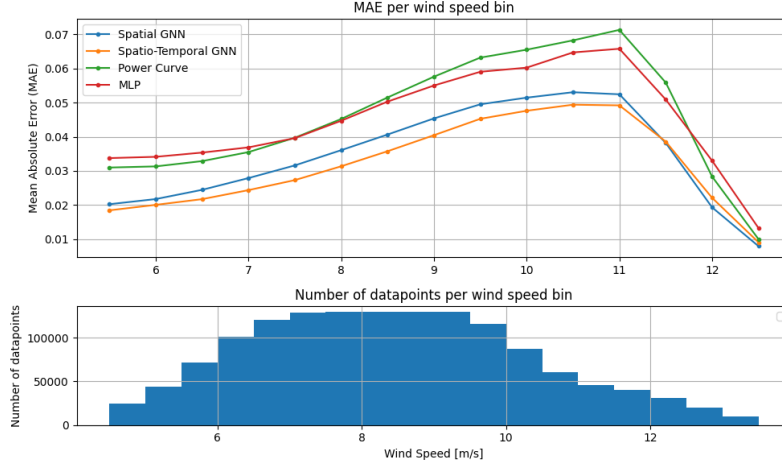
Figure 2. Mean Absolute Error (MAE) per wind speed bin for the two GNN models, the MLP and the power curve binning method. The plot at the bottom shows the number of data points per wind speed bin.

To provide a more granular analysis of our models' predictive performance, Figure 2 presents the Mean Absolute Errors (MAE) for different wind speed bins on the test dataset. While the previous analysis highlighted the overall performance of each model across the entire dataset, this plot shows their performance across varying wind speeds.

The analysis shows a slight increase in MAE at higher wind speeds across all models, with a subsequent decrease in errors near the turbine's rated wind speed. This trend aligns with expectations, as higher wind speeds typically result in higher power ~~outputs andconsequently~~ output and, consequently, larger absolute errors. The observed pattern follows the inherent uncertainties of the power curve, but the models still achieve significantly higher accuracy.

An essential step in model validation involves comparing the model's predictions with historical data, specifically focusing on whether the individual turbine wake losses are accurately predicted. This can be achieved by analyzing the energy ratio between a test turbine and a reference turbine for each wind direction bin (Fleming et al., 2019; Doekemeijer et al., 2022). By examining the predicted energy ratio for each wind direction, we can gain deeper insights into how the model evaluates the impact of wake interactions within the wind farm. In an unbiased model, the energy ratio curves should align closely with those observed in the SCADA data.

The energy ratio R_{Energy} is defined in Eq. (2) and represents the ratio of the sum of all ~~the~~ power measurements between a test turbine and a reference turbine, computed for each wind direction bin. This sum of power measurements over a fixed period is equivalent to the total energy produced during that time. Because the ratio is calculated over the same time period for both the test and reference turbines, the time factor cancels out. In this equation, $P_{i,\theta}^{\text{Turbine 1}}$ and $P_{i,\theta}^{\text{Turbine 2}}$ are the observed
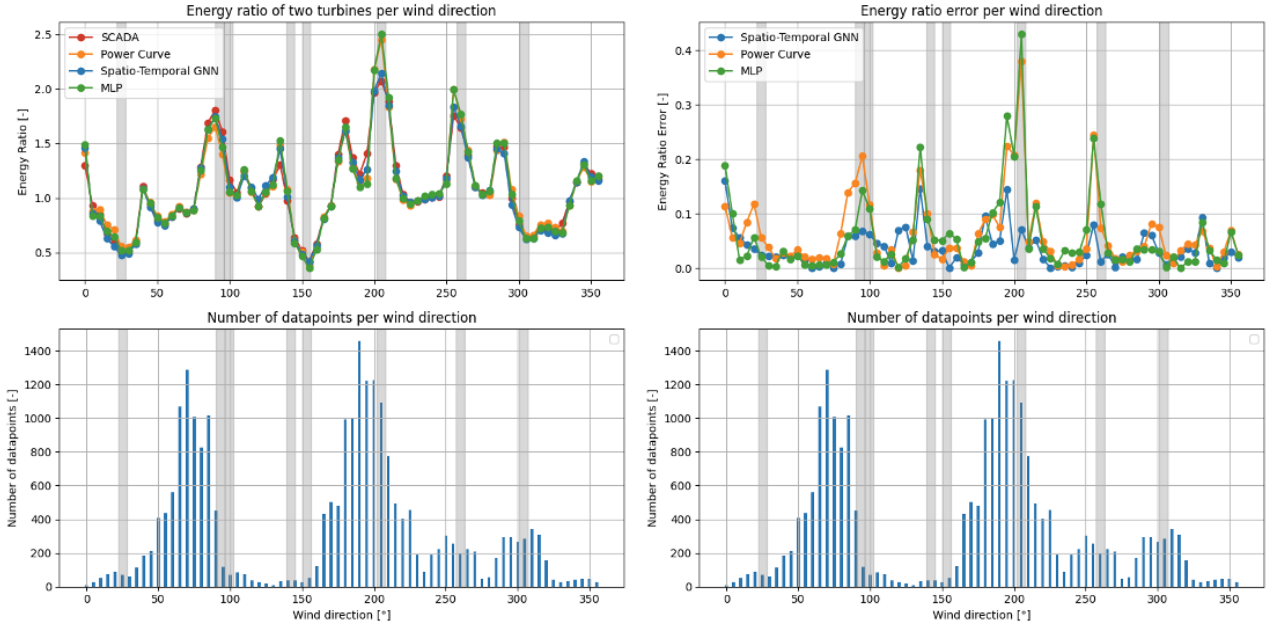
Figure 3. Energy ratio ~~plot (left) and Energy ratio error plot (right)~~ between two turbines ~~per~~ for each wind direction as predicted by the Spatio-Temporal GNN, the power curve binning method, the MLP, and ~~from~~ the SCADA data. The plot at the bottom shows the number of data points ~~per~~ in each wind direction bin. Grey-shaded areas indicate waked conditions.

powers of point i in a given wind direction bin for the two test turbines, and N is the number of points in this wind direction bin.

$$R_{\text{Energy}}(\theta) = \frac{\sum_{i=1}^{N} P_{i,\theta}^{\text{Turbine 1}}}{\sum_{i=1}^{N} P_{i,\theta}^{\text{Turbine 2}}} \qquad (2)$$

320   Figure 3 illustrates the energy ratio curves for two test turbines within the wind farm. One turbine is positioned in the free flow relative to the dominant wind direction, while the other is located in a subsequent row behind the first turbine, thus experiencing the wake effect generated by the upstream turbine ~~.~~ (Turbine 1 and Turbine 2, respectively, as indicated in Eq. (2)). These curves compare the energy ratio, per wind direction bin, as predicted by the ~~hybrid model~~ Spatio-Temporal GNN, the power curve binning method, and the MLP with the energy ratios derived from the SCADA data.

325   Our Spatio-Temporal GNN model demonstrates remarkable agreement with the energy ratios from the SCADA data across the full range of wind directions. Discrepancies are minor and primarily occur in wind directions that are underrepresented in the dataset. This suggests that the model's predictive accuracy is robust, even though it may be slightly less accurate in areas with sparse data. To quantify this, we calculated the wind-direction-frequency-weighted average of the differences between the predictions and the actual energy ratios. The power curve method has an average error of 0.0875, and the MLP

achieves an average error of 0.0828. In contrast, the Spatio-Temporal GNN attains a significantly lower average error of 0.0373, representing a relative improvement of approximately 57.4% over the power curve binning method.

Furthermore, the energy ratios align with intuitive expectations. For instance, in wind directions between 220° and 245°, both turbines experience undisturbed wind inflow, leading to similar power outputs. As the wind direction shifts towards 250° to 260°, one of the test turbines becomes obstructed by another turbine, resulting in reduced power production. Consequently, the energy ratio increases, reflecting the higher power output of the turbine with unobstructed inflow compared to the blocked turbine.

~~This consistency between the model predictions and the intuitive understanding of turbine interactions under different wind conditionsshows the Spatio-Temporal GNN model's effectiveness in capturing the complex dynamics of wake interactions within the wind farm. The minor discrepancies in less common wind directions highlight areas for potential improvement but do not detract significantly from the overall performance of the model~~The energy ratio error plot reveals that the Spatio-Temporal GNN significantly outperforms both the power curve binning method and the MLP, particularly under waked conditions. This highlights the strength of our GNN in accurately predicting potential power in waked flows, where conventional methods like the power curve binning and MLP models fall short.

4.2 ~~Predictability~~ Quantitative detection of abnormal events

Having quantified our models~~'~~' accuracy on the test set containing normal behavior, the next step is to ~~assess how the models handle abnormal behavior. As discussed in the introduction, our objective is to integrate our power prediction model into a normal behavior modeling~~ evaluate their performance under abnormal conditions. While our primary objective is not to develop a Normal Behavior Modeling (NBM) framework~~. With~~, we aim to validate the robustness of our power prediction model ~~trained to predict normal behavior, we propose a methodology to~~ in scenarios where no ground truth is available, such as turbine shutdowns or curtailments. To achieve this, we employ a commonly used anomaly detection methodology: a threshold-based approach that identifies anomalies by analyzing the residuals between the predicted and actual values. This allows us to assess the model's ability to detect abnormal observations based on ~~low model support~~deviations from expected behavior.

In our methodology, ~~low model support is defined as a deviation of the predicted power from the actual power that exceeds three~~ anomalies are flagged when the deviation between the predicted and actual power exceeds two times the standard deviation of the produced power for a given wind speed ~~. The~~ bin. These standard deviations are derived from a data-driven power curve and calculated for wind speed bins of 0.5 m/s width.

At low wind speeds, the standard deviation of the produced power is relatively low ~~because the potential power output is low and more predictable. As the~~ due to the reduced influence of aerodynamic and mechanical complexities; the turbine operates in a more linear torque control regime, where power production closely follows a predictable cubic relationship with wind speed. As wind speed increases, ~~the uncertainty of the potential power output also increases, leading to~~ turbulence and wake interactions introduce greater variability, leading to higher uncertainty in power output and, consequently, a higher standard deviation. ~~Approaching the rated wind speed~~ , However, as the wind speed approaches the rated value, the turbine enters pitch

control mode, where power output stabilizes at the rated capacity, reducing variability and lowering the ~~potential power output~~

365 ~~converges to the rated power, reducing uncertainty and hence decreasing the~~ standard deviation again.

This approach means that our NBM methodology tolerates higher errors in regions where there is more uncertainty about the potential power and enforces stricter error thresholds in regions where the potential power is more certain. This adaptive error tolerance is crucial for accurately identifying abnormal behavior without generating excessive false positives, particularly in regions where power output uncertainty is naturally higher. Thus, the proposed NBM methodology ensures more precise

370 and context-aware detection of anomalies in wind turbine performance.

To validate the NBM methodology, we applied our ~~model~~ Spatio-Temporal GNN to detect anomalous observations ~~within a~~ ~~dataset covering a two-month period. The model's predictions were subsequently~~ in a dataset spanning two months. Similarly, the NBM methodology was applied to the same dataset using the power curve binning method as a predictor. Predictions from both models were then compared against curtailments, shutdowns, and warnings documented in the status logs for the same

375 ~~timeframe~~period. The results are summarized in Figure 4, which presents a confusion matrix comparing the events recorded in the status logs with the anomalies detected by the NBM methodology. In this matrix, True Positives (TP) represent correctly predicted abnormal events; True Negatives (TN) denote correctly predicted normal observations; False Positives (FP) indicate events flagged as abnormal by the model, though no corresponding power loss event was recorded in the status logs; and False Negatives (FN) are abnormal events that the model failed to detect.
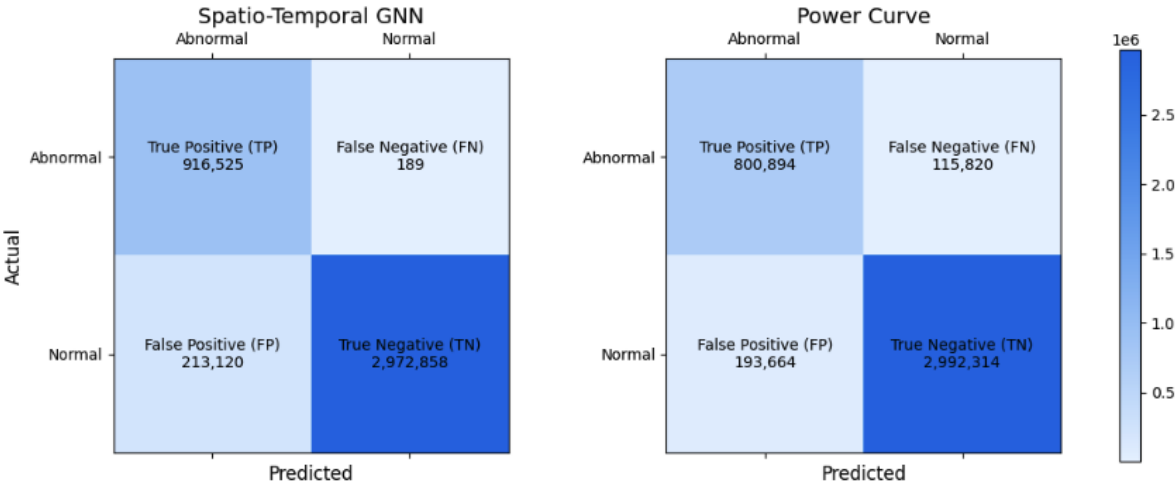
Figure 4. Confusion Matrix for NBM Methodology using the Spatio-Temporal GNN and the Power Curve model. Each entry represents the number of correctly predicted timestamps during the evaluation period.

380 ~~The majority of the~~ With the Spatio-Temporal GNN, the majority of data points in ~~this table fall under~~ the confusion matrix fall into the categories of True Positives or True Negatives, ~~which are~~ representing the desired outcomes of the methodology. Notably, the ~~low number of False Negatives indicates that nearly all known abnormal events were successfully detected by~~

~~the model~~model achieves a very low False Negative rate, with only 0.02% of power loss events going undetected. The few False Negatives that ~~remain are turbine shutdowns in~~ do occur correspond to turbine shutdowns under extremely low wind ~~scenarios~~conditions, where the potential power is ~~predicted to be almost zero. In these cases,~~ nearly zero, and the prediction error does not exceed the threshold for ~~identifying~~ abnormal behavior. The remaining data points are ~~classified as~~ False Positives, representing instances where the model ~~identified anomalies without confirming~~ flagged anomalies without corresponding evidence in the status logs. In contrast, the power curve binning method exhibits a significantly higher False Negative rate, leaving approximately 12.6% of power loss events undetected. While this method yields fewer False Positives, as discussed in Sect. 4.4, the absence of an exhaustive record of all power loss events prevents drawing meaningful conclusions from this metric.
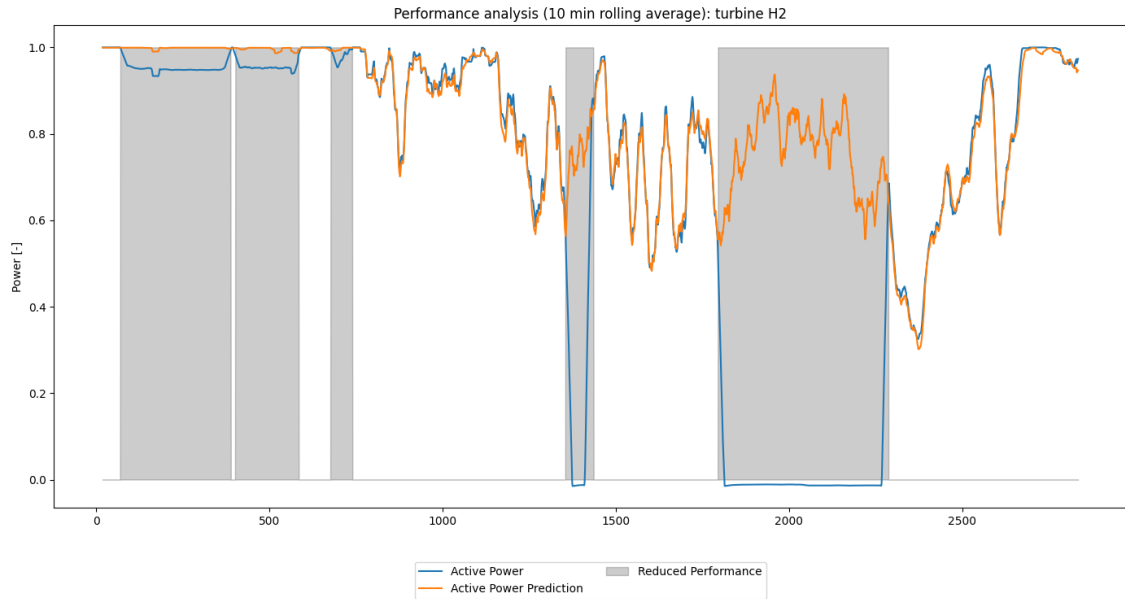
Figure 5. Performance analysis case study. A time series of the turbine's active power and the Spatio-Temporal GNN's prediction, smoothed using a 10-minute rolling average. The grey zones indicate abnormal performance events and are annotated using the NBM.

4.3 Analysis of a known power loss event

To further demonstrate the efficacy of our methodology in detecting power loss events, we conducted a detailed case study. Figure 5 illustrates a 24-hour period during which the active power output of a wind turbine, along with the corresponding predictions by the Spatio-Temporal GNN model, were analyzed. This period was chosen arbitrarily because it includes known loss events, making it suitable for demonstrating the model's ability to detect such anomalies. During this period, five loss events were detected, as indicated by the ~~grey shaded~~ grey-shaded areas. Since these events were confirmed by the status logs, they were categorized as True Positives.

The first three power loss events exhibit similar characteristics, as do the last two. In the first three events, the turbine's active power remains around 95% of its rated capacity, whereas the prediction model suggests that the turbine should be operating at full capacity. These power loss events are briefly interrupted by periods where the active power returns to the rated level. This pattern is consistent with a known curtailment strategy employed by the wind farm, where high wind speeds generate more power than can be transmitted to the grid, necessitating partial curtailment of the farm's output.
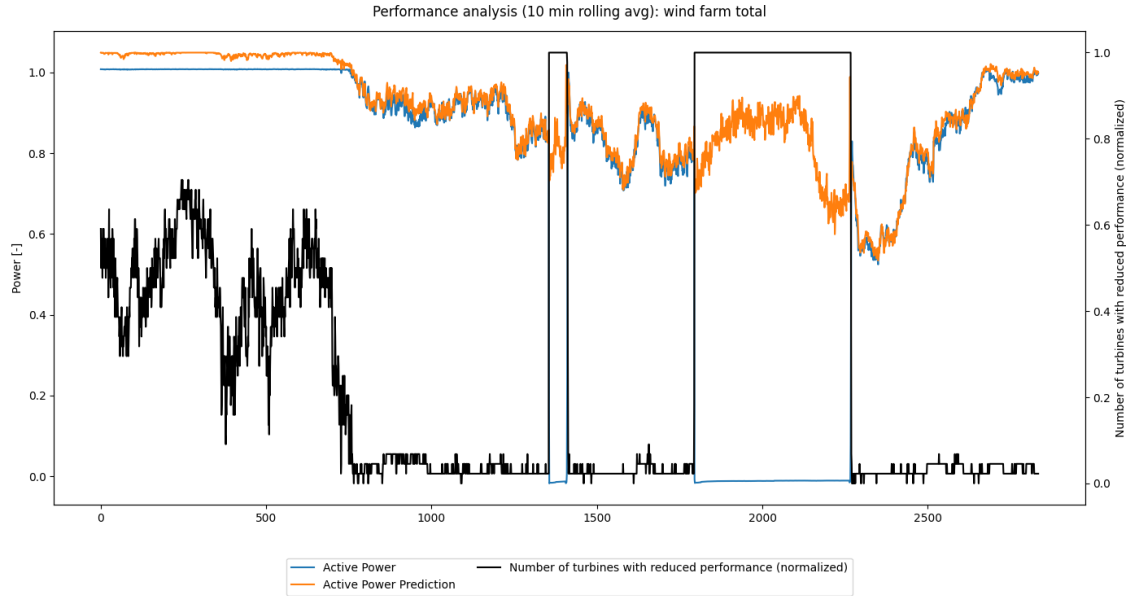
Figure 6. Performance analysis case study. A time series of the wind farm's total active power and the Spatio-Temporal GNN's prediction, smoothed using a 10-minute rolling average. The black line denotes the number of turbines with reduced performance, as detected by the NBM.

Figure 6 provides a broader perspective by aggregating data from individual turbines to represent the total active power output and the Spatio-Temporal GNN model's predictions for the entire wind farm. Additionally, the figure displays the number of turbines experiencing power loss events as detected by the NBM method.

During the period encompassing the first three power loss events, the total active power of the farm remains at its maximum capacity, while the predicted power exceeds this level. During this time, the farm controller actively manages the total output by curtailing certain turbines to avoid exceeding grid capacity limits.

The last two events follow a different pattern, where the turbine shuts down completely while the prediction model continues to forecast nonzero active power. These significant prediction errors enable the NBM method to effectively detect these power loss events. Examination of the total wind farm data reveals that these shutdowns affected the entire farm, not just individual turbines. While a detailed root cause analysis is beyond the scope of this study, these events demonstrate the utility of our

methods in identifying and analyzing power loss events. Furthermore, the ability to predict potential power output during these events provides valuable insights into the magnitude of power losses and their associated revenue impacts.

4.4 Analysis of unknown power loss events

It is worth noting that there is a considerable number of False Positives. This outcome is partially expected, given the lack of an exhaustive list of status logs containing all power loss events. However, upon closer examination, two wind turbines stand out for having a significantly higher number of False Positives compared to the rest of the wind farm.
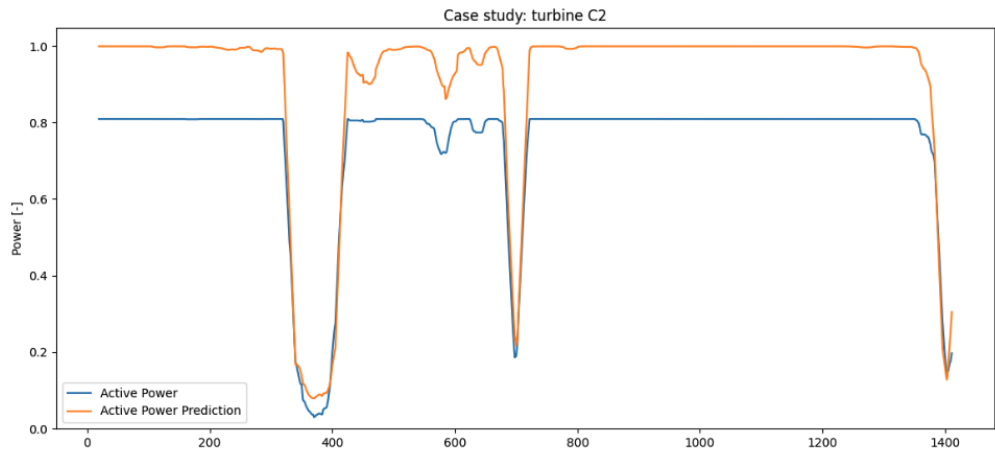
Figure 7. Performance analysis case study. A time series of the turbine's active power and the Spatio-Temporal GNN's prediction, both smoothed with a 10-minute rolling average, over a two-day period.

The first case involves a wind turbine that consistently produces around 80% of its rated power, as shown in Figure 7. A brief analysis indicates that this is a deration of an individual turbine within the farm, a detail not captured in the available status logs. Since derations often result from spontaneous operator decisions and are typically not well-documented in status logs, automated post hoc detection of these events is valuable for ensuring accurate availability data. This enables operators to more precisely assess key performance indicators like turbine availability, quantify revenue losses due to derated turbines, and verify compliance with required performance standards.

The second case is more complex and is depicted in Figure 8. In the days leading up to a maintenance event, we observed large and continuous discrepancies between the turbine's active power output and the model's predictions. The turbine consistently produced less energy than expected, followed by a deration and eventual shutdown. Upon investigating this previously unknown power loss event through historical logs, we discovered issues with the wind sensors on that turbine. These issues led to inconsistent pitch behavior, affecting the power output, as pitch regulation is dependent on accurate wind speed measurements from the sensors.
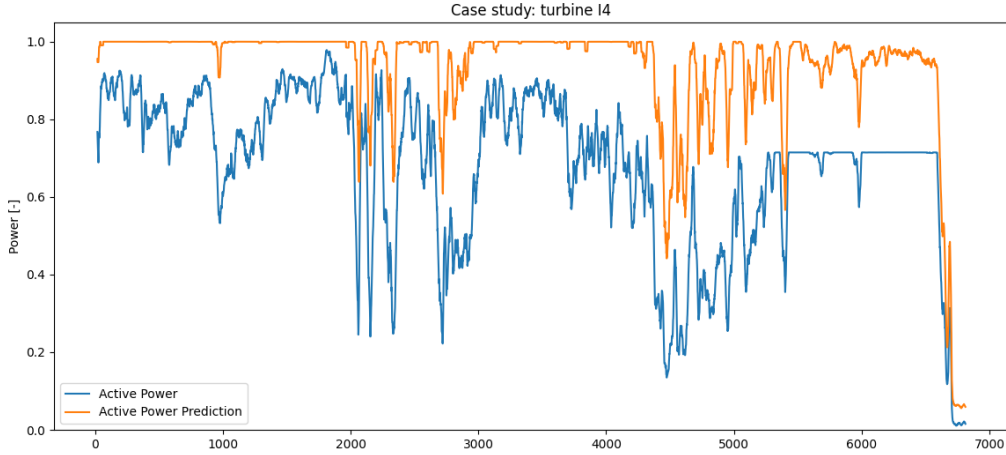
Figure 8. Performance analysis case study. A time series of the turbine's active power and the Spatio-Temporal GNN's prediction, both smoothed with a 10-minute rolling average, over a five-day period.

5 Discussion

As demonstrated in the Sect. 4, our GNN-based power prediction models consistently outperform traditional methods for both predicting power during normal operations and detecting abnormal events. The Spatio-Temporal GNN's unique ability to incorporate lagged input feature values introduces an additional temporal dimension, enabling it to identify and exploit trends over time. This temporal insight proves particularly valuable in capturing dynamic wind condition changes, such as fluctuations in wind speed and direction, while also smoothing out noise in high-frequency SCADA data.

Our findings show that the Spatio-Temporal GNN excels in power prediction, especially under waked conditions. Unlike the power curve binning method and a simple MLP, which struggle to accurately predict energy ratios in these scenarios, the Spatio-Temporal GNN achieves significantly greater accuracy. This consistency between model predictions and the intuitive understanding of turbine interactions under varying wind conditions highlights the model's capability to effectively capture the intricate dynamics of wake interactions within the wind farm.

Additionally, we validated our model in scenarios lacking ground truth data by using it to detect anomalies through deviations from expected behavior. The GNN-based methodology demonstrated remarkable proficiency in detecting nearly all power loss events, whereas the power curve binning method failed to identify a substantial portion of these anomalies.

The superior performance of the GNN models emphasizes the advantage of representing the wind farm as a graph. In this representation, each turbine not only considers its local measurements but also incorporates information from neighboring turbines through message passing, capturing complex spatial dependencies across the wind farm. By integrating wind farm topology into power predictions and learning local features, we anticipate that this model is transferable to new and unseen wind farms. However, this claim needs further investigation.

450 Despite its advanced capabilities, developing and training the Spatio-Temporal GNN requires minimal effort. A single model suffices for an entire wind farm, and training time ranges from minutes to a few hours on a GPU, depending on the selected hyperparameters. This efficiency, combined with its predictive power, makes the GNN-based approach a practical and scalable solution for wind farm power prediction and anomaly detection.

6 Conclusions

455 Our study introduces a robust and innovative approach ~~to~~ for predicting the potential power output of wind turbines within a wind farm at a 30-second temporal resolution by leveraging the spatial and temporal dynamics of the wind environment. By modeling the wind farm as a graph and using Graph Neural Networks (GNNs) to aggregate information from neighboring turbines, we significantly improve ~~the~~ prediction accuracy compared to traditional power curve methods. The addition of a temporal component ~~through a hybrid GNN-LSTM model~~ further enhances the model's ability to capture and leverage

460 temporal patterns in the data, ~~providing superior performance in terms of~~ achieving reductions of 30.3% and 30.5% in Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE), respectively, compared to the power curve method.

The ~~integration of our prediction model~~ Spatio-Temporal GNN demonstrates remarkable agreement with SCADA-derived energy ratios across the full range of wind directions, achieving a weighted average error of 0.0373; an improvement of approximately 57.4% compared to the power curve binning method. Notably, the model excels under waked conditions, where

465 the power curve and MLP models exhibit higher error rates.

Beyond power prediction, our methodology shows exceptional utility in detecting turbine underperformance and power loss events. Integrated into a Normal Behavior Model (NBM) framework~~shows its practical utility in analyzing turbine underperformance and identifying~~, the Spatio-Temporal GNN achieves a remarkably low False Negative rate of just 0.02%, identifying nearly all power loss events accurately. In contrast, the power curve binning method leaves approximately 12.6% of such events

470 undetected, underscoring the practical advantages of the GNN-based approach. ~~Our model's ability to detect abnormal behavior with high accuracy, even in the presence of complex wake interactions and varying wind conditions, demonstrates its potential for real-world applications.~~ The case studies presented validate the model's effectiveness in detecting grid curtailments, shutdowns, individual turbine derations, and anomalous behavior, offering valuable insights into the operational performance of wind farms.

475 Our findings confirm the scalability and practicality of the GNN approach. A single Spatio-Temporal GNN model suffices for an entire wind farm, requiring minimal computational effort and delivering significant accuracy improvements.

Future work could aim to ~~enhance~~ improve model accuracy by incorporating additional environmental factors and expanding its application to different wind farms. ~~Additionally,~~ Developing a model that operates effectively across multiple wind farms presents a promising research direction. Achieving this, however, requires further efforts in data standardization and the

480 consolidation of varying data formats across wind farms. In this context, the use of ontologies and taxonomies (e.g., RDS-PP), could facilitate consistent data integration. Furthermore, we plan to explore the use of this power prediction ~~method~~ framework in wind farm control. ~~In a control setting, the~~ Accurate predictions of potential power output ~~of each turbine is crucial for~~

are critical for optimizing power setpoints; however, such optimization must also consider the load spectrum to balance energy production and turbine longevity (Verstraeten et al., 2019; Nejad et al., 2022).

Author contributions. SD: Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft preparation, & editing; TV: Conceptualization, Supervision, Writing – review & editing; PD: Formal analysis; AN: Funding acquisition, Supervision; JH: Conceptualization, Funding acquisition, Supervision.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors would like to acknowledge the Energy Transition Funds for their support through the POSEIDON and BeFORECAST projects. This research was supported by funding from the Flemish Government under the "Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen" program.

References

495   Akiba, T., Sano, S., Yanase, T., Ohta, T., and Koyama, M.: Optuna: A Next-generation Hyperparameter Optimization Framework, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 2623–2631, https://doi.org/10.1145/3292500.3330701, 2019.

Bentsen, L., Warakagoda, N., Stenbro, R., and Engelstad, P.: Wind Park Power Prediction: Attention-Based Graph Networks and Deep Learning to Capture Wake Losses, J. Phys.: Conf. Ser., 2265, 022 035, https://doi.org/10.1088/1742-6596/2265/2/022035, 2022.

500   Bilendo, F., Badihi, H., Lu, N., Cambron, P., and Jiang, B.: A Normal Behavior Model Based on Power Curve and Stacked Regressions for Condition Monitoring of Wind Turbines, IEEE Transactions on Instrumentation and Measurement, 71, 1–13, https://doi.org/10.1109/TIM.2022.3196116, 2022.

Bleeg, J.: Graph Neural Networks for Power Prediction in Offshore Wind Farms using SCADA Data, J. Phys.: Conf. Ser., 1618, 062 054, https://doi.org/10.1088/1742-6596/1618/6/062054, 2020.

505   Daenens, S., Vervlimmeren, I., Verstraeten, T., Daems, P.-J., Nowé, A., and Helsen, J.: Power prediction using high-resolution SCADA data with a farm-wide deep neural network approach, J. Phys.: Conf. Ser., 2767, 092 014, https://doi.org/10.1088/1742-6596/2767/9/092014, 2024.

de N Santos, F., Duthé, G., Abdallah, I., Élouan Réthoré, P., Weijtjens, W., Chatzi, E., and Devriendt, C.: Multivariate prediction on wake-affected wind turbines using graph neural networks, Journal of Physics: Conference Series, 2647, 112 006, https://doi.org/10.1088/1742-6596/2647/11/112006, 2024.

Doekemeijer, B., Simley, E., and Fleming, P.: Comparison of the Gaussian Wind Farm Model with Historical Data of Three Offshore Wind Farms, Energies, 15, 1964, https://doi.org/10.3390/en15061964, 2022.

European Commission: Delivering on the EU offshore renewable energy ambitions, EUR-Lex, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52023DC0668, COM(2023) 668 final, 2023.

515   Fey, M. and Lenssen, J.: Fast Graph Representation Learning with PyTorch Geometric, https://doi.org/10.48550/arXiv.1903.02428, arXiv:1903.02428, 2019.

Fleming, P., King, J., Dykes, K., Simley, E., Roadman, J., Scholbrock, A., Murphy, P., Lundquist, J., Moriarty, P., Fleming, K., van Dam, J., Bay, C., Mudafort, R., Lopez, H., Skopek, J., Scott, M., Ryan, B., Guernsey, C., and Brake, D.: Initial results from a field campaign of wake steering applied at a commercial wind farm – Part 1, Wind Energ. Sci., 4, 273–285, https://doi.org/10.5194/wes-4-273-2019, 2019.

520   Gilmer, J., Schoenholz, S., Riley, P., Vinyals, O., and Dahl, G.: Neural Message Passing for Quantum Chemistry, in: Proceedings of the 34th International Conference on Machine Learning, vol. 70 of *PMLR*, pp. 1263–1272, 2017.

Hammer, F., Helbig, N., Losinger, T., and Barber, S.: Graph machine learning for predicting wake interaction losses based on SCADA data, Journal of Physics: Conference Series, 2505, 012 047, https://doi.org/10.1088/1742-6596/2505/1/012047, 2023.

International Electrotechnical Commission (IEC): Wind Energy Generation Systems—Part 12-1: Power Performance Measurements of Elec-
525   tricity Producing Wind Turbines, 2017.

International Energy Agency: Wind, International Energy Agency, https://www.iea.org/energy-system/renewables/wind, last updated on 11 July 2023, 2023.

Li, G., Xiong, C., Thabet, A., and Ghanem, B.: DeeperGCN: All You Need to Train Deeper GCNs, https://doi.org/10.48550/arXiv.2006.07739, arXiv:2006.07739, 2020.

530   Lin, Z. and Liu, X.: Wind power forecasting of an offshore wind turbine based on high-frequency SCADA data and deep learning neural network, Energy, 201, 117 693, https://doi.org/10.1016/j.energy.2020.117693, 2020.

Lyons, J. and Gocmen, T.: Applied machine learning techniques for performance analysis in large wind farms, Energies, 14, 3756, https://doi.org/10.3390/en14133756, 2021.

Nejad, A. R., Keller, J., Guo, Y., Sheng, S., Polinder, H., Watson, S., Dong, J., Qin, Z., Ebrahimi, A., Schelenz, R., Gutiérrez Guzmán,

535   F., Cornel, D., Golafshan, R., Jacobs, G., Blockmans, B., Bosmans, J., Pluymers, B., Carroll, J., Koukoura, S., Hart, E., McDonald, A., Natarajan, A., Torsvik, J., Moghadam, F. K., Daems, P.-J., Verstraeten, T., Peeters, C., and Helsen, J.: Wind turbine drivetrains: state-of-the-art technologies and future development trends, Wind Energy Science, 7, 387–411, https://doi.org/10.5194/wes-7-387-2022, 2022.

NREL: Simulator for Wind Farm Applications (SOWFA), https://github.com/NREL/SOWFA, 2012.

NREL: FLORIS, https://github.com/NREL/floris, 2024.

540   Park, J. and Park, J.: Physics-induced graph neural network: An application to wind-farm power estimation, Energy, 187(9), 115 883, https://doi.org/10.1016/j.energy.2019.115883, 2019.

Pedersen, M. M., Forsting, A. M., van der Laan, P., Riva, R., Romàn, L. A. A., Risco, J. C., Friis-Møller, M., Quick, J., Christiansen, J. P. S., Rodrigues, R. V., Olsen, B. T., and Réthoré, P.-E.: PyWake 2.5.0: An open-source wind farm simulation tool, https://gitlab.windenergy.dtu.dk/TOPFARM/PyWake, 2023.

545   Raasch, S. and Schröter, M.: PALM – A large-eddy simulation model performing on massively parallel computers, Meteorologische Zeitschrift, 10, 363–372, https://doi.org/10.1127/0941-2948/2001/0010-0363, 2001.

Robbelein, K., Daems, P., Verstraeten, T., Noppe, N., Weijtjens, W., Helsen, J., and Devriendt, C.: Effect of curtailment scenarios on the loads and lifetime of offshore wind turbine generator support structures, J. Phys.: Conf. Ser., 2507, 012 013, https://doi.org/10.1088/1742-6596/2507/1/012013, 2023.

550   Scarselli, F., Gori, M., Tsoi, A., Hagenbuchner, M., and Monfardini, G.: The Graph Neural Network Model, IEEE Transactions on Neural Networks, 20, 61–80, https://doi.org/10.1109/TNN.2008.2005605, 2009.

Van Binsbergen, D., Daems, P.-J., Verstraeten, T., Nejad, A., and Helsen, J.: Hyperparameter tuning framework for calibrating analytical wake models using SCADA data of an offshore wind farm, Wind Energy Science, 9, 1507–1526, https://doi.org/10.5194/wes-9-1507-2024, 2024a.

555   Van Binsbergen, D., Daems, P.-J., Verstraeten, T., Nejad, A., and Helsen, J.: Scalable SCADA-Based Calibration for Analytical Wake Models Across an Offshore Cluster, Journal of Physics: Conference Series, 2745, 012 014, https://doi.org/10.1088/1742-6596/2745/1/012014, 2024b.

Van Binsbergen, D., Daems, P.-J., Verstraeten, T., Nejad, A., and Helsen, J.: Performance comparison of analytical wake models calibrated on a large offshore wind cluster, Journal of Physics: Conference Series, 2767, 092 059, https://doi.org/10.1088/1742-6596/2767/9/092059,

560   2024c.

Verstraeten, T., Nowé, A., Keller, J., Guo, Y., Sheng, S., and Helsen, J.: Fleetwide data-enabled reliability improvement of wind turbines, Renewable and Sustainable Energy Reviews, 109, 428–437, https://doi.org/https://doi.org/10.1016/j.rser.2019.03.019, 2019.

Verstraeten, T., Daems, P., Bargiacchi, E., Roijers, D. M., Libin, P. J. K., and Helsen, J.: Scalable Optimization for Wind Farm Control using Coordination Graphs, CoRR, abs/2101.07844, https://arxiv.org/abs/2101.07844, 2021.

565   Yin, H., Ou, Z., Fu, J., Cai, Y., Chen, S., and Meng, A.: A novel transfer learning approach for wind power prediction based on a serio-parallel deep learning architecture, Energy, 234, 121 271, https://doi.org/10.1016/j.energy.2021.121271, 2021.

Yousefi, H., Abbaspour, A., and Seraj, H.: Worldwide Development of Wind Energy and CO2 Emission Reduction, Environmental Energy and Economic Research, 3, 1–9, https://doi.org/10.22097/eeer.2019.164295.1064, 2019.

Yu, M., Zhang, Z., Li, X., Yu, J., Gao, J., Liu, Z., You, B., Zheng, X., and Yu, R.: Superposition Graph Neural Network for offshore wind power prediction, Future Gener. Comput. Syst., 113, 145–157, https://doi.org/10.1016/j.future.2020.06.024, 2020.

Zhang, J., Liu, D., Li, Z., Han, X., Liu, H., Dong, C., Wang, J., Liu, C., and Xia, Y.: Power prediction of a wind farm cluster based on spatiotemporal correlations, Applied Energy, 302, 117 568, https://doi.org/10.1016/j.apenergy.2021.117568, 2021.

570