applied sciences

MDPI

Article

Equipment Health Assessment: Time Series Analysis for Wind Turbine Performance

Jana Backhus [1], Aniruddha Rajendra Rao [1,*], Chandrasekar Venkatraman [1], Abhishek Padmanabhan [2], A. Vinoth Kumar [3] and Chetan Gupta [1]

[1] Industrial AI Lab, R&D, Hitachi America, Ltd., Santa Clara, CA 95054, USA
[2] Centre of Excellence in Energy Sciences, Atria University, Bengaluru 560024, India; abhishekp@atriauniversity.edu.in
[3] Atria Brindavan Power Private Limited, Bangalore 560025, India; vinoth.kumar@atriapower.com
* Correspondence: aniruddha.rao@hal.hitachi.com

Abstract: In this study, we leverage SCADA data from diverse wind turbines to predict power output, employing advanced time series methods, specifically Functional Neural Networks (FNN) and Long Short-Term Memory (LSTM) networks. A key innovation lies in the ensemble of FNN and LSTM models, capitalizing on their collective learning. This ensemble approach outperforms individual models, ensuring stable and accurate power output predictions. Additionally, machine learning techniques are applied to detect wind turbine performance deterioration, enabling proactive maintenance strategies and health assessment. Crucially, our analysis reveals the uniqueness of each wind turbine, necessitating tailored models for optimal predictions. These insight underscores the importance of providing automatized customization for different turbines to keep human modeling effort low. Importantly, the methodologies developed in this analysis are not limited to wind turbines; they can be extended to predict and optimize performance in various machinery, highlighting the versatility and applicability of our research across diverse industrial contexts.

Keywords: SCADA data; time series; wind turbine; prediction; classification; ensemble

check for updates

Citation: Backhus, J.; Rao, A.R.; Venkatraman, C.; Padmanabhan, A.; Kumar, A.V.; Gupta, C. Equipment Health Assessment: Time Series Analysis for Wind Turbine Performance. *Appl. Sci.* **2024**, *14*, 3270. https://doi.org/10.3390/app14083270

Academic Editor: José António Correia

Received: 7 March 2024
Revised: 2 April 2024
Accepted: 3 April 2024
Published: 12 April 2024

1. Introduction

In an era marked by mounting concerns about the environmental impact of traditional energy sources, the quest for sustainable and renewable alternatives has taken center stage in the global discourse. The imperative to reduce greenhouse gas emissions, secure energy independence, and mitigate the effects of climate change has driven a fundamental shift in our approach to power generation [1,2]. At the forefront of this transformative journey stands wind energy, a clean and abundant resource harnessed to provide a solution to our growing energy needs while reducing our carbon footprint [3,4]. Wind turbines gracefully punctuate landscapes symbolizing our commitment to a cleaner and more sustainable future. However, the pursuit of harnessing the power of the wind comes with its own set of challenges [5–7]. The intermittent nature of wind, the need for effective grid integration, and the demands of operating and maintaining these complex machines have sparked a revolution in how we manage and optimize wind energy facilities.

Sustainable maintenance practices ensure equipment availability, reliability, and safety while minimizing environmental impact [8,9]. These practices focus on preventive maintenance, efficient resource use, and eco-friendly technologies. Integrating sustainability into maintenance involves aligning activities with broader goals like reducing energy consumption and waste generation. By adopting these practices, organizations enhance operational efficiency, cut costs, and promote environmental conservation. At the heart of Wind farm operation and maintenance revolution lies the Supervisory Control and Data Acquisition (SCADA) system, a vital component that allows us to remotely monitor, control, and collect crucial data from wind turbines [10,11]. The insights offered by data obtained from SCADA systems ensure the efficient operation of wind turbines but also play a pivotal role in their longevity [12,13]. Refs. [14,15] discusses the use of SCADA data for Wind Turbine Performance and explainable artificial intelligence. They also mention the potential of multivariate time series analysis, which we have exploited in this work.

In this study, we leverage SCADA data from 13 wind turbines located in a wind farm in India to predict the power output of the wind turbines, employing advanced time series methods, specifically Functional Neural Networks (FNN) [16,17] and Long Short-Term Memory (LSTM) networks [18]. A key innovation lies in an ensemble of FNN and LSTM model to capitalize on their collective learning to make a more effective prediction. This approach outperforms the individual models, ensuring high accuracy of the power output predictions when wind turbine is in a good state (good timeline) and significantly lower accuracy when the wind turbine is in a bad state (bad timeline). This difference is crucial in making sure we are learning the correct mapping between the features and power output of a wind turbine. Then, statistical techniques are applied to the prediction performance errors to detect wind turbine deterioration, enabling proactive maintenance strategies and health assessment.

Our analysis leads to the important understanding that the uniqueness of each wind turbine necessitates tailored models for optimal predictions. This highlights the significance of offering automated customization for different turbines to minimize human efforts for modeling. Importantly, the methodologies devised in this analysis are not restricted to wind turbines; they can be extended to various other machinery to enhance their performance. This showcases the versatility and relevance of our research across a wide range of industrial settings.

This research is part of an effort to leverage equipment related data to analyze the health of the equipment where multiple entities are placed in a similar environment, but each entity is showing slightly different behavior [19,20]. The goal of this study is to develop methodologies that are not limited to our first target of wind turbines but can be extended to predict and optimize equipment performance in various machinery. Furthermore, we assume that it will be too much human effort to fine-tune a prediction model for each machine entity and therefore we explore machine learning methodologies that are robust enough to perform considerably well over all entities. In summary, this is an initial study that we hope to extend into a more versatile and applicable research across diverse industrial contexts in the future.

The rest of this paper is organized as follows: In Section 2, we first describe the data before diving into the methods, model settings and our approach. We follow this the results of the power output prediction for the wind turbines and the deterioration detection for them in Section 3. Section 4 discusses the results and our understand and insights from the experiments. We also share the validation we got from domain experts. In the final section, we present our concluding remarks and future research directions which pertain explainable analysis, adaption of our work in the wind farms and other improvements.

2. Data and Methods

2.1. Data

A SCADA dataset for wind turbine power generation collects real-time operational, environmental, anomalous, electrical, and communication data from individual turbines or across an entire wind farm [21,22]. It encompasses turbine status (on/off, power output, rotor speed), environmental conditions (wind speed, direction, temperature), alarms for irregularities, electrical parameters (voltage, current), and communication details. The initial SCADA dataset measures approximately 2000 different features. We have limited our scope of interest, however, before downloading the data from a data collection server [7], to keep the data size small. Therefore, we have excluded all features that are metadata like timestamps or Boolean values related to wind turbine status. For our study, we are interested in the power generation performance of each individual wind turbine. Therefore, we mainly focus on the measurements of actual power output and any related factors. We have chosen relevant parameters based on the following turbine components: e.g., rotor, pitch, gearbox, nacelle, power, and environmental information like wind [23].

After downloading an initial dataset with 76 features, we have conducted an exploratory data and correlation analysis on one wind turbine (no. 5) to further downsize the dataset of interest. The wind turbine was selected based on domain information that it has active pitch alignment and is considered one of the better performing wind turbines. We were able to decrease the feature set to 24 features, omitting 9 features because of rare changes in values and 43 features because of very high correlation (larger than 0.9) to at least one other feature. The correlation heatmap of the remaining features of interest is shown in Figure 1. Here, we highlighted the row and column for the actual power output (“WTUR.W.mag.f”) which is our target feature of interest to evaluate the power generation performance of a wind turbine. We also added an additional feature to the heatmap which is a calculation of the sinus radius of the wind direction, where the original wind direction values are given in degrees.

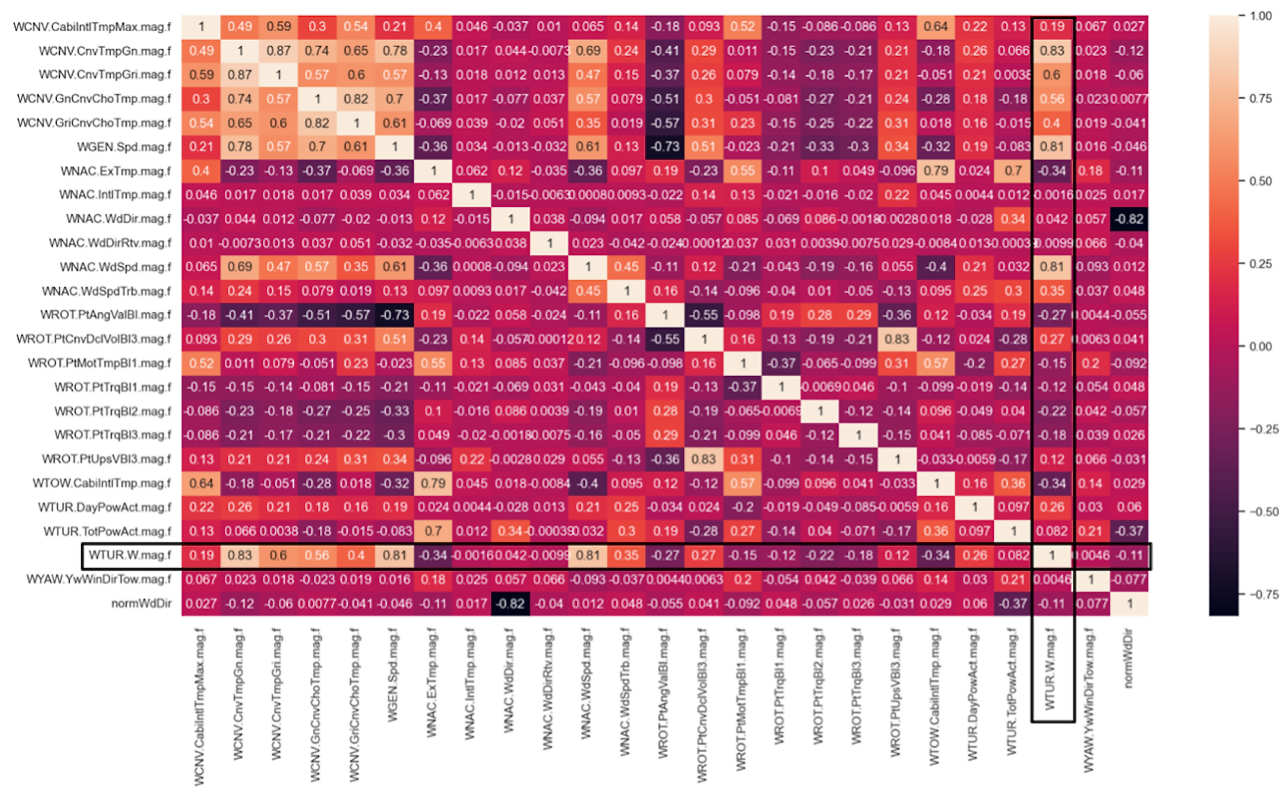


Figure 1. Correlation Heatmap between Features of Interest (the Power output is denoted using the Black box).

Figure 2 shows the actual power output of the 13 wind turbines over the available time frame from end of October 2022 to early June 2023. The timelines colored in orange are considered “good” based on initial information received from the data provider. We also got informed that the whole wind farm is known to show “bad” performance between mid of November to mid of December 2022 due to dust collection. Furthermore, we can observe in Figure 2 that there are several periods of missing data mostly caused by temporary shutdowns of the SCADA data collection system.

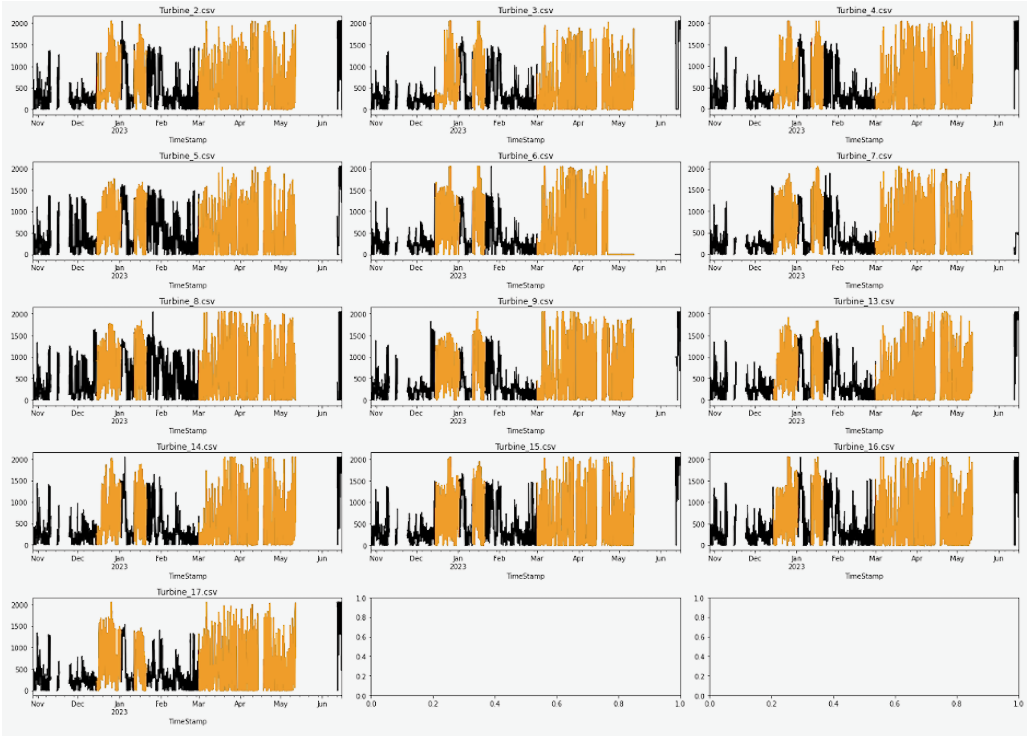


Figure 2. Actual power output of the 13 wind turbines (black: bad performance, orange: good performance).

2.2. Methods

In predicting power outputs, traditional and statistical methods have long been employed to model relationships between variables [24–27]. Traditional machine learning methods like linear regression are fundamental tools, assuming a linear relationship between predictors and the power function output [28]. Polynomial regression, an extension of linear regression, can capture more complex relationships by introducing polynomial terms. These methods, while interpretable and computationally efficient, might struggle to capture intricate nonlinear patterns present in power functions, limiting their accuracy [29]. They also ignore the temporal nature of the data. We therefore consider two different time series prediction approaches using Deep Learning (DL) methods, which have been show to be most successful in the literature [17,30–32], for our wind turbine power output prediction as described in the following two sections.

2.2.1. Long Short-Term Memory (LSTM)

Recently, deep learning methods, especially recurrent neural networks (RNNs) and long short-term memory networks (LSTMs) [18], have revolutionized the time series prediction landscape [31,33]. When applied to power function prediction, these DL methods can capture intricate dependencies and nonlinearities inherent in the data. Additionally, convolutional neural networks (CNNs) can be employed to capture spatial patterns within the data [32]. Deep learning models, while computationally intensive, offer the advantage of automatic feature extraction, enabling them to discern complex relationships, and making them highly effective in accurately predicting power functions even in dynamic and intricate scenarios. Their ability to handle vast amounts of data and learn hierarchical representations often leads to superior performance in power function prediction tasks [34].

LSTM is a type of RNN architecture, specifically designed to capture long-term dependencies in sequential data, making them ideal for predicting power functions over time as seen in Figure 3. Traditional RNNs often struggle with learning patterns in data sequences that are separated by many time steps. LSTMs were introduced to address this issue by incorporating memory cells and sophisticated gating mechanisms. LSTMs are widely used in various applications such as natural language processing, speech recognition, and time series prediction due to their ability to capture long-term dependencies in sequential data. LSTM components are as follows:

1. Hidden state & new inputs—hidden state from a previous timestep (h_{t-1}) and the input at a current timestep (x_t) are combined before passing copies of it through various gates.
2. Forget gate—this gate controls what information should be forgotten. Since the sigmoid function ranges between 0 and 1, it sets which values in the cell state should be discarded (multiplied by 0), remembered (multiplied by 1), or partially remembered (multiplied by some value between 0 and 1).
3. Input gate helps to identify important elements that need to be added to the cell state. Note that the results of the input gate get multiplied by the cell state candidate, with only the information deemed important by the input gate being added to the cell state.
4. Update cell state—first, the previous cell state (c_{t-1}) gets multiplied by the results of the forget gate. Then we add new information from [input gate $\times$ cell state candidate] to get the latest cell state (c_t).
5. Update hidden state—the last part is to update the hidden state. The latest cell state (c_t) is passed through the tanh activation function and multiplied by the results of the output gate.
6. Finally, the latest cell state (c_t) and the hidden state (h_t) go back into the recurrent unit, and the process repeats at timestep $t + 1$. The loop continues until we reach the end of the sequence.

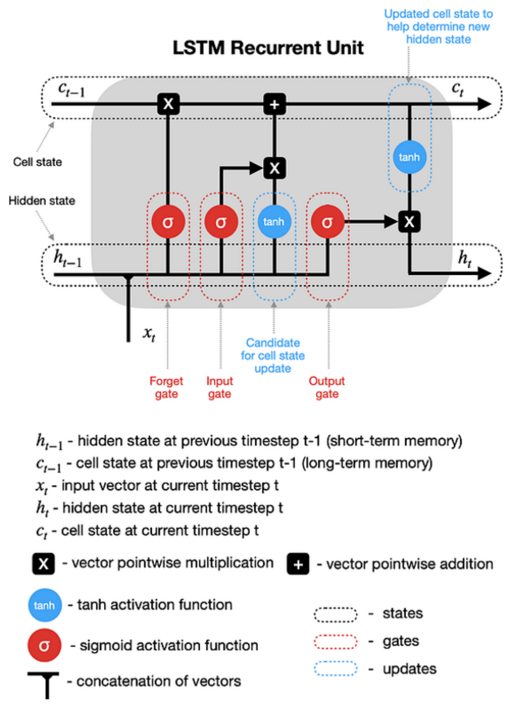


Figure 3. LSTM Cell (https://towardsdatascience.com/lstm-recurrent-neural-networks-how-to-teach-a-network-to-remember-the-past-55e54c2ff22e, accessed on on 15 January 2024).

2.2.2. Functional Neural Networks (FNN)

On the other hand, Functional data analysis (FDA) [35–37] is a branch of statistics that deals with data that are functions or curves, rather than simple numeric or categorical values. In FDA, the data is treated as a continuous function, and the goal is to analyze and model the behavior of the function over time or space. FDA methods allow researchers to extract meaningful information from functional data, such as trends, patterns, and underlying structures. This information can then be used to develop models and make predictions or forecasts. A Functional Linear Model (FLM) [38,39] is a statistical model used in FDA that extends the linear regression model to functional data. It is designed to handle data that consists of a collection of curves or functions, rather than discrete data points. The FLM assumes that the relationship between the response variable and the predictor variables can be represented by a linear function, but with a functional rather than scalar form. The predictor variables can be either functional or scalar and can be continuous or categorical. In FLM, the functional form of the predictor variable is expressed using a basis expansion of a finite set of basis functions. These basis functions can be any set of orthogonal functions that span the space of functions under consideration. For example, Fourier, wavelet, or B-spline basis functions can be used.

Traditional time series techniques rely on statistical methods, such as ARIMA, exponential smoothing, and FLM. These methods have been widely used for decades and have proven to be effective in many applications. However, they have some limitations, such as being sensitive to the assumptions made about the data and the need for manual feature engineering. One of the more popular deep learning models for time series prediction is Functional Neural Networks (FNN) [30,40], that have been shown to outperform traditional time series models in many applications due to their ability to handle complex and nonlinear relationships in the data. This is particularly useful in applications where the relationships between variables are not well understood or there are many interacting factors. A Functional Neural Network (FNN) [16,17] as shown in Figure 4, is composed of a series of interconnected continuous neurons that are designed to process functional data. The input layer of the network takes in functional data, and each continuous neuron given by Equation (1) in the continuous hidden layer performs a non linear transformation on the input data or values coming in from the previous layer. The output layer of the network produces a functional output that can be used for prediction or forecasting.

$$H_{(l)}^{(r)}(s) = \sigma\Big(b_{(l)}^{(r)}(s) + \sum_{j=1}^{J} \int w_{(l)}^{(r,j)}(s,t) H_{(l-1)}^{(r,j)}(t)dt\Big), \tag{1}$$

where l indicates the layer number and r, j are the neuron number, $l = 1, 2, 3, \ldots, L$, $H_{(0)}^{(r)}(s) = X_r(s)$ (input time series), $H_{(L)}^{(r)}(s) = \widehat{Y_r(s)}$ (output time series), $\sigma(\cdot)$ is a non-linear activation function, $b_{(l)}^{(r)} \in \mathcal{L}^2(\mathcal{T})$ is the unknown intercept function, $w_{(l)}^{(r,j)} \in \mathcal{L}^2(\mathcal{T} \times \mathcal{T})$ is the bivariate parameter function for the r^{th} continuous neuron in the l^{th} hidden layer coming from the j^{th} continuous neuron of the $(l-1)^{th}$ hidden layer.

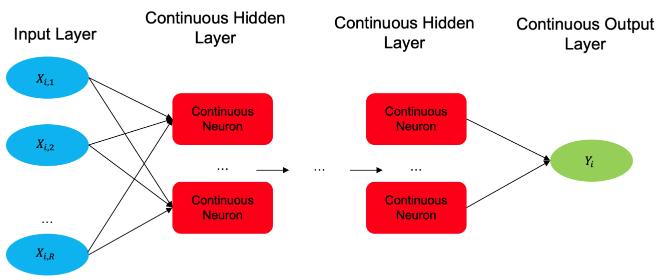


Figure 4. The general architecture of Functional Neural Network (FNN).

The architecture of FNN can vary widely, depending on the specific application and the complexity of the data. The training typically involves backpropagation through functional derivatives or basis expansion, a method that adjusts the weights of the network to minimize the error between the predicted output and the actual output. The choice of loss function and optimization algorithm also impacts the performance of the network.

2.3. Model Settings

2.3.1. Hyperparameter Search

To obtain robust power prediction results over all wind turbines, we have conducted an exhaustive hyperparameter search in a grid search manner to understand the performance of the model settings over the dataset with different feature selections. The experiments were conducted separately for two network types: LSTM and FNN. In the end, we have settled on a small set of selected features which are described in Table 1 for the two model types.

Table 1. Selected Features for LSTM and FNN.

Model	Selected Features	Total No. of Features
LSTM	Wind Speed Generator Speed Rotator Blade Pitch Angle $\times$ Generator Speed Rotator Blade Pitch Angle $\times$ Wind Speed Turbulences	4
FNN	Wind Speed Generator Speed Wind Speed $\times$ Generator Speed Wind Speed Turbulences	4

For the hyperparameters, we have not only investigated different model sizes but also tested different activation functions (i.e., tanh, ReLU, LeakyReLU, ELU, Sigmoid), learning rates, and error criterions (i.e., MSE, MAE). In addition, wind turbine power output is generally limited to a maximum and minimum possible power output and our models were often over- or underpredicting these limits. Therefore, we have also experimented with different output cutoff methods such as tanh, sigmoid, and a hard cutoff at the minimum and maximum. For the experiments presented in this study, we set our hyperparameters as stated in Table 2.

Table 2. Hyperparameter Settings for LSTM and FNN.

Model	Activation Function	Loss Error	Output Cutoff	Learning Rate	Network Size
LSTM	ReLU	MSE	Hard cutoff	0.001	Hidden Size = 2 Common Channel = 24
FNN	ELU	MSE	Sigmoid	0.01	Hidden Size = 1 Common Channel = 20 Common Size = 40

2.3.2. Ensemble Method

Ensemble methods involve the amalgamation of multiple machine learning models', a strategy proven to elevate predictive performance, and the ability to combat overfitting, a prevalent issue in machine learning [41,42]. Ensemble methods have gained a lot of popularity in recent years for different time series tasks [43,44]. Amalgamations can iron out inconsistencies and errors inherent to individual models, reducing bias and variance in the final predictions and leading to enhanced generalization capabilities even in the face of noisy data and outliers. Ensembles can be achieved through a variety of techniques, including employing different algorithms [45], training on distinct data subsets [46], and utilizing varying feature selection strategies [47]. By virtue of their diversity in model composition, ensembles often produce superior, more reliable prediction across various scenarios.

2.3.3. Deterioration Detection Approach

In this study, we leverage the strengths of ensemble methods by amalgamation of the prediction outputs of our separately trained models for FNN and LSTM with an equal weighting of 0.5. Additionally, we investigate a wind turbine equipment health assessment method that leverages the prediction outputs of the ensemble model and compares it to the real outputs by calculating the root mean squared error (RMSE) and the root mean squared percentage error (RMSPE) that are defined in Equations (2) and (3) respectively.

$$\text{RMSE} = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}\left(Y_i - \widehat{Y}_i\right)^2} \tag{2}$$

$$\text{RMSPE} = \sqrt{\frac{\Sigma_{i=1}^{n}\left(Y_i - \widehat{Y}_i\right)^2}{\Sigma_{i=1}^{n} Y_i}} \tag{3}$$

Our expectation is that the prediction model is trained on the good timeline. If the measured prediction errors increase then it means that the wind turbine power output diverges from the expected power output and the wind turbine performance has degraded. We can calculate RMSE and RMSPE cutoff limits based on the validation dataset for the good timeline. With these cutoff limits, we can detect a degradation in performance of the wind turbine from newly arriving data by comparing predicted and true power output values. An overview of the whole proposed method is shown in Figure 5.

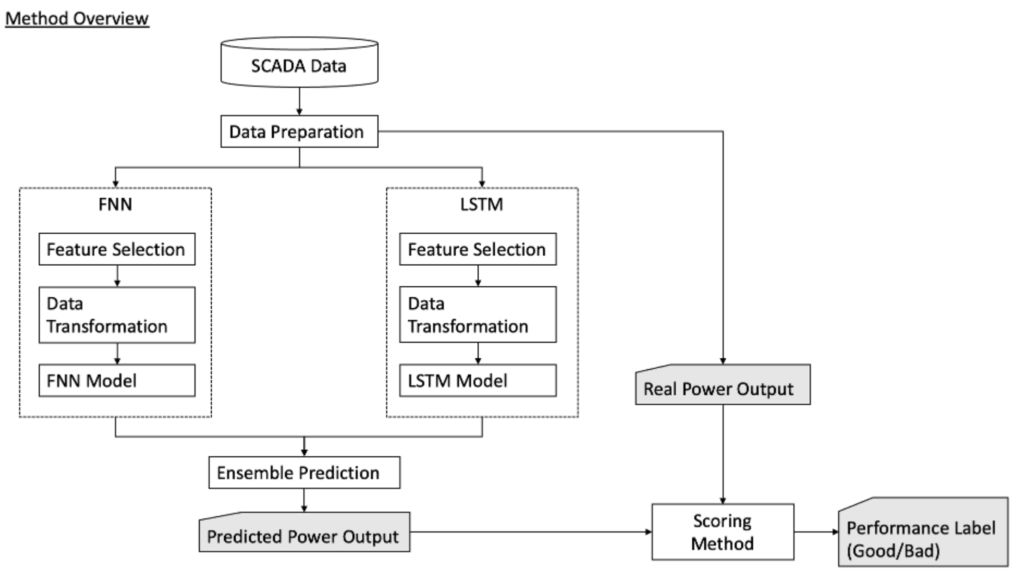


Figure 5. Deterioration Detection Method Overview.

3. Results

3.1. Prediction Models

Figure 6 shows the overall flow of our proposed prediction method. In the following, we present the obtained results for prediction and deterioration detection. We have conducted many trial and error studies, where we experimented with LSTM and FNN time series modeling and compared the prediction outputs based on the RMSE. For these studies, we separated the data into “good” and “bad” performing time periods based on domain knowledge obtained from the data provider. A power output prediction model for each wind turbine was trained based on it’s “good” data timeline and then tested on both a subset of “good” and “bad” data timelines.

The training, validation, and test data for the good timeline are created based on randomly ordered three-day time windows to account for different seasons in the training data. To increase the number of samples in the training, validation, and test data, we have used a sliding window approach as depicted in Figure 7. First, we assign hourly resampled data to a 3-day long day group and then apply a sliding window approach within each 3-day group. This makes sure there is no data leakage. With this approach, we can increase the number of total samples while enabling the random assignment the training, validation, or test dataset. For our first prediction evaluation, we created the test dataset for the bad timeline in a similar manner, and we have aligned the randomization over all wind turbines to compare the same timelines in the same training, validation or test categories.

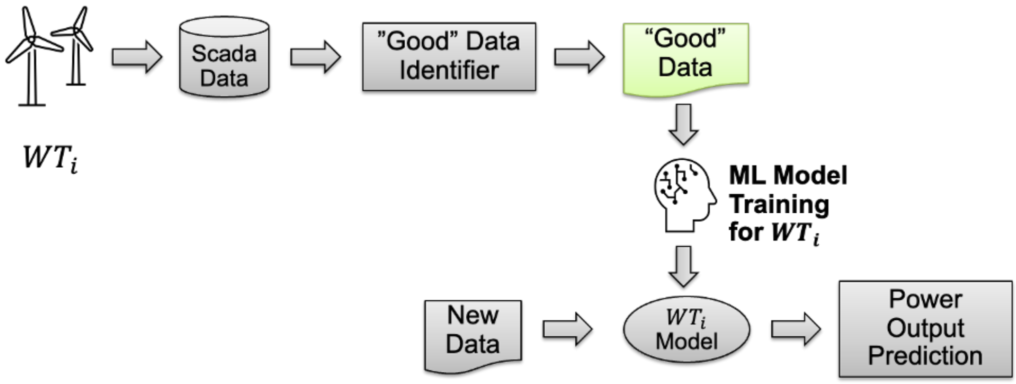
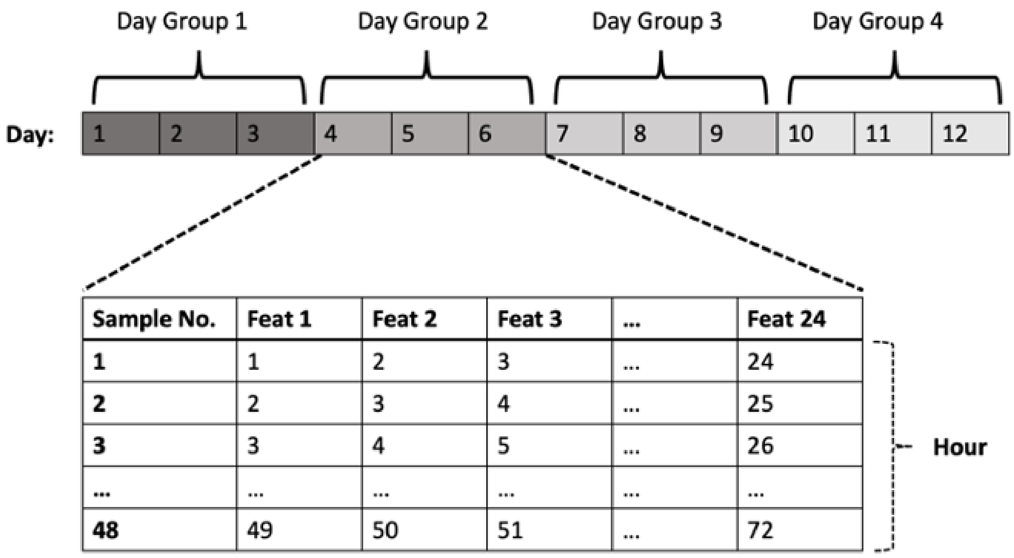


Figure 6. Workflow for wind turbine power output prediction.



Sample No.	Feat 1	Feat 2	Feat 3	...	Feat 24
1	1	2	3	...	24
2	2	3	4	...	25
3	3	4	5	...	26
...	...	...	...	...	...
48	49	50	51	...	72

Figure 7. Sliding window approach for 3-day time periods.

An example of how such power output predictions would look like for a sample wind turbine no. 15 using LSTM and FNN model is shown in Figures 8 and 9 respectively for both "good" and "bad" test datasets. We can clearly observe that the prediction model is often overpredicting the power output for the bad timeline datasets, which indicates that the maximum power output could be potentially higher and there is an issue with the wind turbine performance. The poor performance can also be confirmed on the bad timeline data's power curves for the wind and generator speed (right side of Figures 8 and 9), as it deviates away from it's sigmoid behaviour. Comparing the predictions of LSTM and FNN, we observe that LSTM achieves a higher prediction accuracy on both good and bad test datasets whereas FNN offers more separation between them.

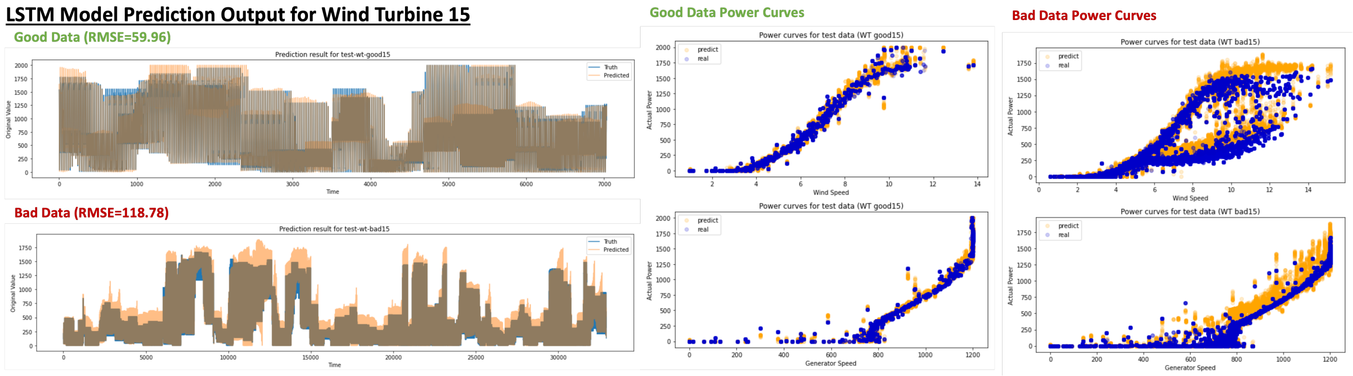


Figure 8. LSTM Model prediction output for both good and bad data periods (blue: true values, orange: predicted values).

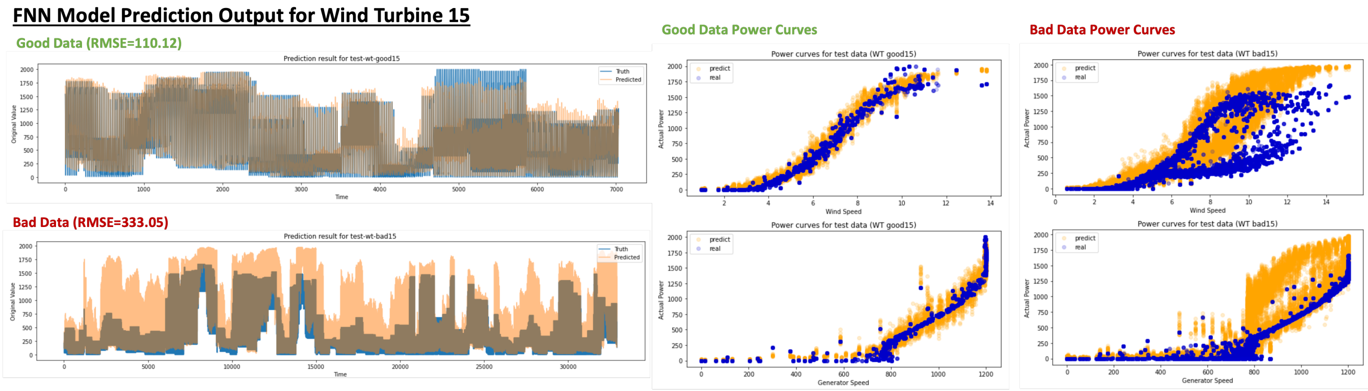
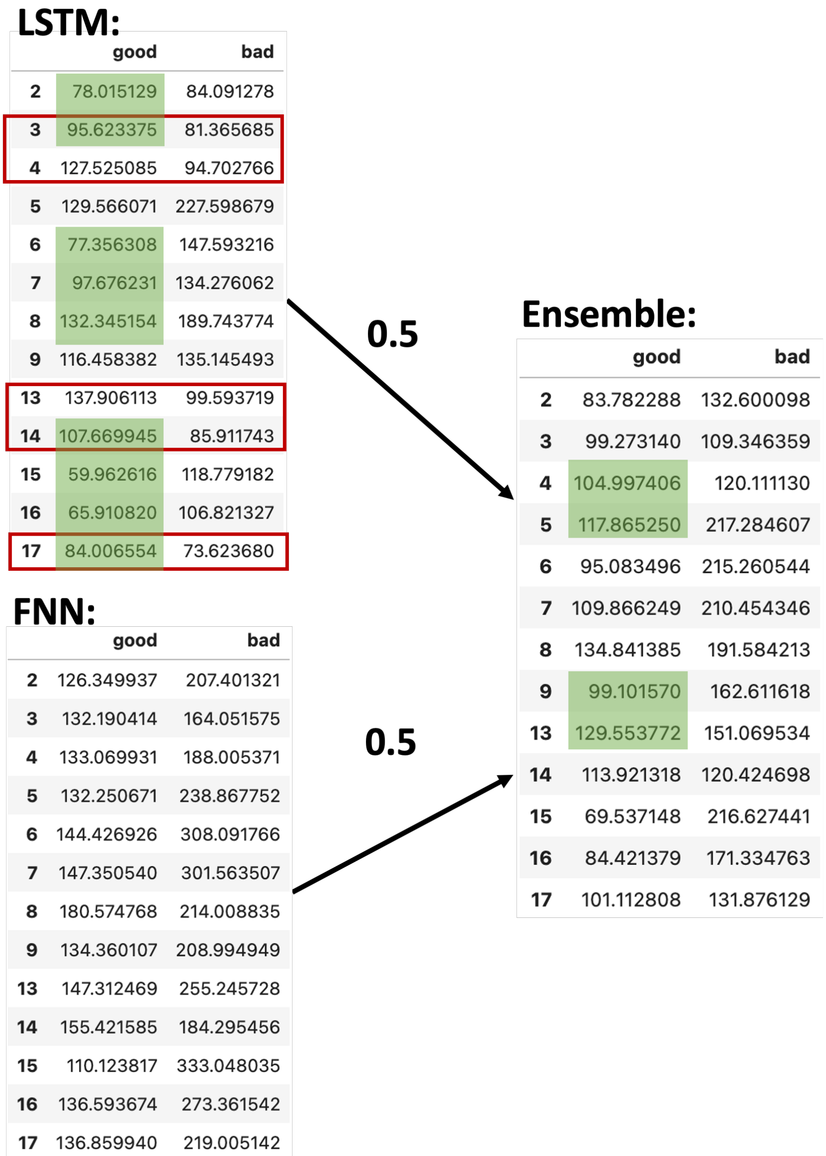


Figure 9. FNN Model prediction output for both good and bad data periods (blue: true values, orange: predicted values).

This observation can be further confirmed in Figure 10 that shows the RMSE for predictions of three different models LSTM, FNN and an ensembled prediction model where LSTM and FNN predictions are accounted for with equal weights. Overall, we can conclude that the LSTM model performs better on all 13 wind turbine indices for both good and bad test datasets in comparison to FNN. However, the LSTM model also performs suspiciously well on the bad test datasets compared to the good datasets for the wind turbines no. 3, 4, 13, 14 and 17 (highlighted in red in Figure 10). The model might be overfitting. Since we want to leverage the prediction accuracy as an indicator of performance deterioration, we need a high performance accuracy on good timelines and a poor performance accuracy on bad timelines, the latter part is provided by FNN. We achieve this by combining the LSTM and FNN model predictions in an equally weighted ensemble shown on the right side of Figure 10. The green highlights indicate the best performing model for each of the wind turbine on the good timeline test datasets. The ensemble model achieves a similar performance on good timeline datasets and even outperforms LSTM prediction accuracy in four cases (for wind turbine no. 4, 5, 9 and 13). We also achieve a bigger gap between prediction accuracy's for the good timeline and bad timeline datasets. For the ensemble method, no wind turbine prediction model performs better on the bad timeline dataset compared to the good timeline dataset.



LSTM:

	good	bad
2	78.015129	84.091278
3	95.623375	81.365685
4	127.525085	94.702766
5	129.566071	227.598679
6	77.356308	147.593216
7	97.676231	134.276062
8	132.345154	189.743774
9	116.458382	135.145493
13	137.906113	99.593719
14	107.669945	85.911743
15	59.962616	118.779182
16	65.910820	106.821327
17	84.006554	73.623680

FNN:

	good	bad
2	126.349937	207.401321
3	132.190414	164.051575
4	133.069931	188.005371
5	132.250671	238.867752
6	144.426926	308.091766
7	147.350540	301.563507
8	180.574768	214.008835
9	134.360107	208.994949
13	147.312469	255.245728
14	155.421585	184.295456
15	110.123817	333.048035
16	136.593674	273.361542
17	136.859940	219.005142

Ensemble:

	good	bad
2	83.782288	132.600098
3	99.273140	109.346359
4	104.997406	120.111130
5	117.865250	217.284607
6	95.083496	215.260544
7	109.866249	210.454346
8	134.841385	191.584213
9	99.101570	162.611618
13	129.553772	151.069534
14	113.921318	120.424698
15	69.537148	216.627441
16	84.421379	171.334763
17	101.112808	131.876129

Figure 10. Ensemble Prediction Performance Comparison.

Figure 11 shows the prediction performance for our previous example wind turbine no. 15 for the good timeline and bad timeline test datasets. The performance stays robust for the good timeline dataset but prediction performance deterioration for the bad timeline dataset is now much clearer. In Figure 12, we show the prediction results (RMSE) for the "good timeline" test dataset over all wind turbines for the ensemble prediction model. The first column of the table is the wind turbine no. on whose training data the wind turbine prediction model was built. Each wind turbine prediction model was used to not only predict the power output for its own test dataset but for all other wind turbines' "good timeline" test datasets as well. The obtained results are interesting, there's an overall tendency we can observe that each wind turbine is performing comparably well on its own test dataset as shown by the light green highlights indicating the minimum performance error of each column that often aligns with a wind turbine's own prediction model. Also, we can observe that performance results are mostly worst for wind turbine 8 (red highlight indicates max performance error of each row) and best for wind turbine 15 (dark green highlight indicates minimum performance error of each row). Wind turbine 8 might be affected by factors that are not considered in the selected features (or some unknown external features) making the power prediction more difficult while wind turbine 15 is the most straight forward to predict.

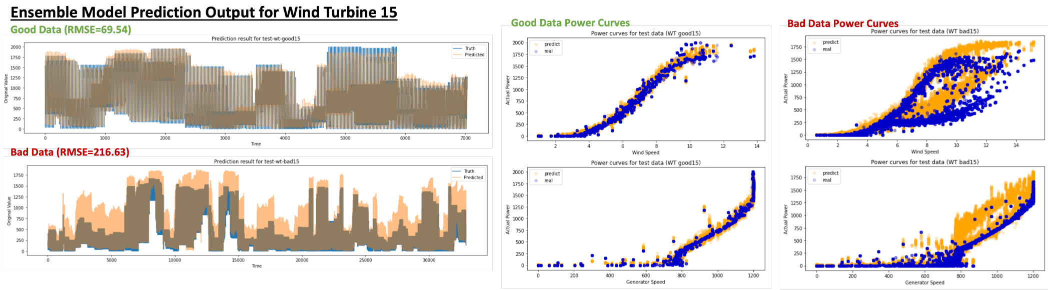


Figure 11. Ensemble Model prediction output for both good and bad data periods (blue: true values, orange: predicted values).

	test-wt-good2	test-wt-good3	test-wt-good4	test-wt-good5	test-wt-good6	test-wt-good7	test-wt-good8	test-wt-good9	test-wt-good13	test-wt-good14	test-wt-good15	test-wt-good16	test-wt-good17
2	83.782	136.378	145.027	189.965	146.243	143.661	215.312	135.840	150.201	122.800	91.826	112.109	128.756
3	100.288	99.273	106.015	159.075	114.818	116.629	174.302	152.144	149.175	135.106	113.652	106.475	108.651
4	81.848	107.145	104.997	168.576	94.985	114.710	181.450	136.649	138.985	111.810	76.291	81.508	91.501
5	90.443	109.272	122.697	117.865	130.996	143.482	186.734	125.696	156.098	129.231	86.651	108.195	119.620
6	98.245	117.292	134.090	165.168	95.084	121.167	178.005	124.580	170.488	141.462	94.243	98.273	97.301
7	74.204	107.099	137.028	169.688	108.985	109.866	172.274	116.351	137.093	127.775	75.475	88.765	113.370
8	135.387	129.111	118.323	168.297	120.190	142.705	134.841	160.686	177.519	162.611	127.345	130.332	137.721
9	84.956	123.596	117.650	164.512	114.485	124.697	194.774	99.102	135.069	108.556	77.995	91.992	90.182
13	109.344	112.434	136.330	177.503	147.250	134.154	200.567	149.085	129.554	116.064	113.678	124.505	134.138
14	95.419	114.176	120.120	163.659	121.011	126.099	200.988	144.124	135.289	113.921	94.676	111.053	126.778
15	81.313	129.334	136.916	177.497	119.428	141.093	188.175	123.075	153.901	115.319	69.537	90.609	128.987
16	100.890	113.684	108.436	186.700	98.218	126.376	189.219	126.721	165.212	136.509	99.955	84.421	96.825
17	75.743	122.189	127.592	180.115	108.983	110.301	183.779	107.923	151.943	136.777	66.739	98.054	101.113

Figure 12. Ensemble Prediction Results over all good timeline test datasets (best and worst values are highlighted per row/column as dark green: best in row, red: worst in row, light green: best in column, orange: worst in column).

In Figure 13, we can see the ensemble model prediction results (RMSE) for the bad timeline test dataset over all wind turbines. The first column of the table is the wind turbine no. on whose good timeline training dataset the prediction model was built. We used the prediction models for each wind turbine to not only predict the power output for its own test dataset but for all other wind turbines' "bad timeline" test datasets. Overall, we can observe that each wind turbine is performing worse compared to the "good timeline" test datasets in Figure 12. Also, we can observe that prediction performance results are worst for wind turbine 8 and best for wind turbine 2 regardless of the chosen prediction model. On the other hand, the prediction model for wind turbine no. 3 achieves the overall best prediction performance (light green highlights for the best entry in each column) and wind turbine no. 5 the worst (orange highlights for worst entry in each column).

In summary, Figure 12 showed us that the best prediction model for a wind turbine is based on its own training data. Results of Figure 13 on the other hand showed that there might be one wind turbine prediction model that is overachieving on the bad timeline of another wind turbine. Therefore, we can't use any one of the wind turbine's prediction model for other wind turbines' performance prediction. This supports our claim that each wind turbine (or equipment) is unique and has different feature settings, hence they can't be considered as homogeneous. Therefore, we have decided to predict power output for each wind turbine separately.

	test-wt-bad2	test-wt-bad3	test-wt-bad4	test-wt-bad5	test-wt-bad6	test-wt-bad7	test-wt-bad8	test-wt-bad9	test-wt-bad13	test-wt-bad14	test-wt-bad15	test-wt-bad16	test-wt-bad17
2	132.600	171.730	194.434	197.432	192.764	156.987	299.866	158.547	151.569	141.448	151.209	188.362	142.763
3	79.150	109.346	116.926	122.524	100.906	88.839	198.096	91.730	80.404	83.317	83.843	120.622	77.459
4	84.717	103.370	120.111	134.740	128.929	101.588	217.482	94.160	95.326	91.251	96.341	117.878	92.910
5	191.687	211.523	271.145	217.285	340.564	229.893	273.626	226.115	210.835	239.394	250.235	253.653	280.653
6	139.472	199.256	235.854	169.015	215.261	164.135	210.676	175.423	150.232	159.602	162.726	227.976	183.756
7	179.023	198.777	242.174	204.156	294.108	210.454	261.672	208.445	198.810	226.047	222.308	230.125	258.690
8	146.183	160.035	208.813	171.127	247.240	165.313	191.584	167.880	159.400	180.986	185.014	200.347	206.268
9	145.951	158.951	195.462	196.159	242.880	175.173	271.268	162.612	162.253	167.724	186.291	186.587	197.287
13	141.939	152.993	179.288	190.639	207.069	164.770	280.657	157.190	151.070	152.582	166.121	163.264	169.269
14	115.811	132.728	143.023	158.781	150.205	125.007	282.793	127.250	119.073	120.425	123.621	149.449	119.958
15	184.911	213.498	246.173	204.093	275.839	215.390	288.102	212.490	193.492	217.131	216.627	240.540	237.778
16	109.163	161.660	187.781	172.412	169.308	126.484	199.414	133.874	123.189	126.110	127.427	171.335	130.848
17	110.880	136.300	156.617	169.633	171.182	134.979	251.342	128.701	127.698	115.477	132.279	145.124	131.876

Figure 13. Ensemble Prediction Results over all bad timeline test datasets (best and worst values are highlighted per row/column as dark green: best in row, red: worst in row, light green: best in column, orange: worst in column).

3.2. Deterioration Detection

For deterioration detection, we have to find the balance between good performance on the good timeline and bad performance on the bad timeline. This needs to be achieved over a range of wind turbines (or other equipment) individually. The ensemble prediction method shows the most robust performance and therefore is our time series prediction method of choice.

For deterioration detection, we use a continuous time period with good and bad timelines to test the performance of our approach. Figure 14 is showing an example test timeline for wind turbine no. 15 that has good and bad timelines. The bad timelines are labeled by the value 1 indicated by the orange line on the secondary y-axis. All labels were manually assigned according to the available information from the data provider.

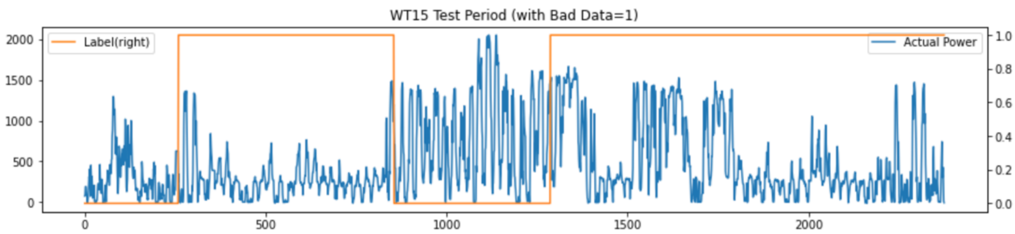


Figure 14. Test Data for Wind Turbine 15 with Labels for deterioration detection.

Figure 15 shows the power output prediction results (orange line) in regard to the actual power output (blue line). For overlaps between the actual and the predicted power, the line appears slightly grey. The predicted power outputs are obtained for each wind turbine based on the workflow presented in Figure 6 using the ensemble prediction method. We can observe that the RMSE (red line, scale on secondary y-axis), calculated for 24-h windows, is generally lower in the very beginning and in the middle of the presented test timeline. This observation matches with the time periods that we have identified as “good timelines” based on our domain knowledge.

To detect deterioration in the performance of a wind turbine, we want to establish a prediction model performance baseline using RMSE or RMSPE metrics leveraging validation data, i.e., a known good timeline dataset that wasn’t used during model training. This baseline can then be compared to the prediction performance of a test dataset where the real wind turbine output performance is unknown. Figure 16 shows an example of this process for wind turbine no. 15. The first graph shows the validation data and its RMSPE (green line on secondary y-axis). Once RMSPE value is calculated for each 24-h window hence the green line is cut off early in the plot and RMSE can be calculated in the same manner. Then, we use the obtained RMSE/RMSPE values to calculate value cutoff limits using simple statistics. For this research, the cutoff limits are calculated individually for each wind turbine prediction model using the 90% percentile RMSE value and the 95% percentile RMSPE value based on the validation data. Thus, dynamic cutoff limits can be calculated based on the prediction quality for each wind turbine. The cutoff limits are then used to identify if an RMSE/RMSPE value observed for the test data suggests a wind turbine performance deterioration or not. Examples of this can be seen in Figure 16 for wind turbine no. 15, where the second plot shows the RMSE based deterioration labels in red and true labels in purple. The third plot shows similar information for RMSPE with true labels in purple and identified labels in green. The two plots in the last row of Figure 16 show the confusion matrices for RMSE and RMSPE respectively. The plots give us insights into the performance deterioration classification accuracy of the two metrics. Our proposed performance deterioration approach performs better when we use the RMPSE metric. We observe 100% precision, meaning that if the wind turbine is in a good state, we are always able to classify this correctly. There are cases when the method is unable to detect deterioration correctly, especially when the expected power output is in a lower value range or at the beginning/end of a bad performing time period. Performance deterioration might therefore be harder to detect when wind speed and actual power output are low. We conclude that the impact of deterioration is felt more during times where the expected power output is supposed to be high.

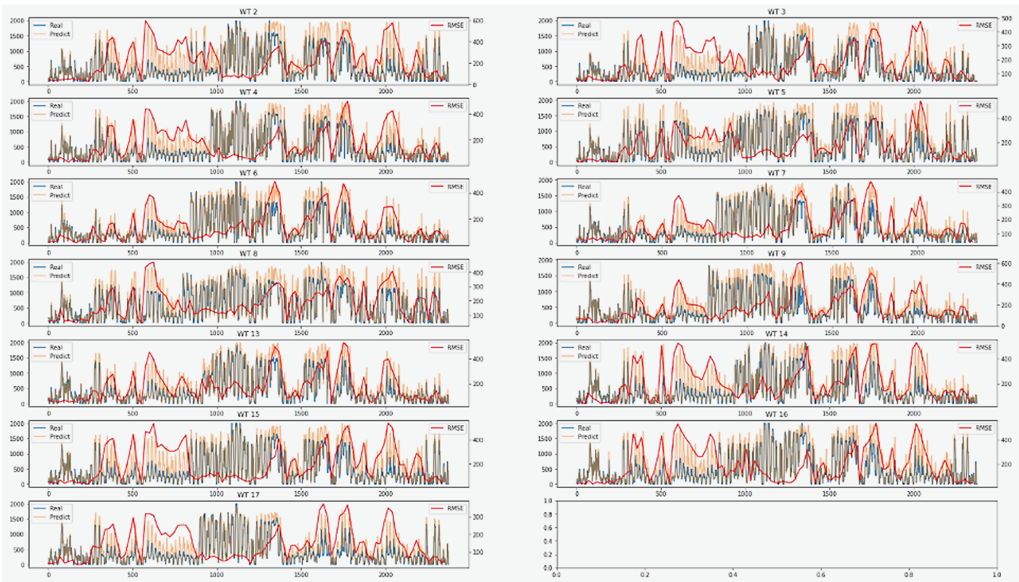


Figure 15. Power Output Prediction Results for all Wind Turbines.

Figure 17 shows a second deterioration detection example for wind turbine no. 8, which we observed previously as the most difficult wind turbine to achieve good prediction results for as shown in Figures 12 and 13. For this wind turbine, the classification accuracy for the deterioration task is slightly lower. The observed F1-score is 0.631 for the RMSPE metric and the lower score is observed because less of the bad performing time periods are actually identified as bad and some good performing time periods also have been wrongly identified as bad. In the following Table 3, we calculate the overall weighted F1 classification score to account for the imbalance in the classes for each wind turbine to understand the overall performance of our proposed method. In addition to the RMSE and RMSPE metrics, we also added a mixed metric version where a bad time period is identified if either RMSE or RMSPE value cutoff limits are triggered. Overall, we observe the highest classification performance for the mixed metric that outperforms the other two approaches in 10 out of 13 cases. In the other cases, the performance is equal or slightly worse compared to the RMSPE approach. The RMSE approach never performs the best and only beats the RMSPE approach for wind turbine no. 3 and 8. The reason why the relative percentage error like RMSPE works better is that the bad performing time periods include expected low value power outputs that are harder to detect when only using statistically calculated cutoff limits based on an absolute value. Therefore, our conclusion is to use the mixed approach for getting the best performance in deterioration detection.

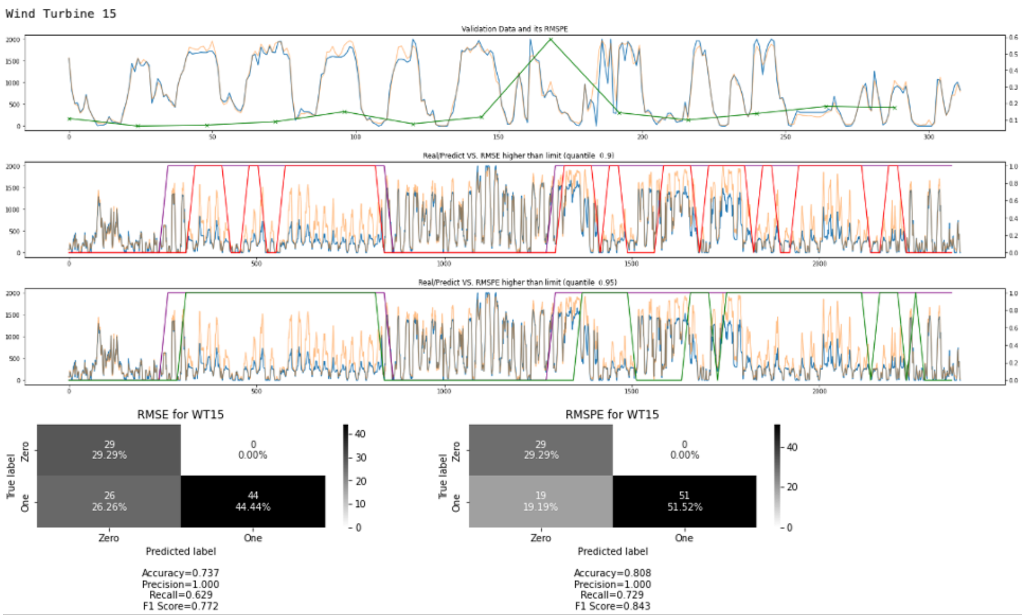


Figure 16. Deterioration Detection for Wind Turbine 15 where (**a**) denotes the true (blue line) vs. predicted (orange line) power output for a validation dataset with RMSPE for groups of samples (green line), and for a test dataset: (**b**) the real (purple line) and deterioration labels using RMSE (red line), (**c**) the real (purple line) and deterioration labels using RMSPE (green line), (**d**) confusion matrix based on RMSE and (**e**) confusion matrix based on RMSPE.

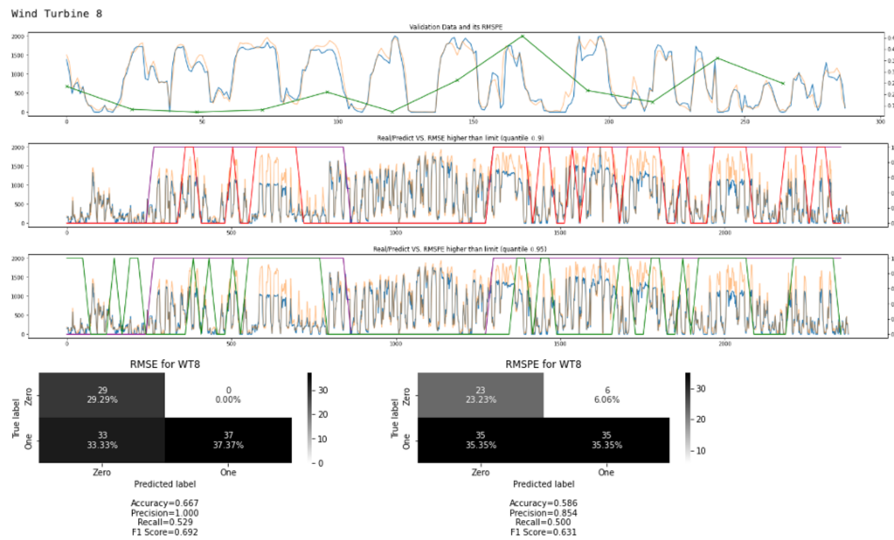


Figure 17. Deterioration Detection for Wind Turbine 8 where (**a**) denotes the true (blue line) vs. predicted (orange line) power output for a validation dataset with RMSPE for groups of samples (green line), and for a test dataset: (**b**) the real (purple line) and deterioration labels using RMSE (red line), (**c**) the real (purple line) and deterioration labels using RMSPE (green line), (**d**) confusion matrix based on RMSE and (**e**) confusion matrix based on RMSPE.

Table 3. Weighted F1-Score Results for RMSE, RMSPE, and Mixed Metric.

Wind Turbine	RMSE	RMSPE	Mixed
2	0.661	0.737	**0.783**
3	0.710	0.704	0.775
4	0.537	0.747	0.793
5	0.407	0.516	0.634
6	0.549	0.890	0.878
7	0.490	0.876	0.876
8	0.675	0.600	0.720
9	0.526	0.823	0.812
13	0.457	0.720	0.768
14	0.485	0.605	0.648
15	0.748	0.816	0.873
16	0.717	0.754	0.757
17	0.573	0.883	0.892

4. Discussion

In the following section, we will discuss the results and benefits of the proposed method. In summary, we have proposed a wind turbine output prediction modeling and deterioration detection method that takes data from a wind turbine and then trains a prediction model to conduct deterioration detection on newly observed data. While the approach processes each wind turbine separately, the workflow is automatized in a manner where fine-tuning to the needs of each turbine is not necessary.

The results for the power output prediction show that not all wind turbines achieve the same prediction performance due to differences in their feature settings even though the turbines were physically placed in a similar environment within a short distance from each other. It is therefore important to create separate prediction models for each wind turbine. Furthermore, our previous discussion based on Figure 10 showed that it is hard to choose a specific time series prediction method, LSTM or FNN, over another for all the wind turbines and an ensemble is a good approach to stabilize the prediction results achieving good prediction accuracy on the good timelines and distinguishably worse prediction performance on the bad timelines.

For the deterioration detection, we have shown an overview of all wind turbines' power output prediction performance plotted for a continuous test dataset in Figure 15. The overall periods of good and bad performance are similar for the deterioration detection of all wind turbines showing that the wind turbines are operating under similar conditions. On the other hand, our experiments showed the uniqueness of each wind turbine. This makes us confident that there is value in using our ensemble model for prediction to detect deterioration of wind turbine individually. The overall F1-scores as shown in Table 3 are between 0.634 and 0.892 for all wind turbines for a mixed error metrics. In addition, it is not necessary for deterioration detection to identify all labels correctly. If we can keep the number of false positives low during "good timeline" data periods while detecting continuous phases of bad performance during "bad timeline" periods, it will be possible for a wind turbine operator to use these results to improve overall wind turbine power output in a quick manner. Even a fraction of an improvement will result in an improved business value.

Our work can identify deteriorating performance in wind turbines therefore enabling better planning for the maintenance and repairs of wind turbines. The proposed deterioration detection approach can also be applied to other equipment and help improve the performance of those industry machinery. Running machines at lower efficiency leads to faster wear and tear as well as revenue loss, and our work supports to mitigate these risks. The obtained results are promising and it will be interesting to apply our work to other machines/equipment. The accurate prediction of power output over a group of wind turbines using the same method is challenging because of the differences in external and internal factors and how they affect the wind turbines. Currently, data is limited, and training/validation has to be conducted on a relatively small dataset, especially for the second part of deterioration detection. More experiments need to be done over a longer horizon to understand the performance better. Also, it is essential to consider the implications of reducing or increasing the number of features in wind turbine deterioration analysis. A sensitivity analysis could provide valuable insights into the impact of different features on the model's performance. For instance, determining which features have the most significant influence on deterioration predictions could help prioritize maintenance efforts. Additionally, conducting bootstrapping or cross-validation for parameter setting can enhance the model's robustness and reliability. These techniques can help ensure that the deterioration analysis accurately reflects the real-world behavior of wind turbines, thereby improving maintenance planning and overall operational efficiency.

We shared our comprehensive analysis with Atria Power, the owner of the wind turbine farm, and they found it to be valuable. The insights derived from our analysis align closely with the operational dynamics of their wind turbines, providing them with actionable intelligence to enhance their operations. Atria Power is excited about the prospect of implementing these insights and is eager to explore how our findings can be integrated into their existing processes. They perceive a significant potential in this collaboration, recognizing the opportunity to improve efficiency, reliability, and overall performance of their wind turbine operations.

5. Conclusions

We have demonstrated how to efficiently predict the power output of a wind turbine using SCADA data and then to leverage these predictions to detect possible performance deterioration. This analysis can be extended to any machine and helps to proactively maintain such equipment in a healthy state, to run them optimally and to get the best performance while maximizing profit for the business/company.

Our presented results showed that we were able to predict deterioration for all wind turbines where the classification F1-score was between 0.634 and 0.892. We achieved this with an automatized workflow that can create a prediction model based on training data, predict outputs for a validation dataset to calculate performance cutoff limits, and then label newly observed test data as good or bad performing. Although we did some initial trial and error to achieve a stable prediction performance over all wind turbines, we refrained from fine-tuning/specializing prediction models to the needs of specific wind turbines or do anything that might make application in a real-world business scenario difficult. Thus, the our approach demonstrates to be a valuable asset in wind farm maintenance, offering additional insights to enhance the current maintenance strategy, which relies on scheduled inspections and alarms triggered by exceeding static threshold values.

For future work, we aim to extend the scope of our research by exploring additional deep neural networks and time series prediction approaches to further enhance our approach. Extending our analysis to encompass broader class of machinery will allow us to generalize our findings and contribute to a more comprehensive understanding of predictive maintenance. Additionally, incorporating a Functional Basis Neural Network with Fourier or wavelet functions holds promise for improving modeling results, warranting further investigation. We plan to incorporate our work to more Atria Power Wind Farms as next steps. Another important avenue is understanding the important features and their impact. Finally, expanding our time scope beyond six months will enable us to capture and account for all seasonalities, other kinds of bad states, providing a more comprehensive assessment of performance over time.

Author Contributions: Conceptualization, J.B., A.R.R., C.V. and C.G.; Methodology, J.B., A.R.R. and C.V.; Validation, J.B., A.R.R., A.P. and A.V.K.; Formal analysis, J.B. and A.R.R.; Investigation, J.B. and A.R.R.; Writing—original draft, J.B. and A.R.R.; Writing—review & editing, J.B., A.R.R., C.V., A.P., A.V.K. and C.G.; Supervision, C.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets presented in this article are not readily available because the data are part of an ongoing study or is not permitted to share readily with everyone, etc. Requests to access the datasets should be directed to Atria Power.

Acknowledgments: We would like to extend our sincere gratitude to Atria Power for their invaluable contribution of data and insights that greatly enriched our research. Their willingness to share their expertise and resources has been instrumental in enhancing the depth and quality of our work. We are truly thankful for their collaboration and support, which has been pivotal in achieving our objectives.

Conflicts of Interest: Jana Backhus, Aniruddha Rajendra Rao, Chandrasekar Venkatraman, and Chetan Gupta were employed by Hitachi America, Ltd. A. Vinoth Kumar was employed by Atria Brindavan Power Private Limited. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DNN	Deep Neural Network
FDA	Functional Data Analysis
FLM	Functional Linear Model
FNN	Functional Neural Network
TS	Time Series
LSTM	Long Term Short Memory
RMSE	Root Mean Squared Error
RMSPE	Root Mean Squared Percentage Error
RNN	Recurrent Neural Network
SCADA	Supervisory control and data acquisition

References

1. Gernaat, D.E.; de Boer, H.S.; Daioglou, V.; Yalew, S.G.; Müller, C.; van Vuuren, D.P. Climate change impacts on renewable energy supply. *Nat. Clim. Chang.* **2021**, *11*, 119–125. [CrossRef]
2. Koch, H.; Vögele, S.; Hattermann, F.F.; Huang, S. The impact of climate change and variability on the generation of electrical power. *Meteorol. Z.* **2015**, *24*, 173–188. [CrossRef]
3. Porté-Agel, F.; Bastankhah, M.; Shamsoddin, S. Wind-Turbine and Wind-Farm Flows: A Review. *Bound.-Layer Meteorol.* **2019**, *174*, 1–59. [CrossRef] [PubMed]
4. Pryor, S.C.; Barthelmie, R.J.; Bukovsky, M.S.; Leung, L.R.; Sakaguchi, K. Climate change impacts on wind power generation. *Nat. Rev. Earth Environ.* **2020**, *1*, 627–643. [CrossRef]
5. Veers, P.S.; Dykes, K.; Lantz, E.J.; Barth, S.; Bottasso, C.L.; Carlson, O.; Clifton, A.; Green, J.B.; Green, P.; Holttinen, H.; et al. Grand challenges in the science of wind energy. *Science* **2019**, *366*, eaau2027. [CrossRef] [PubMed]
6. Yang, W.; Tavner, P.J.; Crabtree, C.J.; Feng, Y.; Qiu, Y. Wind turbine condition monitoring: Technical and commercial challenges. *Wind Energy* **2014**, *17*, 673–693. [CrossRef]
7. Narasinh, V.; Mital, P.; Chakravortty, N.; Mittal, S.; Kulkarni, N.; Venkatraman, C.; Rajakumar, A.G.; Banerjee, K. Investigating power loss in a wind turbine using real-time vibration signature. *Eng. Fail. Anal.* **2024**, *159*, 108010. [CrossRef]
8. Vasic, S.; Orosnjak, M.D.; Brkljac, N.; Vrhovac, V.; Ristic, K. Identification of Criteria for Enabling the Adoption of Sustainable Maintenance Practice: An Umbrella Review. *Sustainability* **2024**, *16*, 767. [CrossRef]
9. Franciosi, C.; Iung, B.; Miranda, S.; Riemma, S. Maintenance for Sustainability in the Industry 4.0 context: A Scoping Literature Review. *IFAC-PapersOnLine* **2018**, *51*, 903–908. [CrossRef]
10. Jin, X.; Xu, Z.; Qiao, W. Condition Monitoring of Wind Turbine Generators Using SCADA Data Analysis. *IEEE Trans. Sustain. Energy* **2021**, *12*, 202–210. [CrossRef]
11. Tautz-Weinert, J.; Watson, S.J. Using SCADA data for wind turbine condition monitoring—A review. *Iet Renew. Power Gener.* **2017**, *11*, 382–394. [CrossRef]
12. Murgia, A.; Verbeke, R.; Tsiporkova, E.; Terzi, L.; Astolfi, D. Discussion on the Suitability of SCADA-Based Condition Monitoring for Wind Turbine Fault Diagnosis through Temperature Data Analysis. *Energies* **2023**, *16*, 620. [CrossRef]
13. Santolamazza, A.; Dadi, D.; Introna, V. A Data-Mining Approach for Wind Turbine Fault Detection Based on SCADA Data Analysis Using Artificial Neural Networks. *Energies* **2021**, *14*, 1845. [CrossRef]

14. Astolfi, D.; Pandit, R.K.; Terzi, L.; Lombardi, A. Discussion of Wind Turbine Performance Based on SCADA Data and Multiple Test Case Analysis. *Energies* **2022**, *15*, 5343. [CrossRef]
15. Astolfi, D.; De Caro, F.; Vaccaro, A. Condition Monitoring of Wind Turbine Systems by Explainable Artificial Intelligence Techniques. *Sensors* **2023**, *23*, 5376. [CrossRef] [PubMed]
16. Rao, A.R.; Reimherr, M.L. Nonlinear Functional Modeling Using Neural Networks. *J. Comput. Graph. Stat.* **2021**, *32*, 1248–1257. [CrossRef]
17. Rao, A.R.; Reimherr, M.L. Modern non-linear function-on-function regression. *Stat. Comput.* **2021**, *33*, 1–12. [CrossRef]
18. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef]
19. Zhang, W.; Yang, D.; Wang, H. Data-Driven Methods for Predictive Maintenance of Industrial Equipment: A Survey. *IEEE Syst. J.* **2019**, *13*, 2213–2227. [CrossRef]
20. Zio, E. *Prognostics and Health Management of Industrial Equipment*; GIG: Hershey, PA, USA, 2013; p. 24
21. Maldonado-Correa, J.L.; Martín-Martínez, S.; Artigao, E.; Gómez-Lázaro, E. Using SCADA Data for Wind Turbine Condition Monitoring: A Systematic Literature Review. *Energies* **2020**, *13*, 3132. [CrossRef]
22. Pandit, R.K.; Astolfi, D.; Hong, J.; Infield, D.; Santos, M. SCADA data for wind turbine data-driven condition/performance monitoring: A review on state-of-art, challenges and future trends. *Wind Eng.* **2022**, *47*, 422 – 441. [CrossRef]
23. Wang, H.; Zhao, X.; Wang, W. Fault diagnosis and prediction of wind turbine gearbox based on a new hybrid model. *Environ. Sci. Pollut. Res.* **2022**, *30*, 24506–24520. [CrossRef]
24. Wang, M.; Peng, J.; Luo, Y.; Shen, Z.; Yang, H. Comparison of different simplistic prediction models for forecasting PV power output: Assessment with experimental measurements. *Energy* **2021**, *224*, 120162. [CrossRef]
25. Dolara, A.; Leva, S.; Manzolini, G. Comparison of different physical models for PV power output prediction. *Sol. Energy* **2015**, *119*, 83–99. [CrossRef]
26. Theocharides, S.; Makrides, G.; Georghiou, G.E.; Kyprianou, A. Machine learning algorithms for photovoltaic system power output prediction. In Proceedings of the 2018 IEEE International Energy Conference (ENERGYCON), Limassol, Cyprus, 3–7 June 2018; pp. 1–6.
27. Massidda, L.; Marrocu, M. Use of multilinear adaptive regression splines and numerical weather prediction to forecast the power output of a PV plant in Borkum, Germany. *Sol. Energy* **2017**, *146*, 141–149. [CrossRef]
28. Backhus, J.; Kono, Y. Cooling Power Consumption Dependency Simulation for Real-World Data Center Data. In Proceedings of the 2022 International Conference on Software, Telecommunications and Computer Networks (SoftCOM), Split, Croatia, 22–24 September 2022; pp. 1–6.
29. Parmezan, A.R.S.; Souza, V.M.A.; Batista, G.E.A.P.A. Evaluation of statistical and machine learning models for time series prediction: Identifying the state-of-the-art and the best conditions for the use of each model. *Inf. Sci.* **2019**, *484*, 302–337. [CrossRef]
30. Wang, Q.; Wang, H.; Gupta, C.; Rao, A.R.; Khorasgani, H. A Non-linear Function-on-Function Model for Regression with Time Series Data. In Proceedings of the 2020 IEEE International Conference on Big Data (Big Data), Atlanta, GA, USA, 10–13 December 2020; pp. 232–239.
31. Song, X.; Liu, Y.; Xue, L.; Wang, J.; Zhang, J.; Wang, J.; Jiang, L.; Cheng, Z. Time-series well performance prediction based on Long Short-Term Memory (LSTM) neural network model. *J. Pet. Sci. Eng.* **2020**, *186*, 106682. [CrossRef]
32. Jin, X.B.; Yu, X.H.; Wang, X.; Bai, Y.T.; Su, T.; Kong, J. Prediction for Time Series with CNN and LSTM. In *Proceedings of the 11th International Conference on Modelling, Identification and Control (ICMIC2019)*; Springer: Singapore, 2019.
33. Lindemann, B.; Müller, T.; Vietz, H.; Jazdi, N.; Weyrich, M. A survey on long short-term memory networks for time series prediction. *Procedia CIRP* **2021**, *99*, 650–655. [CrossRef]
34. Han, Z.; Zhao, J.; Leung, H.; Ma, K.; Wang, W. A Review of Deep Learning Models for Time Series Prediction. *IEEE Sens. J.* **2019**, *21*, 7833–7848. [CrossRef]
35. Ramsay, J.O.; Silvermann, B.W. *Functional Data Analysis*, 2nd ed.; Springer: New York, NY, USA, 2005.
36. Kokoszka, P.; Reimherr, M.L. *Introduction to Functional Data Analysis*, 1st ed.; Chapman and Hall/CRC: New York, NY, USA, 2017.
37. Ferraty, F.; Romain, Y. The Oxford Handbook of Functional Data Analysis. In *Oxford Handbooks Online*; Oxford University Press: Oxford, UK, 2011.
38. ling Wang, J.; Chiou, J.M.; Müller, H.G. Functional Data Analysis. *Annu. Rev. Stat. Its Appl.* **2016**, *3*, 257–295. [CrossRef]
39. Reiss, P.T.; Goldsmith, J.; Shang, H.L.; Ogden, R.T. Methods for Scalar-on-Function Regression. *Int. Stat. Rev.* **2017**, *85*, 228–249. [CrossRef] [PubMed]
40. Rossi, F.; Conan-Guez, B.; Fleuret, F. Functional data analysis with multi layer perceptrons. In Proceedings of the 2002 International Joint Conference on Neural Networks, IJCNN'02 (Cat. No.02CH37290), Honolulu, HI, USA, 12–17 May 2002; Volume 3, pp. 2843–2848.
41. Liu, Y.; Cheng, H.; Kong, X.; Wang, Q.B.; Cui, H. Intelligent wind turbine blade icing detection using supervisory control and data acquisition data and ensemble deep learning. *Energy Sci. Eng.* **2019**, *7*, 2633–2645. [CrossRef]
42. Sagi, O.; Rokach, L. Ensemble learning: A survey. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2018**, *8*, e1249. . [CrossRef]
43. Lee, X.Y.; Kumar, A.; Vidyaratne, L.S.; Rao, A.R.; Farahat, A.K.; Gupta, C.R. An ensemble of convolution-based methods for fault detection using vibration signals. In Proceedings of the 2023 IEEE International Conference on Prognostics and Health Management (ICPHM), Montreal, QC, Canada, 5–7 June 2023; pp. 172–179.

44. Phyo, P.P.; Byun, Y. Hybrid Ensemble Deep Learning-Based Approach for Time Series Energy Prediction. *Symmetry* **2021**, *13*, 1942. [CrossRef]
45. Zhou, Z.H. *Ensemble Methods: Foundations and Algorithms*; CRC Press: Boca Raton, FL, USA, 2012.
46. Tian, Y.; Feng, Y. RaSE: Random subspace ensemble classification. *J. Mach. Learn. Res.* **2021**, *22*, 2019–2111.
47. Guan, D.; Yuan, W.; Lee, Y.K.; Najeebullah, K.; Rasel, M.K. A Review of Ensemble Learning Based Feature Selection. *IETE Tech. Rev.* **2014**, *31*, 190–198. [CrossRef]