Article

Data Enrichment as a Method of Data Preprocessing to Enhance Short-Term Wind Power Forecasting

Yingya Zhou [1] , Linwei Ma [1,] *, Weidou Ni [1] and Colin Yu [2]

1   State Key Laboratory of Power Systems, Department of Energy and Power Engineering, Tsinghua-BP Clean Energy Research and Education Centre, Tsinghua University, Beijing 100084, China
2   Chenqiao Smart Technology, Inc., Shanghai 201306, China
*   Correspondence: malinwei@tsinghua.edu.cn; Tel.: +86-10-62795734-8023

Abstract: Wind power forecasting involves data preprocessing and modeling. In pursuit of better forecasting performance, most previous studies focused on creating various wind power forecasting models, but few studies have been published with an emphasis on new types of data preprocessing methods. Effective data preprocessing techniques and the fusion with the physical nature of the wind have been called upon as potential future research directions in recent reviews in this area. Data enrichment as a method of data preprocessing has been widely applied to forecasting problems in the consumer data universe but has not seen application in the wind power forecasting area. This study proposes data enrichment as a new addition to the existing library of data preprocessing methods to improve wind power forecasting performance. A methodological framework of data enrichment is developed with four executable steps: add error features of weather prediction sources, add features of weather prediction at neighboring nodes, add time series features of weather prediction sources, and add complementary weather prediction sources. The proposed data enrichment method takes full advantage of multiple commercially available weather prediction sources and the physical continuity nature of wind. It can cooperate with any existing forecasting models that have weather prediction data as inputs. The controlled experiments on three actual individual wind farms have verified the effectiveness of the proposed data enrichment method: The normalized root mean square error (NRMSE) of the day-ahead wind power forecast of XGBoost and LSTM with data enrichment is 11% to 27% lower than that of XGBoost and LSTM without data enrichment. In the future, variations on the data enrichment methods can be further explored as a promising direction of enhancing short-term wind power forecasting performance.

Keywords: wind power forecasting; data enrichment; data preprocessing; weather prediction

1. Introduction

Wind energy is one of the world's most promising renewable energy resources. Nevertheless, wind energy is inherently intermittent, uncontrollable, and random. Short-term wind power forecasting is of significant interest for unit commitment and scheduling. Wind power forecasting involves data preprocessing and modeling. Many studies have been conducted on short-term wind power forecasting models, especially on models using artificial intelligence and hybrid techniques [1–6]. Meanwhile, data preprocessing is widely used, but few studies have focused on developing improved data preprocessing methods for wind power forecasting [7]. Instead, data preprocessing has often been mentioned as an auxiliary part of each wind power forecasting model, and has been very limited to either simple data cleaning and organization or individualized data preparation steps for specific wind power forecasting algorithms. A detailed literature review on the data preprocessing methods used in wind power forecasting are referred to Section 2.1.

While the authors kept trying different wind power forecasting algorithms from the literature to seek higher forecasting performance during the operation of actual wind

farms, it was observed that the forecast accuracy for some wind farms was influenced more significantly by inputting weather prediction data rather than the use of various forecasting models. Recent studies in the field have also called for more attention to be paid to data preprocessing [7,8]. It has been suggested that approaches to fully utilize the original datasets are worthy of future study. Moreover, insights from general earth system science highlighted that combining machine learning with relationships derived from the natural sciences should be a promising way to improve forecasting accuracy [9]. These phenomena and insights prompted the authors to explore new ways of data preprocessing to introduce more useful weather prediction information for better forecast accuracy.

Data enrichment refers to the process of appending or otherwise enhancing collected data, with the relevant context being obtained from additional sources [10]. Originating in the realm of processing consumer data [11], data enrichment is an increasingly popular approach to data preprocessing to substantially enhance the performance of various forecasting problems in business [12], industrial network security [10], and commodity price forecasting [13]. However, data enrichment has not been carried out in previous research on short-term wind power forecasting, as revealed by the literature search in Section 2.1. Inspired by the application of data enrichment in other areas, the authors were motivated to apply the principles of data enrichment to wind power forecasting.

This research aimed to develop a data enrichment method as a novel data preprocessing method to improve the short-term wind power forecasting performance of various forecast models. The proposed data enrichment method is physically interpretable, easy to execute in the day-to-day operation of engineering, and has demonstrated effectiveness in elevating the accuracy of wind power forecasting. The method has the following two characteristics:

- It can be combined with other data preprocessing methods and is also applicable to various modeling algorithms;
- It serves the purpose of adding as much valuable weather prediction information as possible into the inputs of the wind power forecasting models.

The main contributions of this study are as follows:

- To propose the concept and a methodological framework of data enrichment to improve short-term wind power forecasting performance;
- To put forward a set of executable data enrichment steps and validate the effectiveness of each step in improving wind power forecasting performance;
- To verify the general applicability of the proposed data enrichment method by cooperating with one machine learning and one deep learning short-term wind power forecasting models for three different actual wind farms.

Based on these contributions, other researchers in wind power forecasting can incorporate the proposed data enrichment method as a way of data preprocessing before applying their forecasting models or develop other ways of data enrichment for enhanced wind power forecasting performance.

The rest of the paper is organized as follows: Section 2 reviews the data preprocessing methods employed in previous research on wind power forecasting and the application of data enrichment in other areas. Section 3 introduces the concept and the step-by-step guidance of the proposed data enrichment method. Section 4 describes in detail the experimental setting and data. Section 5 presents the results and discussion of the verification of the proposed method's general effectiveness and the effectiveness of each step in the method. Section 6 reviews the limitations of the proposed method. Finally, Section 7 concludes the paper.

2. Literature Review

2.1. Data Preprocessing Methods Used in Wind Power Forecasting

In wind power forecasting, the preprocessing of raw data is necessary before the data can be utilized to train forecasting models. The terminology of data preprocessing,

however, means that different scopes have been applied in different studies and there have been ambiguous boundaries regarding the terms 'data preparation' and 'feature engineering'. Liu and Chen [7] summarized the data preprocessing methods applied in wind power forecasting into five categories: data decomposition, feature selection, feature extraction, denoising, and outlier detection. Lipu et al. [8] provided a list of various data preparation techniques during the review of the 140 most recent papers on wind power forecasting using artificial intelligence. In their review paper, data preparation was categorized into data preprocessing, data filtering, data sampling, downscaling, and outlier detection. Data preprocessing was then further categorized into data division, decomposition, standardization, and normalization. In another recent review on wind power forecasting with deep neural networks [6], data preparation was divided into signal processing and outlier detection approaches, with the sole objective of addressing uncertainty in wind data.

Table 1 provides a summary of the occasions of use and recent applications of various data preprocessing methods in wind power forecasting. Despite the diverse library of detailed data preprocessing methods, all methods can be divided into three groups according to their purpose:

- Data organization methods reorganize raw datasets into datasets of different sizes, scales, and sampling frequencies;
- Data cleaning methods correct abnormal data and errors;
- Dimensionality reduction methods reduce the number of features or transform the original features to shrink the feature space and prevent overfitting.

Table 1. Summary on data preprocessing methods and their recent research applications in wind power forecasting from the literature.

Purpose	Data Preprocessing Method	Occasion of Use and Recent Research Applications
Data organization	Data sampling	• Unifies the data sampling rates according to the forecast horizon [14,15] • Re-samples the continuous data after signal processing [16] • Mitigates the class imbalance problem [17]
	Data division/Data splitting	• Separates datasets into two sets: training and testing [18] • Divides datasets for individual model building [19]
	Data standardization/Data normalization	• Converts variables in the input dataset in different scales into those of the same scale [20,21]
	Data clustering	• Clusters similar input datasets into one group to reduce computational burden in modeling while maintaining characteristics of the datasets [22,23]
	Data decomposition	• Breaks down non-stationary original time series into several relatively stationary subseries, and then builds a forecasting model on each subseries [20,24,25]
	Data augmentation	• Enlarges training data in case of lacking input data for machine learning models [26,27]
Data cleaning	Data correction/denoising	• Deletes or substitutes abnormal data, noise, and missing values [16,28] • Corrects systematic errors in numerical weather prediction data [26,29]

Table 1. *Cont.*

Purpose	Data Preprocessing Method	Occasion of Use and Recent Research Applications
Dimensionality reduction	Data filtering	• Eliminates any possible systematic and random errors [30–33]
	Feature selection	• Selects useful features from feature candidates to reduce the complexity of forecasting models and prevents overfitting [28,34–36]
	Feature extraction	• Maps the original feature set into a new one to reduce dimensionality and prevent overfitting [37,38]

In all of the studies introduced in Table 1, data preprocessing was either mentioned incidentally as an auxiliary part of the wind power forecasting models or as part of a newly proposed specialized forecasting algorithm. It can also be observed from Table 1 that the data preprocessing methods used in previous studies took data only as pure time series signals without unfolding the physics behind the data.

Few studies have focused on developing a new kind of data preprocessing method. The results of the literature search with the keywords 'data preprocessing wind forecasting' or 'data processing wind forecasting' in ScienceDirect and MDPI revealed the gap. As is shown in Table 2, only eight journal articles fulfilled the search criteria, but none of these focused on developing new types of data preprocessing methods.

Table 2. Literature search results by relevant keywords in ScienceDirect and MDPI conducted on 15 December 2022.

Keywords	Number of Results [1]	Topic
Data preprocessing wind forecast or data processing wind forecast	7	All of the seven studies proposed a specific combination of hybrid algorithms. In the data preprocessing part, the authors in [39] proposed singular spectrum analysis (SSA) for data denoising and variational mode decomposition (VMD) as the data decomposition method. The authors in [40] proposed VMD as the data decomposition method and PSR as the feature extraction method. The authors in [41] proposed complete ensemble empirical mode decomposition (CEEMD) as the data decomposition method. The authors in [42] proposed a complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) as the data decomposition method. The authors in [43] proposed empirical wavelet transform (EWT) as the data decomposition method. The authors in [44] proposed combining CEEMD and VMD as the data decomposition method. The authors in [45] proposed combining ICEEMDAN and fuzzy time series as the data decomposition and feature extraction methods.
	1	A review paper of the data processing methods used in wind energy forecasting models [7].

[1] The number of results is the number of the search results where the article title or article keywords contains the search keywords.

Meanwhile, researchers have also recently observed the gap and called for attention to be paid to future research on data preprocessing methods in wind power forecasting. Liu and Chen [7] stressed that data preprocessing methods have not drawn much attention and the approaches to fully explore the original high-resolution dataset are worthy of future study. Lipu et al. [8] suggested that further studies were also required concerning effective data preparation strategies. However, how to utilize data more effectively was not clearly stated.

2.2. Data Enrichment

Within the internet and consumer data universe the concept of data enrichment has been extensively applied in practice. As D. Needham described in [46], data enrichment is to make data powerful by showing more context about the meaning of the data. Operationally,

data enrichment is the process of enhancing existing information by supplementing relevant context. Typically, data enrichment is achieved by using external data sources, but that is not always the case [47]. For example, a customer relationship management (CRM) system may store two data series: the customers' names and identification card numbers. Before customer recommendation algorithms are trained, the input customer data can be enriched by decoding birthday information from the identification card numbers or relating other information such as the addresses and hobbies of the customers to the customer names from external systems.

It is uncommon to see the concept of data enrichment applied to engineering problems. The literature search conducted on 15 December 2022 with the search keywords 'data enrichment forecast' in ScienceDirect and MDPI revealed only one entry: The authors in [13] published an approach in 2022 on enriching time series using domain-specific terms for forecasting agricultural commodity prices. The results indicated that data enrichment is promising in reducing the forecast uncertainty of agricultural commodity prices. Data enrichment has not been found in previous studies regarding wind power forecasting.

On the other hand, the rough direction of enriching valuable information and knowledge in inputs to enhance forecast accuracy was indicated by Markus Reichstein et al. in *Nature* in 2019 [9]. They pointed out that classical machine learning approaches rarely exploit spatial–temporal dependencies exhaustively when dealing with data-driven earth system science. For example, previous time steps and neighboring grid cells contain hidden information on the state of the system, i.e., the 'memory effect'. It was proposed to tackle the critical challenges in earth system science by extracting knowledge from the data deluge and combining machine learning research with physically based relationships derived from the natural sciences.

Therefore, it is of value to introduce data enrichment to wind power forecasting.

2.3. Summary

Data preprocessing is an important step for wind power forecasting. The data preprocessing methods applied in previous wind power forecasting studies can be clustered into three categories by their purpose: data organization, data cleaning, and dimensionality reduction. All of these categories take data only as pure time series signals. Few studies have focused on developing new types of data preprocessing methods.

Data enrichment has been widely applied in the consumer data area to enhance forecast performance by enriching the valuable information content of the input data. However, data enrichment was not discovered in earlier studies on wind power forecasting. The authors were motivated by the current lack of consideration of the physical traits of wind in the current data preprocessing methods and the idea of data enrichment. Consequently, this study's emphasis was on bringing data enrichment to the field of wind power forecasting with consideration given to the physical features of wind before that data was put into forecasting models.

3. Wind Data Enrichment Method

3.1. Concept

Data enrichment aims to introduce more valuable information into the input dataset. Wind power forecasting has a strong correlation with weather prediction. In this study, weather prediction data were enriched by taking advantage of multiple commercial weather prediction sources such as IBM and the European Centre for Medium-Range Weather Forecasts (ECMWF), as well as the intrinsic characteristics of wind and wind prediction. The method made available the inherent forecast error feature of different weather prediction sources as well as the spatial and temporal continuity of the atmosphere. This allowed the wind power forecasting models to learn, hence improving the accuracy of wind power forecasting. This method is easy to execute and of broad applicability, so that it can be deployed with various existing wind power forecasting models using weather prediction data as inputs.

Figure 1 shows how the proposed method differs from the data preprocessing strategies summarized in Table 1. In fact, the proposed method may well position itself as a step before the original data preprocessing and forecasting models. The effectiveness of the proposed data enrichment method was verified by three different wind farms and two different types of classical wind power forecasting models in the following sections of this paper.
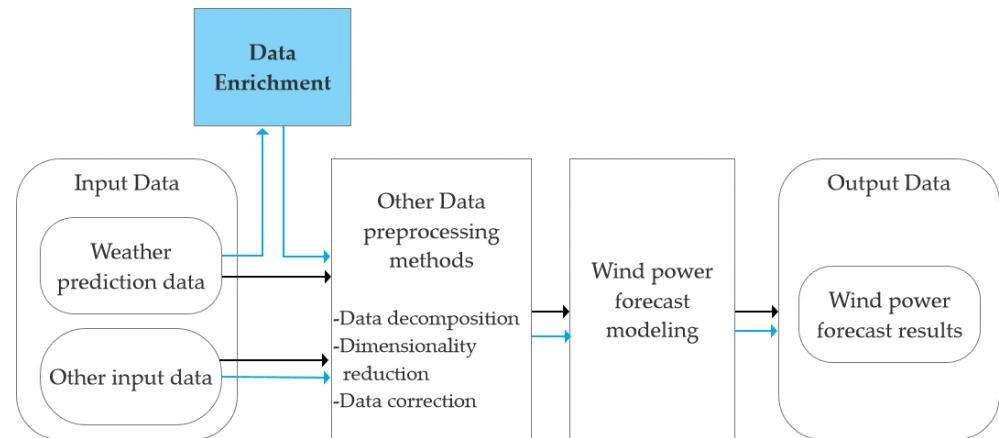

Figure 1. Illustration [1] of the relationship between the proposed data enrichment method, other incumbent data preprocessing methods, and wind power forecasting models. ([1] The black lines represent the data flow without data enrichment; The blue lines represent the data flow with the proposed data enrichment method).

3.2. Overall Framework

The proposed data enrichment method, as illustrated in Figure 2, can be broken down into four steps:

- Add error features of the weather prediction sources;
- Add features of neighboring weather prediction data nodes to take advantage of the spatial continuity of the atmosphere;
- Add time series features of the weather prediction sources to take advantage of the temporal continuity of the atmosphere;
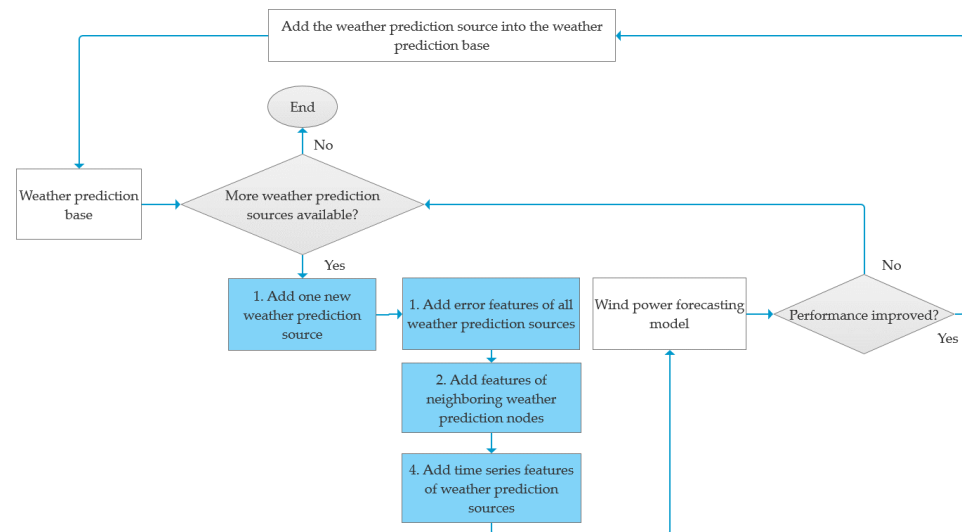- Add multiple complementary weather prediction sources.

Figure 2. The framework of the data enrichment method.

The detailed concept for each of the four steps is introduced in Sections 3.3–3.6, with exemplary equations to realize the steps proposed.

3.3. Add Error Features of Weather Prediction Sources

"The error is the feature", Schön et al. from the German weather forecast agencies shared in their publication [48]. They successfully forecasted lightning by taking the historical error between the results given by two-dimensional optical flow algorithms and the truth value measured from meteorological satellites as a feature fed into the machine learning forecast model. Similarly, the historical prediction error indicates an important feature of each weather prediction source relative to the weather truth value. For example, some weather prediction sources have better accuracy for rainy days, while others are good at predicting calm weather. Therefore, for each weather prediction source, new features representing the historical weather prediction error can be built as additional inputs for wind power forecasting models. The simplest way to represent the error feature is the historical difference between weather prediction and weather true value (WTV) according to Equation (1):

$$\delta y(t) = \hat{y}(t) - y(t) \tag{1}$$

where y and $\hat{y}$ represent the true value and prediction value of each physical quantity included in weather prediction data, such as wind speed, air temperature, air pressure, and air density.

3.4. Add Features of Weather Prediction at Neighboring Nodes

Commercial weather prediction data divides the real atmosphere into a grid and performs the prediction at each grid node. Since the atmosphere is a continuous physical body, the physical quantities at neighboring nodes interact with each other. For example, the wind speed at an adjacent node in the opposite direction of the wind speed may largely influence the wind speed at the current node for the next moment. Accordingly, the weather prediction data at neighboring nodes contain valuable information for the weather prediction of the present nodes. Recent studies have also shown the importance of the meteorological information of adjacent areas to the forecast of a target area by a data-driven approach [49,50]. However, for two reasons the weather prediction at adjacent nodes cannot be directly added as the inputs of wind power forecasting models. On the one hand, sole weather prediction data of neighboring nodes miss the information on the relative position of adjacent nodes to the present nodes. On the other hand, excessive inputs could lead to overfitting. A simple way to combine information on the relative location and the weather prediction values of neighboring nodes into new features can be conducted in the following procedures:

- Determine the adjacent nodes: As Figure 3 shows, for models forecasting at the wind turbine level the adjacent nodes are the eight nodes surrounding the box where the turbine is located. For models forecasting at the wind farm level, the adjacent nodes can be the eight areas surrounding the wind farm area. Each adjacent node is as large as the wind farm area. Neighboring nodes are selected based on physical proximity to reflect the physical continuity of the atmosphere;
- Calculate the combined weather prediction feature of each adjacent node: If the weather variable is a vector, such as wind speed, the new feature can be formulated by projecting the vector in the direction from the center of the wind farm node to the center of the adjacent node, as Equation (2) shows. If the weather variable is a scalar, the new feature can be formulated by calculating the gradient of the scalar first and then transforming the gradient as the other vector variables.

$$\hat{y}_{project,i} = |\vec{\hat{y}}_{adjacent,i}| \cdot \cos \theta_{project,i} \tag{2}$$
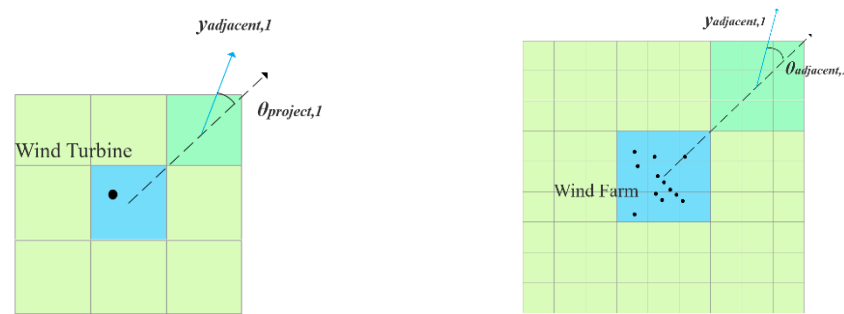
Figure 3. Illustration [1] on the adjacent node. ([1] The blue area is the forecasting node where the wind turbine or wind farm is located; the green areas are the adjacent nodes).

Here, $\hat{y}_{\text{project},i}$ is the new feature representing the influence from the weather condition at the ith neighboring node, $\vec{\hat{y}}_{\text{adjacent},i}$ is the weather variable vector or the gradient vector of scalar weather variable of the ith adjacent node, and $\theta_{\text{project},i}$ is the angle between $\vec{\hat{y}}_{\text{adjacent},i}$ and the vector from the center of the wind farm node to the center of the ith adjacent node.

3.5. Add Time Series Features of Weather Prediction Sources

Since the atmosphere is also continuous over time, the historical and future weather data have a strong relationship with the current weather data. The addition of time series features into the inputs of wind power forecasting models can extract valuable upstream and downstream information regarding the weather along the time into the models. Typical time series features include lag features, difference features, and rolling window features [51,52].

A lag feature is a variable that contains data from previous time steps. The past values are called lags, so t-1 is lag 1, and t-2 is lag 2. A lag N feature can be made according to Equation (3) [52]:

$$y_{\text{lagN}}(t) = y(t - N) \tag{3}$$

where y_{lagN} is the lag N feature of y, N means N times the time stamp.

A difference feature is a variable that contains the difference between the present and historical data points. A difference N feature can be calculated according to Equation (4) [52]:

$$y_{\text{diffN}}(t) = y(t) - y(t - N) \tag{4}$$

where y_{diffN} is the difference N feature of y.

Rolling means creating a rolling window with a specified size and performing calculations on the data in this window which rolls through the data. The rolling mean feature can be calculated according to Equation (5) [52]:

$$y_{\text{rollN,mean}}(t) = [y(t) + y(t - 1) \ldots + y(t - N + 1)]/N \tag{5}$$

where $y_{\text{rollN,mean}}$ is the rolling mean of y with a rolling window size of N.

3.6. Add Complementary Weather Prediction Sources

Multiple worldwide weather prediction sources are available for commercial use; some of these are listed in Table 3. They differ in a variety of settings, such as model principle, whether they are deterministic or probabilistic, their spatial and temporal resolutions, and other parameters. For different locations at different times, different weather prediction sources perform differently. Dabernig [53] compared the weather prediction of the ECMWF (European Centre for Medium-Range Weather Forecasts), ZAMG (Zentralanstalt für Meteorologie und Geodynamik—the Austrian national weather service), and GEFS (North American Global Ensemble Forecast System) as inputs for the statistical wind power forecasts of seven turbines located in Austria, Germany, and the Czech Republic. The

comparison showed that while the wind power forecast with GEFS as the input was always worse for all turbines, the wind power forecast with ECMWF as the input had a better performance than that with ZAMG for some turbines; however, this was worse for the other turbines.

Table 3. Examples of commercial weather prediction sources.

Abbreviation	Description of the Weather Prediction Sources
IBM	Global high-resolution atmospheric forecasting by IBM Weather Operations Center
GFS	Global forecast system operated by the United States' National Weather Service
ECMWF	European Centre for Medium-Range Weather Forecasts
CMC	Canadian Meteorological Centre
DWD	Deutscher Wetterdienst

Accordingly, multiple weather prediction sources can be utilized to improve the accuracy of wind power forecasting. However, the added weather prediction sources must be complementary to the current weather prediction source(s) as inputs to wind forecast models. Otherwise, adding a new weather prediction source could only result in more inaccurate information in the input, leading to a lower accuracy of wind power forecasting.

The detailed process to add complementary weather prediction sources is as follows:

- Calculate the average forecast accuracy and forecast data availability for the problem period. The problem period can be all the historical periods when the weather prediction and WTV for the problem location are available, or for artificial intelligence algorithms, the time covering the training and test dataset. The average forecast accuracy can be measured by its root mean square error (RMSE):

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{T} (\hat{y}(t) - y(t))^2}{T}} \tag{6}$$

where $\hat{y}(t)$ and $y(t)$ are the prediction and true values of weather variables, and T is the number of the data points;

- Select weather prediction source(s) that the wind power forecasting model used to set as inputs, or one weather prediction source with the lowest RMSE and highest data availability, as the first weather prediction source in the weather prediction base. Calculate the accuracy of the wind power forecasting model as $\text{RMSE}_{\text{before}}$;
- Add one more weather prediction source to the wind power forecasting model as a new input. Calculate the accuracy of the wind power forecasting model with the new weather prediction base as $\text{RMSE}_{\text{after}}$. If $\text{RMSE}_{\text{after}} < \text{RMSE}_{\text{before}}$, the newly added weather prediction source is then set as one of the sources in the weather prediction base;
- Repeat the previous step until all of the available weather prediction sources have been tried. An optimal weather prediction base is then set in which all the available complementary weather prediction sources are contained.

4. Experimental Setting

4.1. Baseline Wind Power Forecasting Models

For the sake of proving the efficacy of the data enrichment method presented in Section 3, the accuracy of wind power forecasting with and without the data enrichment method in the step of data preprocessing had to be compared to see whether the data enrichment had brought any 'added value' to the incumbent wind power forecasting models. To make the comparison of broad applicability and show the effect of the data enrichment method, two different types of classical and established wind power forecasting models were selected as benchmarks instead of distinctive derivatives: extreme gradient boosting (XGBoost) was chosen to represent machine learning algorithms for wind power

forecasting. Long short-term memory (LSTM) served as the representative of deep learning wind power forecasting models.

XGBoost, proposed in 2016 [54], is one of the most efficient and popular implementations of gradient boosting algorithms. It has been widely applied in wind speed forecasting [55] as well as wind power forecasting at the wind turbine level [56], wind farm level [57] and country level [58], all with demonstratable outstanding model performance.

LSTM was proposed in 1997 [59]. As one of the improved variants of the recurrent neural network, LSTM can capture the dynamic behavior of time series and has proven performance in forecasting various time series [60–62], wind speed forecasting [63] and wind power forecasting [64,65].

4.2. Datasets

Datasets of three different wind farms located in totally different cities that fed into the North China Power Grid were obtained to verify the general effectiveness of the proposed method. The details of the datasets are listed in Table 4.

Table 4. Description of three wind farm datasets.

Wind Farm	№ 1	№ 2	№ 3
Number of turbines	45	51	23
Dataset start time	2019/4/26 00:00:00	2019/4/26 00:00:00	2019/4/26 00:00:00
Dataset end time	2020/5/28 23:00:00	2020/5/28 23:00:00	2020/5/28 23:00:00

The dataset of each wind farm included 24-h-ahead weather prediction data from GFS, ECMWF, IBM, and CWC, the truth value of wind data measured by anemometer towers, and the truth value of wind farm power output from the supervisory control and data acquisition (SCADA) systems. The data were obtained at 1-h intervals. The variables from each weather prediction source and anemometer data were wind speed, wind direction, air temperature, air pressure, and air density. The data points from the previous 12 months from each wind farm were used as the training dataset, and the remaining data points from each wind farm were used as the test dataset. Consequently, for each wind farm, the training dataset contained 8784 time points for 366 days, while the test dataset contained 792 time points for 33 days. The original data had at least five significant digits and saved in double-precision floating-point format. All calculated data in the intermediate steps were saved in full length of the double-precision format.

The variables in the datasets were normalized according to Equation (7) to eliminate the influence of different scales:

$$\overline{y} = (y - \mu)/\sigma \tag{7}$$

where y is the original variable in the dataset, $\overline{y}$ is the normalized variable, μ is the mean of the original variable through the whole dataset, and σ is the standard deviation of the original variable through the whole dataset.

4.3. Performance Evaluation Metric

To date, several performance metrics have been employed to evaluate forecast accuracy, but no single performance metric has been recognized as the universal standard [66]. RMSE was chosen as the performance evaluation metric of wind power forecasting in this study due to its wide employment in academic research [67,68] and industry applications [69–71]. Since the three wind farms are of different installed capacities, the RMSE metric was normalized by capacity into NRMSE and calculated by Equation (8) [72]. The NRMSE is

also called root mean square relative error (RMSRE) in some literature [68]. Accuracy can be calculated as the complement of NRMSE, as Equation (9) shows:

$$\text{NRMSE} = \frac{\sqrt{\Sigma_{i=1}^{n} \left(P_i - \widehat{P_i} \right)^2}}{\text{Cap}\sqrt{n}} \times 100\% \tag{8}$$

$$\text{ACC} = 100\% - \text{NRMSE} \tag{9}$$

Here, n is the sample number, P_i is the generated power of the wind farm at the ith time point, $\widehat{P_i}$ is the wind power forecast of the wind farm at the ith time point, Cap is the total installed capacity of the wind farm, and ACC is the accuracy of day-ahead wind power forecast.

The relative reduction in NRMSE and the relative improvement in the accuracy, according to Equations (10) and (11), were applied to show the relative performance improvement by the incorporation of data enrichment.

$$\text{NRMSE}_{\Delta\%} = 100\% - \left(\text{NRMSE}_{\text{withDE}} / \text{NRMSE}_{\text{withoutDE}} \right) \tag{10}$$

$$\text{ACC}_{\Delta\%} = \text{ACC}_{\text{withDE}} / \text{ACC}_{\text{withoutDE}} - 100\% \tag{11}$$

Here, the subscript $\Delta\%$ denotes the relative change, the subscripts withoutDE and withDE represent the metrics achieved without and with data enrichment, and $\text{ACC}_{\Delta\%}$ is the relative improvement.

In order to make the performance improvement more tangible, the penalty resulting from the wind power forecast error that the grid operators in the North China Grid enforce was also calculated as an evaluation metric. The penalty and absolute penalty reduction achieved by the method were calculated according to Equations (12) [69] and (13):

$$\text{Penalty} = (85\% - \text{ACC}) \times 40\% \tag{12}$$

$$\text{Penalty}_{\triangle} = \text{Penalty}_{\text{withDE}} - \text{Penalty}_{\text{withoutDE}} \tag{13}$$

Here, Penalty is the penalized power production reduction compared to the installed capacity of the wind farm.

5. Results and Discussion

5.1. General Effectiveness of the Data Enrichment Method

The baseline was to input the weather prediction data from GFS without adding the data enrichment method proposed in Section 3 using XGBoost and LSTM for all three wind farms. The comparison was to employ the data enrichment method proposed in Section 3. First, the historical differences between the weather prediction and WTV of all the weather variables calculated according to Equation (1) were added as error features of weather prediction resources. Secondly, the projected wind speeds of eight adjacent nodes calculated according to Equation (2) were added as features of neighboring weather prediction. Thirdly, the lag 1 feature, difference 1 feature, and 3-h-rolling feature of all the weather variables calculated according to Equations (3)–(5) were added as time series features of the weather prediction sources. Lastly, three additional weather prediction sources, i.e., ECMWF, IBM, and CWC, were added to the inputs according to the criteria described in Section 3.6.

The results in Table 5 show that the forecast results using XGBoost and LSTM with data enrichment are better than those using XGBoost and LSTM without data enrichment. For the XGBoost model, the data enrichment method enabled a relative accuracy improvement from 5.0% to 15.9% and a relative error reduction from 11.2% to 27.5% for the three wind farms. For the LSTM model, the data enrichment method enabled a relative accuracy improvement from 16.7% to 17.4% and a relative error reduction from 27.2% to 36.9%. It is

indicated that the proposed data enrichment method can effectively enhance the accuracy of wind power forecasting for different wind farms and forecasting models.

Table 5. Comparison of the forecast accuracy, accuracy improvement, and RMSE reduction in two wind energy forecasting models with and without the proposed data enrichment method for three wind farms.

Model Wind Farm	XGBoost				LSTM			
	$ACC_{withoutDE}$ [1]	ACC_{withDE}	$ACC_{\triangle\%}$	$NRMSE_{\triangle\%}$	$ACC_{withoutDE}$	ACC_{withDE}	$ACC_{\triangle\%}$	$NRMSE_{\triangle\%}$
№ 1	69.1%	72.6%	5.0%	11.2%	62.0%	72.3%	16.7%	27.2%
№ 2	72.1%	77.2%	7.2%	18.5%	63.0%	76.7%	21.6%	36.9%
№ 3	63.4%	73.4%	15.9%	27.5%	61.0%	71.6%	17.55%	27.3%

[1] DE stands for data enrichment.

It should be pointed out that this study aimed to validate the relative improvement in accuracy made possible by applying the intended data enrichment method along with each wind power forecasting model, rather than to compare the absolute values of the accuracy achieved by wind power forecasting models. However, a comparison with the model improvements in the literature and the savings created from avoided penalties may become two approaches that can shed some light on the significance of the improvement brought about by the proposed method:

- The review in [68] published in 2021 listed the percentage error reduction in 41 hybrid wind power forecasting models compared to their benchmarks. There were 26 models designed for short-term wind power or wind speed forecasting, from which eight were evaluated by RMSE. The average RMSE reduction in the models proposed in the eight studies was 24.0%, which is comparable to the error reduction achieved by the proposed data enrichment method in this paper;

- The improvement brought by other published methods that also have XGBoost and LSTM as benchmarks is also comparable to the improvement brought by the proposed data enrichment method. Xiong et al. [73] proposed an improved XGBoost algorithm via Bayesian hyperparameter optimization (BH-XGBoost) and verified the efficacy of the improvement relative to XGBoost on a 200 MW wind farm. The verification results showed that the BH-XGBoost achieved 10.2% to 21.4% of RMSE reduction. Qin et al. proposed an improved LSTM algorithm that combines variational mode decomposition (VMD), maximum relevance and minimum redundancy algorithm (mRMR), long short-term memory neural network (LSTM), and firefly algorithm (FA) together. Compared to LSTM, the combined method achieved an RMSE reduction of 27.9%;

- Table 6 shows that the proposed data enrichment method can help wind farm operator to avoid 1.4% to 5.5% of the penalized power. Assuming that a wind farm had an installed capacity of 180 MW, 2000 annual full load hours, and 0.4 RMB/kWh of the feed-in tariff for a wind farm, the annual savings could amount to CNY 2 to 7 million.

Table 6. Comparison of the penalized power resulting from wind power forecasting errors of two wind energy forecasting models with and without the proposed data enrichment method for three wind farms.

Model Wind Farm	XGBoost			LSTM		
	$Penalty_{withoutDE}$	$Penalty_{withDE}$	$Penalty_{\triangle}$	$Penalty_{withoutDE}$	$Penalty_{withDE}$	$Penalty_{\triangle}$
№ 1	6.4%	5.0%	1.4%	9.2%	5.1%	4.1%
№ 2	5.2%	3.1%	2.1%	8.8%	3.3%	5.5%
№ 3	8.8%	4.6%	4.0%	9.6%	5.4%	4.3%

It can also be seen from Figure 4 that the accuracy improvement brought about by the proposed data enrichment method varied from model to model and wind farm to wind

farm. A definite accuracy improvement rate for other forecasting models and wind farms is not assured. Moreover, three interesting trends were observed:

- XGBoost outperformed LSTM for all three wind farms, regardless of the application of data enrichment or not;
- For all of the wind farms, the accuracy improvement brought by data enrichment with LSTM through the proposed method was more than that with XGBoost;
- The accuracy of the two models was closer with the aid of the proposed data enrichment method, as can be observed in Figure 4. For wind farms № 1 and № 2 the forecast accuracy of LSTM caught up and became almost the same as that of XGBoost after the introduction of the data enrichment step.
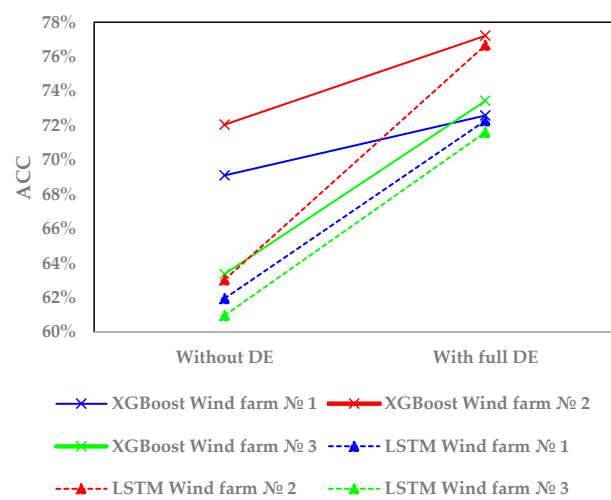

Figure 4. Accuracy [1] improvement in the proposed method on two models for three wind farms. ([1] ACC was rounded to the first decimal place, which are with three significant digits. Compared to the original data with five significant digits and intermediate data saved to the full use of double-precision format, the calculation error of ACC was negligible).

Possible reasons for the three trends are as follows:

- The data enrichment enables the original information in the input data to be better learned by LSTM; XGBoost has already learned this prior to data enrichment;
- LSTM can better learn the additional information brought by the data enrichment than XGBoost.

5.2. Effectiveness of Each Step of the Data Enrichment Method

The four steps of the data enrichment method were applied incrementally to both the XGBoost and LSTM models for wind farm № 1 to validate the efficacy of each step. Consequently, the forecast accuracies with different numbers of data enrichment steps applied were compared.

The baseline was to use the NWP data of GFS as the input and without adding any steps in the data enrichment method on both models for wind farm № 1. The four steps of the data enrichment were executed in the same way as in Section 5.1. The results in Table 7 and Figure 5 show clear evidence that all four steps of the data enrichment method contributed to the accuracy improvement in both wind power forecasting models. Figure 5 also illustrates an interesting phenomenon that the enhancement in accuracy differed from model to model and from step to step:

- While the addition of error features of weather prediction sources did little to improve the forecast accuracy of XGBoost, it resulted in a significant increase in the forecast accuracy of LSTM. On the contrary, the addition of time series features of weather prediction sources led to a considerable increase in the forecast accuracy of XGBoost, but it hardly improved the accuracy of LSTM. The contrast might well reflect the

merits and demerits of the two models: XGBoost is good at analyzing the relationship between different variables simultaneously, while the advantage of LSTM lies in learning the relationship between the historical values and future values of time series. That being the case, XGBoost might already infer the error features of different weather prediction sources from the historical weather prediction and wind power production, but LSTM does not. The step of adding error features explicitly expresses the error characteristics of weather prediction for LSTM to learn. Similarly, it might be easy for LSTM to learn the temporal continuity of wind from time series prior to data enrichment, making the step of adding time series features of NWP almost redundant to LSTM. However, the addition of the time series features of NWP is a helpful complement to XGBoost to allow it to learn the time series characteristics. Similar phenomena have also appeared and been discussed in studies on forecasting models in other areas [2,74,75];

- Adding features of neighboring weather prediction nodes only slightly improved the accuracy for both models;
- Adding complementary weather prediction sources could be instrumental in improving the forecast accuracy of both models since the step supplied more weather prediction information to the model.

Table 7. Forecast accuracy of the wind power forecasting models with the addition of each step in the data enrichment method for wind farm № 1.

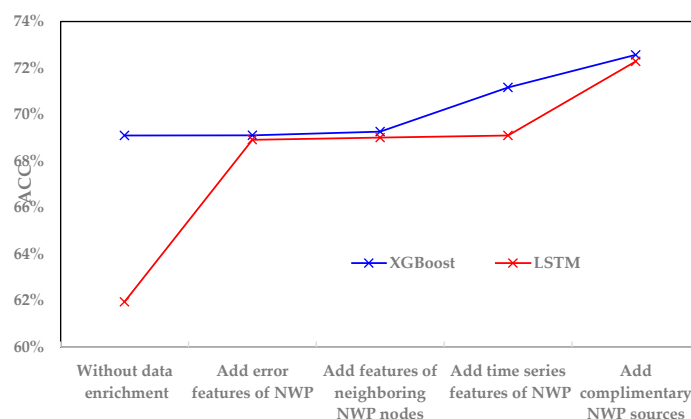| Steps	Models	XGBoost for Wind Farm № 1	LSTM for Wind Farm № 1
	Without data enrichment (baseline)	69.1%	62.0%
With data enrichment	Add error features of weather prediction sources	69.1%	68.9%
	Add features of neighboring nodes	69.3%	69.0%
	Add time series features of weather prediction sources	71.2%	69.1%
	Add complimentary weather prediction sources	72.6%	72.3%

Figure 5. Accuracy improvement in each step in the proposed method applied to both XGBoost and LSTM for wind farm № 1.

6. Limitations

The experiment with two different types of wind power forecasting models and three different actual wind farms can only provide preliminary results for the added value brought by the proposed data enrichment method. More models and wind farms are required for subsequent testing to further validate the method's effectiveness and to explore the degree of accuracy improvement that the method can achieve. Based on more

data, further analysis can be conducted to discover the relationship between the degree of performance brought by the proposed data enrichment method and its influencing factors.

Although the findings in Section 5.2 demonstrate the effectiveness of each proposed step in the data enrichment method, more forecasting models and wind farms should be tested to provide further verification in the future. It is also necessary to note that the presented equations for each step are not compulsory in realizing the data enrichment method, but they are rather proposed as examples. Other calculation equations reflecting the concepts of each data enrichment step can be tried in the future in search of the best calculation method for each data enrichment step. Researchers could omit some steps or design new steps depending on their problems and models, as long as each step enriches the data according to the concept proposed in this paper.

7. Conclusions and Recommendations

Wind power is one of the fast-growing renewable energy sources in the world. Accurate short-term wind power forecasting is of great significance to power dispatching and grid security, as well as the profitability of wind farms in terms of power trading and imbalance penalty. To pursue better forecast performance, extensive research has been carried out on different wind power forecasting models. However, there is a paucity of research on data preprocessing methods that aim to add more valuable information to forecasting models.

Based on substantial engineering experience and inspiration from other areas, a data enrichment method was proposed to enhance the performance of wind power forecasting. The method effectively involved multiple complementary commercial weather prediction sources and extracted valuable wind intrinsic physical features as additional inputs to the wind power forecasting models. Experimental results showed that the addition of the proposed data enrichment method could effectively reduce the NRMSE produced by XGBoost and LSTM by 11% to 25% for three different actual wind farms. Moreover, all four steps in the proposed data enrichment method were verified to contribute to the improvement in forecast accuracy.

This paper is, however, only a preliminary study that demonstrates the idea of improving the accuracy of wind power forecasting by data enrichment. Further research can be carried out on the following aspects:

- Application of the data enrichment method to more types of wind power forecasting models to further identify the adaptability and performance of the method;
- Extension of the data enrichment method to long-term and very-short-term wind power forecasting models or other forecasting problems related to weather data;
- Exploration of the relationship between the intrinsic strength and weakness of forecasting models and the data enrichment method;
- In-depth study into the optimized calculation of each data enrichment step;
- Design of other possible methods of data enrichment.

Author Contributions: Conceptualization, Y.Z.; methodology, Y.Z.; software, Y.Z.; validation, Y.Z. and C.Y.; formal analysis, Y.Z.; investigation, Y.Z.; resources, Y.Z. and C.Y.; data curation, C.Y.; writing—original draft preparation, Y.Z.; writing—review and editing, L.M.; visualization, Y.Z.; supervision: L.M. and W.N.; project administration, Y.Z. All authors have read and agreed to the published version of the manuscript.

References

1. Alkhayat, G.; Mehmood, R. A review and taxonomy of wind and solar energy forecasting methods based on deep learning. *Energy AI* **2021**, *4*, 100060. [CrossRef]
2. Hanifi, S.; Liu, X.; Lin, Z.; Lotfian, S. A critical review of wind power forecasting methods—Past, present and future. *Energies* **2020**, *13*, 3764. [CrossRef]
3. Tascikaraoglu, A.; Uzunoglu, M. A review of combined approaches for prediction of short-term wind speed and power. *Renew. Sustain. Energy Rev.* **2014**, *34*, 243–254. [CrossRef]
4. Soman, S.S.; Zareipour, H.; Malik, O.; Mandal, P. In A review of wind power and wind speed forecasting methods with different time horizons. In Proceedings of the North-American Power Symposium (NAPS) 2010, Arlington, TX, USA, 26–28 September 2010; pp. 1–8.
5. Zhao, E.; Sun, S.; Wang, S. New developments in wind energy forecasting with artificial intelligence and big data: A scientometric insight. *Data Sci. Manag.* **2022**, *5*, 84–95. [CrossRef]
6. Wang, Y.; Zou, R.; Liu, F.; Zhang, L.; Liu, Q. A review of wind speed and wind power forecasting with deep neural networks. *Appl. Energy* **2021**, *304*, 117766. [CrossRef]
7. Liu, H.; Chen, C. Data processing strategies in wind energy forecasting models and applications: A comprehensive review. *Appl. Energy* **2019**, *249*, 392–408. [CrossRef]
8. Lipu, M.S.H.; Miah, M.S.; Hannan, M.A.; Hussain, A.; Sarker, M.R.; Ayob, A.; Saad, M.H.M.; Mahmud, M.S. Artificial intelligence based hybrid forecasting approaches for wind power generation: Progress, challenges and prospects. *IEEE Access* **2021**, *9*, 102460–102489. [CrossRef]
9. Reichstein, M.; Camps-Valls, G.; Stevens, B.; Jung, M.; Denzler, J.; Carvalhais, N.; Prabhat. Deep learning and process understanding for data-driven Earth system science. *Nature* **2019**, *566*, 195–204. [CrossRef] [PubMed]
10. Knapp, E.D.; Langill, J.T. Exception, anomaly, and threat detection. In *Industrial Network Security—Securing Critical Infrastructure Networks for Smart Grid, SCADA, and Other Industrial Control Systems*; Elsevier: Amsterdam, The Netherlands, 2015.
11. Buckinx, W.; Verstraeten, G.; Van den Poel, D. Predicting customer loyalty using the internal transactional database. *Expert Syst. Appl.* **2007**, *32*, 125–134. [CrossRef]
12. Azad, S.A.; Wasimi, S.; Ali, A.B.M.S. Business Data Enrichment: Issues and Challenges. In Proceedings of the 2018 5th Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE), Nadi, Fiji, 10–12 December 2018; pp. 98–102.
13. Reis Filho, I.J.; Marcacini, R.M.; Rezende, S.O. On the enrichment of time series with textual data for forecasting agricultural commodity prices. *MethodsX* **2022**, *9*, 101758. [CrossRef]
14. Pombo, D.V.; Gocmen, T.; Das, K.; Sorensen, P. Multi-horizon data-driven wind power forecast: From nowcast to 2 days-ahead. In Proceedings of the 2021 International Conference on Smart Energy Systems and Technologies (SEST), Vaasa, Finland, 6–8 September 2021; pp. 1–6.
15. He, Y.; Cao, C.; Wang, S.; Fu, H. Nonparametric probabilistic load forecasting based on quantile combination in electrical power systems. *Appl. Energy* **2022**, *322*, 119507. [CrossRef]
16. Chen, X.; Zhao, J.; Jia, X.; Li, Z. Multi-step wind speed forecast based on sample clustering and an optimized hybrid system. *Renew. Energy* **2021**, *165*, 595–611. [CrossRef]
17. Takahashi, Y.; Fujimoto, Y.; Hayashi, Y. Forecast of infrequent wind power ramps based on data sampling strategy. *Energy Procedia* **2017**, *135*, 496–503. [CrossRef]
18. Wang, J.; Li, Q.; Zhang, H.; Wang, Y. A deep-learning wind speed interval forecasting architecture based on modified scaling approach with feature ranking and two-output gated recurrent unit. *Expert Syst. Appl.* **2023**, *211*, 118419. [CrossRef]
19. Liu, Z.; Hara, R.; Kita, H. Hybrid forecasting system based on data area division and deep learning neural network for short-term wind speed forecasting. *Energy Convers. Manag.* **2021**, *238*, 114136. [CrossRef]
20. Qiao, B.; Liu, J.; Wu, P.; Teng, Y. Wind power forecasting based on variational mode decomposition and high-order fuzzy cognitive maps. *Appl. Soft Comput.* **2022**, *129*, 109586. [CrossRef]
21. Manero, J.; Béjar, J.; Cortés, U. "Dust in the wind...", deep learning application to wind energy time series forecasting. *Energies* **2019**, *12*, 2385. [CrossRef]
22. Huang, Y.; Liu, G.; Hu, W. Priori-guided and data-driven hybrid model for wind power forecasting. *ISA Trans.* **2022**, *in press*. [CrossRef] [PubMed]
23. Zhang, Y.; Li, Y.; Zhang, G. Short-term wind power forecasting approach based on Seq2Seq model using NWP data. *Energy* **2020**, *213*, 118371. [CrossRef]
24. Qian, Z.; Pei, Y.; Zareipour, H.; Chen, N. A review and discussion of decomposition-based hybrid models for wind energy forecasting applications. *Appl. Energy* **2019**, *235*, 939–953. [CrossRef]
25. Chen, Y.; Yu, S.; Islam, S.; Lim, C.P.; Muyeen, S.M. Decomposition-based wind power forecasting models and their boundary issue: An in-depth review and comprehensive discussion on potential solutions. *Energy Rep.* **2022**, *8*, 8805–8820. [CrossRef]
26. Dong, W.; Sun, H.; Tan, J.; Li, Z.; Zhang, J.; Zhao, Y.Y. Short-term regional wind power forecasting for small datasets with input data correction, hybrid neural network, and error analysis. *Energy Rep.* **2021**, *7*, 7675–7692. [CrossRef]
27. Chen, H.; Birkelund, Y.; Zhang, Q. Data-augmented sequential deep learning for wind power forecasting. *Energy Convers. Manag.* **2021**, *248*, 114790. [CrossRef]

28. Niu, D.; Sun, L.; Yu, M.; Wang, K. Point and interval forecasting of ultra-short-term wind power based on a data-driven method and hybrid deep learning model. *Energy* **2022**, *254*, 124384. [CrossRef]

29. Yan, J.; Zhang, H.; Liu, Y.; Han, S.; Li, L.; Lu, Z. Forecasting the high penetration of wind power on multiple scales using multi-to-multi mapping. *IEEE Trans. Power Syst.* **2018**, *33*, 3276–3284. [CrossRef]

30. Delle Monache, L.; Nipen, T.; Liu, Y.; Roux, G.; Stull, R. Kalman filter and analog schemes to postprocess numerical weather predictions. *Mon. Weather Rev.* **2011**, *139*, 3554–3570. [CrossRef]

31. Cassola, F.; Burlando, M. Wind speed and wind energy forecast through Kalman filtering of Numerical Weather Prediction model output. *Appl. Energy* **2012**, *99*, 154–166. [CrossRef]

32. Hur, S. Short-term wind speed prediction using Extended Kalman filter and machine learning. *Energy Rep.* **2021**, *7*, 1046–1054. [CrossRef]

33. Louka, P.; Galanis, G.; Siebert, N.; Kariniotakis, G.; Katsafados, P.; Pytharoulis, I.; Kallos, G. Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering. *J. Wind. Eng. Ind. Aerodyn.* **2008**, *96*, 2348–2362. [CrossRef]

34. Zhang, C.; Zhou, J.; Li, C.; Fu, W.; Peng, T. A compound structure of ELM based on feature selection and parameter optimization using hybrid backtracking search algorithm for wind speed forecasting. *Energy Convers. Manag.* **2017**, *143*, 360–376. [CrossRef]

35. Li, Y.; Peng, T.; Zhang, C.; Sun, W.; Hua, L.; Ji, C.; Muhammad Shahzad, N. Multi-step ahead wind speed forecasting approach coupling maximal overlap discrete wavelet transform, improved grey wolf optimization algorithm and long short-term memory. *Renew. Energy* **2022**, *196*, 1115–1126. [CrossRef]

36. Zha, W.; Liu, J.; Li, Y.; Liang, Y. Ultra-short-term power forecast method for the wind farm based on feature selection and temporal convolution network. *ISA Trans.* **2022**, *129*, 405–414. [CrossRef] [PubMed]

37. Lu, P.; Ye, L.; Zhao, Y.; Dai, B.; Pei, M.; Li, Z. Feature extraction of meteorological factors for wind power prediction based on variable weight combined method. *Renew. Energy* **2021**, *179*, 1925–1939. [CrossRef]

38. Lu, P.; Ye, L.; Pei, M.; Zhao, Y.; Dai, B.; Li, Z. Short-term wind power forecasting based on meteorological feature extraction and optimization strategy. *Renew. Energy* **2022**, *184*, 642–661. [CrossRef]

39. Ai, X.; Li, S.; Xu, H. Short-term wind speed forecasting based on two-stage preprocessing method, sparrow search algorithm and long short-term memory neural network. *Energy Rep.* **2022**, *8*, 14997–15010. [CrossRef]

40. Wang, Y.; Wang, J.; Li, Z.; Yang, H.; Li, H. Design of a combined system based on two-stage data preprocessing and multi-objective optimization for wind speed prediction. *Energy* **2021**, *231*, 121125. [CrossRef]

41. Yang, Z.; Wang, J. A combination forecasting approach applied in multistep wind speed forecasting based on a data processing strategy and an optimized artificial intelligence algorithm. *Appl. Energy* **2018**, *230*, 1108–1125. [CrossRef]

42. Niu, X.; Wang, J. A combined model based on data preprocessing strategy and multi-objective optimization algorithm for short-term wind speed forecasting. *Appl. Energy* **2019**, *241*, 519–539. [CrossRef]

43. Deng, Y.; Wang, B.; Lu, Z. A hybrid model based on data preprocessing strategy and error correction system for wind speed forecasting. *Energy Convers. Manag.* **2020**, *212*, 112779. [CrossRef]

44. Tian, C.; Hao, Y.; Hu, J. A novel wind speed forecasting system based on hybrid data preprocessing and multi-objective optimization. *Appl. Energy* **2018**, *231*, 301–319. [CrossRef]

45. Li, C.; Zhu, Z.; Yang, H.; Li, R. An innovative hybrid system for wind speed forecasting based on fuzzy preprocessing scheme and multi-objective optimization. *Energy* **2019**, *174*, 1219–1237. [CrossRef]

46. Needham, D. *The Enrichment Game: A Story about Making Data Powerful*, 1st ed.; Technics Publications: Basking Ridge, NJ, USA, 2021.

47. Allen, M.; Cervo, D. Chapter 9—Data Quality Management. In *Multi-Domain Master Data Management*; Allen, M., Cervo, D., Eds.; Morgan Kaufmann: Boston, MA, USA, 2015; pp. 131–160.

48. Schön, C.; Dittrich, J.; Müller, R. The Error is the Feature: How to Forecast Lightning using a Model Prediction Error. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Anchorage, AK, USA, 4–8 August 2019.

49. Miao, L.; Yu, D.; Pang, Y.; Zhai, Y. Temperature Prediction of Chinese Cities Based on GCN-BiLSTM. *Appl. Sci.* **2022**, *12*, 11833. [CrossRef]

50. Xie, H.; Zheng, R.; Lin, Q. Short-Term Intensive Rainfall Forecasting Model Based on a Hierarchical Dynamic Graph Network. *Atmosphere* **2022**, *13*, 703. [CrossRef]

51. Ozdemir, S. *Feature Engineering Bookcamp*; Manning Publications: Shelter Island, NY, USA, 2022.

52. Hamilton, J.D. *Time Series Analysis*; Princeton University Press: Princeton, NJ, USA, 1994.

53. Dabernig, M. Comparison of different numerical weather prediction models as input for statistical wind power forecasts. Master's Thesis, University of Innsbruck, Innsbruck, Austria, 2013.

54. Chen, T.; Guestrin, C. XGBoost: A scalable tree noosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016.

55. Cai, R.; Xie, S.; Wang, B.; Yang, R.; Xu, D.; He, Y. Wind speed forecasting based on extreme gradient boosting. *IEEE Access* **2020**, *8*, 175063–175069. [CrossRef]

56. Browell, J.; Gilbert, C.; McMillan, D. Use of turbine-level data for improved wind power forecasting. In Proceedings of the 2017 IEEE Manchester PowerTech, Manchester, UK, 18–22 June 2017; pp. 1–6.

57. Gebin, L.G.G.; Salgado, R.M.; Nogueira, D.A. Wind power forecast: Ensemble model based in statistical and machine learning models. *Res. Soc. Dev.* **2020**, *9*, e38291211251. [CrossRef]

58. Bochenek, B.; Jurasz, J.; Jaczewski, A.; Stachura, G.; Sekuła, P.; Strzyżewski, T.; Wdowikowski, M.; Figurski, M. Day-ahead wind power forecasting in poland based on numerical weather prediction. *Energies* **2021**, *14*, 2164. [CrossRef]

59. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef]

60. Shokouhifar, M.; Ranjbarimesan, M. Multivariate time-series blood donation/demand forecasting for resilient supply chain management during COVID-19 pandemic. *Clean. Logist. Supply Chain.* **2022**, *5*, 100078. [CrossRef]

61. AlRassas, A.M.; Al-qaness, M.A.A.; Ewees, A.A.; Ren, S.; Sun, R.; Pan, L.; Abd Elaziz, M. Advance artificial time series forecasting model for oil production using neuro fuzzy-based slime mould algorithm. *J. Pet. Explor. Prod. Technol.* **2022**, *12*, 383–395. [CrossRef] [PubMed]

62. Kader, A.; Izzati, N. A review of long short-term memory approach for time series analysis and forecasting. In Proceedings of the 2nd International Conference on Emerging Technologies and Intelligent System: ICETIS 2022, Virtual, 2–3 September 2022; Springer Nature: Cham, Switzerland, 2023.

63. Puspita Sari, A.; Suzuki, H.; Kitajima, T.; Yasuno, T.; Arman Prasetya, D.; Rabi', A. Deep convolutional long short-term memory for forecasting wind speed and direction. *SICE J. Control. Meas. Syst. Integr.* **2021**, *14*, 30–38. [CrossRef]

64. Moharm, K.; Eltahan, M.; Elsaadany, E. Wind speed forecast using LSTM and BILSTM algorithms over Gabal El-Zayt wind farm. In Proceedings of the 2020 International Conference on Smart Grids and Energy Systems (SGES), Perth, Australia, 23–26 November 2020; pp. 922–927.

65. Su, Y.; Yu, J.; Tan, M.; Wu, Z.; Xiao, Z.; Hu, J. A LSTM based wind power forecasting method considering wind frequency components and the wind turbine states. In Proceedings of the 2019 22nd International Conference on Electrical Machines and Systems (ICEMS), Harbin, China, 11–14 August 2019; pp. 1–6.

66. Li, G.; Shi, J. On comparing three artificial neural networks for wind speed forecasting. *Appl. Energy* **2010**, *87*, 2313–2320. [CrossRef]

67. González-Sopeña, J.M.; Pakrashi, V.; Ghosh, B. An overview of performance evaluation metrics for short-term statistical wind power forecasting. *Renew. Sustain. Energy Rev.* **2021**, *138*, 110515. [CrossRef]

68. Yang, B.; Zhong, L.; Wang, J.; Shu, H.; Zhang, X.; Yu, T.; Sun, L. State-of-the-art one-stop handbook on wind forecasting technologies: An overview of classifications, methodologies, and analysis. *J. Clean. Prod.* **2021**, *283*, 124628. [CrossRef]

69. China Energy Regulatory Bureau. *Implementation Rules for Grid-Connected Operation Management of Wind Farms in North China*; No. 381; China Energy Regulatory Bureau: Beijing, China, 2015. (In Chinese)

70. Entsoe Enhanced RES Infeed Forecasting-Wind. Available online: https://www.entsoe.eu/Technopedia/techsheets/enhanced-res-infeed-forecasting-wind (accessed on 3 June 2022).

71. European Wind Energy Association. *Powering Europe: Wind Energy and the Electricity Grid*; European Wind Energy Association: Brussels, Belgium, 2010.

72. Shcherbakov, M.V.; Brebels, A.; Shcherbakova, N.L.; Tyukov, A.P.; Janovsky, T.A.; Kamaev, V.A.E. A Survey of Forecast Error Measures. *World Appl. Sci. J.* **2013**, *24*, 171–176.

73. Xiong, X.; Guo, X.; Zeng, P.; Zou, R.; Wang, X. A Short-Term Wind Power Forecast Method via XGBoost Hyper-Parameters Optimization. *Front. Energy Res.* **2022**, *10*, 574. [CrossRef]

74. Liu, Y.; Zhang, S.; Zhang, J.; Tang, L.; Bai, Y. Assessment and comparison of six machine learning models in estimating evapotranspiration over croplands using remote sensing and meteorological factors. *Remote Sens.* **2021**, *13*, 3838. [CrossRef]

75. Du, Q.; Yin, F.; Li, Z. Base station traffic prediction using XGBoost-LSTM with feature enhancement. *IET Netw.* **2020**, *9*, 29–37. [CrossRef]