energies

Review

Review of AI-Based Wind Prediction within Recent Three Years: 2021–2023

Dongran Song [1], Xiao Tan [1], Qian Huang [1], Li Wang [1], Mi Dong [1,*], Jian Yang [1] and Solomin Evgeny [2]

[1] School of Automation, Central South University, Changsha 410083, China; humble_szy@163.com (D.S.); 224611003@csu.edu.cn (X.T.); tomhuangq@163.com (Q.H.); 224601008@csu.edu.cn (L.W.); jian.yang@csu.edu.cn (J.Y.)
[2] Department of Electric Stations, Grids and Power, Supply Systems, South Ural State University, 76 Prospekt Lenina, 454080 Chelyabinsk, Russia; solominev@susu.ru
* Correspondence: mi.dong@csu.edu.cn

Abstract: Wind prediction has consistently been in the spotlight as a crucial element in achieving efficient wind power generation and reducing operational costs. In recent years, with the rapid advancement of artificial intelligence (AI) technology, its application in the field of wind prediction has made significant strides. Focusing on the process of AI-based wind prediction modeling, this paper provides a comprehensive summary and discussion of key techniques and models in data preprocessing, feature extraction, relationship learning, and parameter optimization. Building upon this, three major challenges are identified in AI-based wind prediction: the uncertainty of wind data, the incompleteness of feature extraction, and the complexity of relationship learning. In response to these challenges, targeted suggestions are proposed for future research directions, aiming to promote the effective application of AI technology in the field of wind prediction and address the crucial issues therein.

Keywords: wind prediction; artificial intelligence; data preprocessing; feature extraction; parameter optimization

Citation: Song, D.; Tan, X.; Huang, Q.; Wang, L.; Dong, M.; Yang, J.; Evgeny, S. Review of AI-Based Wind Prediction within Recent Three Years: 2021–2023. *Energies* **2024**, *17*, 1270. https://doi.org/10.3390/en17061270

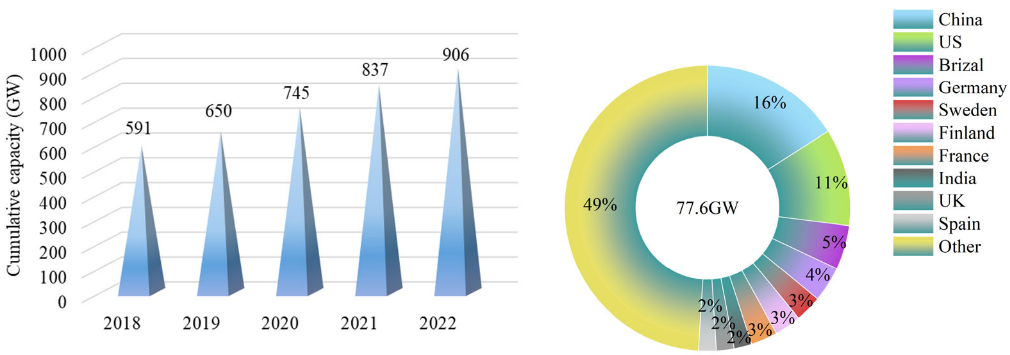
Academic Editors: Galih Bangga, Martin Otto Laver Hansen and Alessandro Bianchini

Received: 26 January 2024
Revised: 1 March 2024
Accepted: 5 March 2024
Published: 7 March 2024

1. Introduction

With the development of society and economy, the issues of energy shortage and environmental pollution become increasingly prominent, making it imperative to promote the transformation from traditional fossil energy to renewable energy [1]. Wind energy, as a widely distributed and pollution-free renewable energy, has been widely used around the world [2,3]. In the past 30 years, global wind energy technology and industry have experienced rapid development, achieving significant progress in both theoretical and applied research [4,5]. The future outlook indicates a promising trend for further advancement. According to the Global Wind Report 2023 published by the Global Wind Energy Council [6], 77.6 GW of new wind power capacity was connected to power grids in 2022, bringing total installed wind capacity to 906 GW, a growth of 9% compared with 2021 (as shown in Figure 1a). The world's top ten markets for new installations in 2022 are depicted in Figure 1b, from which China contributes 16% of its 2022 additions, followed by the United States with 11%. Moreover, the report anticipates that by 2024, the global installed capacity of onshore wind power will exceed 100 GW for the first time, and the newly added capacity for offshore wind power worldwide will also reach a new record high by 2025.

Due to the high randomness and volatility of wind, accurate wind prediction is of great significance for improving the efficiency and quality of wind power generation, reducing fatigue loads, and extending the service life of wind turbines [7,8]. AI-based wind prediction can assist in achieving maximum wind energy capture, thereby improving power generation efficiency. Accurate wind forecasting helps avoid unnecessary starts and stops of wind turbines, subsequently enhancing turbine lifespan and reducing operational costs. Wind prediction mainly includes wind speed prediction (WSP) and wind power prediction (WPP), and the prediction methods are general, so a unified discussion is chosen. Many methods have been proposed for WSP/WPP, which can be classified into several groups according to different classification criteria [9], as shown in Figure 2. According to the time scale [10], they are divided into ultra-short-term (a few seconds to 30 min ahead), short-term (30 min to 6 h ahead), medium-term (6 h to 1 day ahead) and long-term (1 day to 1 week or more ahead) models [11–13]. This classification reflects a focus on the time span of predictions and provides more specific time frameworks for different application scenarios. Ultra-short-term prediction is primarily utilized for real-time control and load tracking. By providing accurate predictions within a short time frame, it enables timely adjustments to cope with momentary changes in wind speed. Short-term prediction, on the other hand, provides support for load scheduling planning in scenarios with a larger time span. This is crucial for efficiently organizing the operation and scheduling of the power system in the coming hours. When conducting short-term data-driven wind predictions at the wind farm level, it may be particularly important to consider the impact of other turbines, such as wake effects [14,15] or the spatiotemporal correlation among turbines [16]. Medium-term prediction extends over a longer time span and is mainly applied in energy trading and power system management. Wind prediction several hours to a day in advance facilitates better planning for the supply and demand balance in the electricity market. Long-term forecasting is employed to guide optimal maintenance plans to allow for the proper planning of maintenance and repair work, ensuring the reliability and stability of the system. According to the prediction objective, they are divided into wind turbines [17] and wind farms [18] WSP/WPP. The prediction for wind turbines primarily focuses on the output of individual wind turbines, while the prediction for wind farms is more intricate, requiring consideration of the collaborative operation of multiple wind turbines within the entire wind farm [19]. According to the prediction types, WSP/WPP models are divided into deterministic models and probabilistic models [20,21]. The former provides only certain prediction results, but its prediction performance is easily limited by the complexity of the environment. The latter characterizes uncertainty through prediction intervals, which can often provide more information to decision makers [22,23]. Liu et al. [24] coupled the light gradient boosting machine (LGB) model and the Gaussian process regression (GPR) model. The LGB model provided high-precision deterministic wind speed prediction results, while the GPR model provided reliable probabilistic prediction results. Zhang et al. [25] proposed a probabilistic prediction model of wind power based on quantile regression and evaluated the uncertainties with different confidence levels.



(**a**) (**b**)

Figure 1. (**a**) Wind power global cumulative capacity, 2018–2022; (**b**) Top 10 markets for installations in 2022.

According to modeling theory, they can be divided into three categories [26]: physical models, traditional statistical models, and artificial intelligence (AI)-based models. Classical physical models (such as numerical weather forecasting (NWP) [27] and weather research forecasting (WRF) [28], etc.) typically incorporate various meteorological factors (such as temperature, humidity, and air pressure) and terrain, combining atmospheric science and fluid mechanics, to predict wind speed or wind power through complex calculations. In contrast to physical models, traditional statistical models establish linear statistical relationships between input and predicted output by analyzing historical wind speed data, including autoregressive moving average (ARMA) [2], autoregressive integrated moving average (ARIMA) [29], etc. However, both physical models and traditional statistical models struggle to handle the complex nonlinear relationships in wind data.

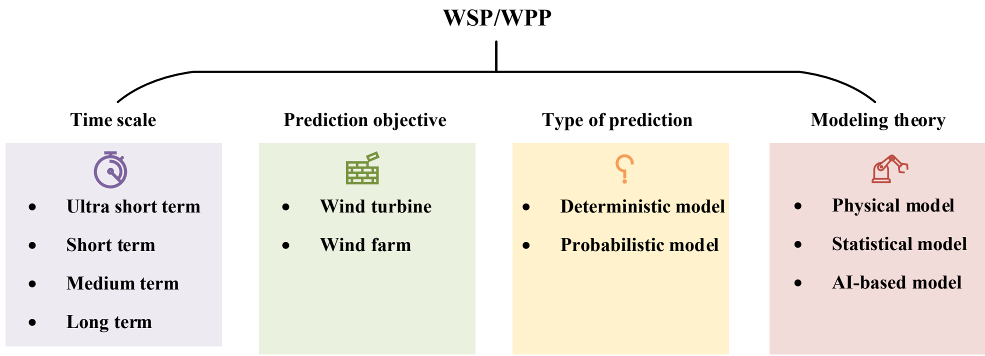


Figure 2. Classification for WSP/WPP.

With the rapid development of AI technology, there is increasing attention to its advantages in handling complex nonlinear problems in WSP/WPP [30]. AI-based data-driven [31] models include various machine learning (ML) algorithms [32,33], such as support vector machine (SVM), extreme learning machines (ELM), and various traditional artificial neural networks (ANNs) represented by back propagation neural network (BPNN). In many studies [34–36], AI-based models have shown better predictive accuracy and robustness in complex nonlinear problems [37]. Li et al. [38] successfully achieved accurate wind power prediction using the proposed enhanced crow search algorithm optimization-extreme learning machine model, keeping root mean square error (*RMSE*) and mean absolute percentage error (*MAPE*) values below 20% and 4%, respectively, effectively reducing the impact of large-scale wind power integration on the grid. Tan et al. [28] used a multi-input single-output ANN to correct the WRF model, and upon validation, it was demonstrated to enhance the WPP performance for a power plant in western Turkey. However, ML algorithms often require manual feature selection, which demands domain knowledge and expertise. In recent years, deep learning (DL) models such as recurrent neural networks (RNN) and convolutional neural networks (CNN) have gradually gained prominence and proven their effectiveness in the field of wind prediction [39]. Benti et al. [40] provided a detailed review of the progress and prospects of ML and DL technologies in the renewable energy generation field, discussing their advantages and limitations. Compared to ML models, DL models can automatically learn features from raw data, alleviating the burden of feature engineering. It is noteworthy that DL models excel in capturing and representing complex nonlinear relationships in large datasets but may suffer from overfitting when data are limited. Additionally, integrating multiple single models into hybrid models through serial (stacking) and parallel (weighted) methods can compensate for the limitations of individual models, improving overall performance and robustness across different types of data and prediction tasks. Qu et al. [41] employed a stacking integration algorithm to combine bagging, long short-term memory (LSTM), and random forest (RF) for high-precision prediction of target wind farm power. Huang et al. [42] generated preliminary predictions using five single models, namely k-nearest neighbors, RNN, LSTM, support vector regression, and random forest regression. By optimizing their weight distribution based on a population-based intelligent algorithm, this combination model surpassed individual models in prediction accuracy.

In summary, with the continuous development of AI technology, significant contributions have been made to the further development and utilization of wind energy, particularly in the context of WSP and WPP. Over the past three years, numerous AI-based WSP/WPP studies have been published, and it is necessary to systematically classify and analyze these articles. The contributions of this paper are summarized as follows:

- This paper systematically summarizes and organizes AI-based wind prediction research conducted from 2021 to 2023 to promote the effective application of AI technology in the field of wind prediction and address key challenges;
- This paper, focusing on critical steps, including data preprocessing, feature extraction, relationship learning, and parameter optimization, delves into in-depth analysis and discussion of the key technologies and models involved in these crucial processes;
- This paper highlights the challenges faced by AI technology in wind prediction, such as data uncertainty, model complexity, and hyperparameter selection. Specific recommendations for future research directions are provided to address these challenges.

The remaining sections are arranged as follows: Section 2 describes how to select and screen the literature in detail. Section 3 provides reviews of the essential methods and technologies in key steps of WSP/WPP, including data preprocessing, feature extraction, relationship learning, and parameter optimization, respectively. Section 4 presents the predictive performance evolution metrics. Section 5 discusses the challenges in AI-based wind prediction over the past three years and future trends. Lastly, Section 6 concludes this review.

2. Literature Selection and Screening

The dataset utilized in this study was extracted from journals indexed in the Web of Science (WOS). The WOS database encompasses over 20,000 authoritative and high-impact academic journals and conference proceedings. Its inclusion criteria are rigorous, considering various factors such as expert evaluations and measurement metrics. Users can search for data in WOS by specifying topics and publication periods, followed by reliable data collection through keyword or term co-occurrence analysis.

To obtain articles relevant to wind prediction, we initially conducted searches in the database using the primary keyword “wind prediction”, covering articles related to both WSP and WPP. The search period was set from 2021 to 2023, resulting in a total of 13,349 articles. Recognizing that AI is a broad concept encompassing machine learning and deep learning methods, using “AI” as a keyword for searching was deemed less prudent. Instead, we separately used the keywords “data preprocessing”, “feature extraction”, and “parameter optimization”, corresponding to the AI-based wind prediction modeling process, resulting in 3026, 258, and 630 articles, respectively. These articles were comprehensively summarized and classified. As “relationship learning” is integral to the prediction process, it did not need a separate keyword search. During the search for the aforementioned three keywords, the same article might appear in different search results. However, through keyword and abstract analysis, we could better understand the focal points of the articles, providing valuable insights for further quality screening.

Subsequently, Zotero 6.0.30 was utilized to deduplicate and rank articles obtained from searches with different keywords. This software not only has deduplication functionality but also extracts keywords and abstracts. Given potential duplications in articles applying similar methods, analyzing keywords and abstracts helped identify research emphases and provided better references for subsequent high-quality article selection.

3. Methods and Technologies

In the modeling process of AI-based wind prediction, data preprocessing, feature extraction, relationship learning, and parameter optimization are crucial steps. Firstly, data preprocessing is considered the cornerstone of the entire modeling process. The quality of input data are ensured through outlier detection and other operations, and then the data structure is analyzed by decomposition methods. Secondly, classical methods such as correlation analysis or approaches based on neural networks (NNs) are employed to extract effective features from the input data. Relationship learning is a central element in the modeling process, utilizing either a single prediction model or a combination of prediction models to learn the mapping relationship between input features and target outputs. Finally, parameter optimization aims to further enhance the performance of the prediction model by adjusting model parameters using various optimization algorithms. The framework of this research review is primarily illustrated in Figure 3.

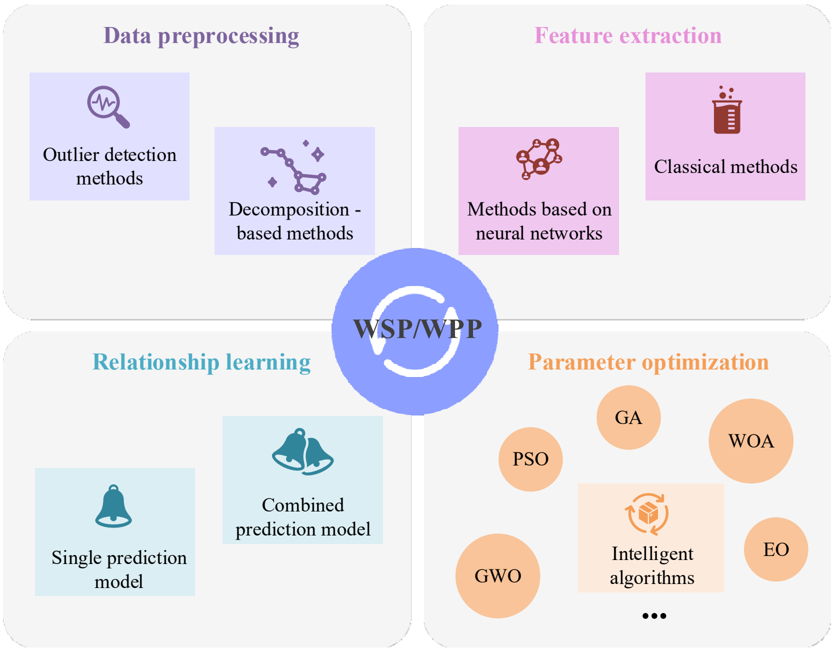


Figure 3. Framework of this research review.

3.1. Data Preprocessing

Due to the complexity of the environment and stochastic factors such as equipment failures, the quality of collected wind data may be significantly influenced, leading to issues such as missing values, noise, and outliers. To mitigate the uncertainty in the data and accurately extract features for training precise prediction models, preprocessing the collected wind data before prediction is crucial. Data preprocessing methods can be broadly categorized into two types: outlier detection methods and decomposition-based methods.

3.1.1. Outlier Detection Methods

Training prediction models with wind data that includes outliers can lead to incorrect information, thereby reducing the effectiveness of prediction. Therefore, employing outlier detection methods to categorize the raw data into outliers and normal values and discarding or correcting the outliers is crucial to mitigate their impact.

Traditional outlier detection methods are primarily based on statistical, clustering, distance, and density knowledge. Local outlier detection can utilize sliding windows or local density to detect outliers in data points, and another commonly used method involves using a weighted least squares algorithm to identify outliers. Khazaei et al. [43] employed a method combining automatic clustering with T2 statistics for outlier detection and removal, achieving accurate WPP. Fahim et al. [44] eliminated outliers based on the histogram-based outlier score and utilized a genetic algorithm for feature extraction, providing high-quality training data for DL models. He et al. [45] used the quartile method for outlier detection in initial data processing and replaced outliers with a cubic spline interpolation function before predicting wind power. Fu et al. [46] employed boxplot and medcouple for outlier detection and correction to improve the quality and smoothness of wind speed sequences. Ammar et al. [47] detected outliers using interquartile range and then corrected them using the last observation carried forward method. Özen et al. [48] coupled k-means with the isolation forest method to detect outliers in SCADA data and used the output of the WRF model to impute the missing data.

Additionally, in the literature review, models based on DL, primarily the denoising autoencoder, have been found to be effective in reducing noise in raw wind data while simultaneously extracting robust features. Unfortunately, no further AI-based methods specifically designed for identifying and handling outliers were discovered.

3.1.2. Decomposition-Based Methods

Unlike outlier detection methods that focus on processing abnormal data, decomposition-based methods treat all raw data equally. These methods primarily draw from knowledge in the field of signal processing. By decomposing the raw wind data into different subsequences, these methods extract implicit modes from the original sequence while reducing noise. Ultimately, the final prediction results are reconstructed by summarizing the predictions of the subsequences. Decomposition-based methods can be broadly categorized into two types. One generates all subsequences at the same hierarchical level, including variational mode decomposition (VMD) [49], adaptive VMD (AVMD), optimal VMD (OVMD) [50], local mean decomposition (LMD), empirical mode decomposition (EMD) [51], ensemble EMD (EEMD) [52], fast EEMD (FEEMD), complete EEMD (CEEMD) [53], CEEMD with adaptive noise (CEEMDAN), improved CEEMDAN (ICEEMDAN) [54], singular spectrum analysis (SSA), improved SSA (ISSA) [55], and symmetric geometry mode decomposition (SGMD). The VMD itself, as well as AVMD and OVMD, demonstrate excellent performance in handling nonlinear and non-stationary signals. They exhibit high accuracy in extracting local signal features, making them suitable for wind speed or wind power data with complex dynamic characteristics. However, these methods may perform less effectively in the presence of significant noise and prove sensitive to outliers within the data. EMD and its derivatives, such as EEMD, FEEMD, and CEEMD, also show promising results in dealing with nonlinear and non-stationary signals, displaying robustness to noise and outliers. Nevertheless, the computational complexity of EEMD and its fast version, FEEMD, is relatively high, potentially leading to suboptimal performance on large-scale datasets. While CEEMD and its variants, CEEMDAN and ICEEMDAN, have made progress in mitigating mode mixing issues, their computational costs still need to be carefully considered. SSA and its improved version, ISSA, excel in the frequency domain analysis of signals. They are proficient at capturing the periodicity and spectral structure within signals, exhibiting superior performance when dealing with wind speed or wind power data characterized by distinct frequency domain features. However, their performance in handling nonlinear and non-stationary signals might be comparatively weaker. SGMD is a relatively new method with the potential to handle nonlinear and non-stationary signals. By leveraging the advantages of symmetry and geometric structure, it is capable of extracting local features from signals. Nevertheless, due to its novelty, further empirical research is needed to confirm its applicability and performance across different types of wind speed or wind power data. The other type mainly relies on wavelet-based methods, including wavelet transform (WT), discrete wavelet transform (DWT) [56], wavelet packet transform (WPT), empirical wavelet transform (EWT), and wavelet soft threshold denoising (WSTD) [57]. These methods have certain advantages in handling high-frequency information and sudden events. These methods offer multiscale resolution in the frequency domain, making them suitable for wind speed or wind power data with different time-scale characteristics. However, they may have limitations when dealing with nonlinear and non-stationary signals, potentially lacking sensitivity to some complex dynamic features of wind fields. In addition to individual decomposition methods, there are some proposed combination decomposition methods designed to overcome the drawbacks of single algorithms and random errors. Table 1 presents several combination decomposition methods.

Table 1. Combined decomposition methods.

Article	Year	Type	Methods
[58]	2022	WSP	EEMD + WPT
[59]	2023	WSP	VMD + SSA
[60]	2023	WPP	CEEMDAN + VMD
[61]	2022	WSP	CEEMDAN + LMD
[62]	2023	WSP	WT + VMD
[63]	2023	WSP	OVMD + DWT

In summary, data preprocessing methods can be categorized into outlier detection methods and decomposition-based methods. Outlier detection methods focus on handling abnormal data in the original dataset to reduce the adverse impact of outliers on model training. These methods typically identify and remove outliers from the data by setting thresholds or using statistical approaches. However, their effectiveness depends on the sensitivity to the definition of outliers, requiring careful selection of appropriate methods and parameters. On the other hand, decomposition-based methods treat all data equally, achieving denoising while discovering implicit patterns in the data, thereby enhancing data predictability. These methods decompose the data to extract primary features and suppress noise. However, the scalability and applicability of these methods depend on the structure of the data and the choice of decomposition methods. In practical applications, a combination of these methods may be used as needed to achieve a more comprehensive and robust data preprocessing effect. There are some limitations and potential drawbacks in existing data processing methods when dealing with the uncertainty associated with wind data. Firstly, data processing methods are inevitably influenced by the quality of the data. Secondly, certain data processing methods may be built upon assumptions that overly simplify the characteristics of wind data, potentially hindering their ability to comprehensively consider the complexity of wind data uncertainty. Additionally, in the process of data processing, uncertainty may be transmitted, and these errors could be magnified in subsequent stages, impacting the reliability of the final predictions. Lastly, some methods rely on offline data processing or training static models, which may limit their applicability when facing dynamic meteorological conditions.

3.2. Feature Extraction

A thoughtfully designed feature extraction process contributes to mitigating the adverse effects of redundant information during the prediction model training phase. Simultaneously, it helps discover more effective features, thereby improving the accuracy and robustness of WSP/WPP. Feature extraction methods for WSP/WPP can be broadly categorized into two types: classical feature extraction methods and neural network-based feature extraction methods.

3.2.1. Classical Feature Extraction Methods

Classical feature extraction methods encompass phase space reconstruction (PSR) [16], granger causality testing (GCT) [64], autocorrelation function (ACF), partial ACF (ACF), recursive feature elimination (RFE) [65], mutual information (MI), grey relation analysis (GRA) [66], and principal component analysis (PCA) [67]. Unlike other selection-based feature extraction methods, PCA is a dimensionality reduction-based feature selection method that generates new features during the feature extraction process. Due to their different principles, each feature extraction method represents features from different aspects of the data. For instance, PSR is more suitable for nonlinear feature extraction; GCT is applicable for analyzing causal relationships between sequences; and MI reveals nonlinear relationships and information transfer between variables. ACF is suitable for analyzing periodic patterns in sequences, and PCA identifies major features through dimensionality reduction, thereby preventing overfitting. In order to select the optimal features, multiple feature extraction methods were presented several years ago. However, unfortunately, no similar combination methods have been found in the recent three years.

3.2.2. Neural Network-Based Methods

While combining various classical feature extraction methods helps in selecting better input features, it is ineffective in extracting deep and highly nonlinear features from complex wind data. Methods such as autoencoder (AE), variational AE (VAE), restricted Boltzmann machine (RBM), CNN [68], temporal convolutional network (TCN) [69], and attention mechanism [70] have been proven to be effective tools for nonlinear feature extraction and widely applied in the field of WSP and WPP. In [65], the VMD technique was employed to decompose the original wind speed sequence, obtaining relatively stable wind speed sequences. Subsequently, sparse AE was used to learn various latent influential features in these stable sequences, crucial for enhancing the predictive performance of subsequent models. In [71], Gauss–Bernoulli RBM and Bernoulli–Bernoulli RBM were combined in a deep belief network (DBN), using GBRBM as the initial RBM to transform the continuity features of the original wind speed data into binomial distribution features. In [72], a one-dimensional CNN was employed as an encoder to extract important features and form latent representations. The decoding network, bidirectional LSTM (BiLSTM) [73], predicted wind speed by interpreting the encoded features. In [69], TCN was used to extract hidden temporal features from the dataset, which were then passed to the informer model for WPP. The results indicated that introducing TCN could better extract potential correlations between data, making the prediction curve closer to the actual values. AE, VAE, and RBM excel in learning hierarchical representations and handling nonlinear relationships; however, their performance may be influenced by hyperparameter choices and might struggle to capture long-range dependencies in wind data. CNN and TCN demonstrate exceptional performance in extracting spatial and temporal features from wind data. CNN is adept at capturing spatial patterns, while TCN excels in modeling long-range dependencies in time series data. Nevertheless, these methods may require substantial data for effective training, and they could be sensitive to changes in hyperparameters. Inspired by human visual attention, attention mechanisms effectively capture important information when dealing with the importance of different parts of input data. However, attention mechanisms may necessitate more complex model architectures and longer training times.

In addition, NNs such as RNN, LSTM, gated recurrent unit (GRU), bidirectional GRU (BiGRU), BiLSTM, graph neural network (GNN), graph convolutional network (GCN), Transformer, and graph Transformer network (GTN) are widely used for learning the mapping relationships between inputs and outputs. These NNs can be applied to feature extraction tasks, automatically learning deep nonlinear features in a hierarchical and interpretable manner while learning the input–output mapping relationships. Table 2 summarizes more feature extraction methods based on NNs.

Table 2. Neural network-based feature extraction methods.

Article	Year	Type	Methods
[74]	2023	WPP	GRU + Transformer + Attention
[75]	2022	WSP	Attention + GTN
[76]	2023	WSP	VMD + GNN
[77]	2022	WPP	CNN + BiGRU
[78]	2023	WPP	GCN + CNN
[79]	2023	WPP	CNN + GRU

In summary, feature extraction methods can be classified into classical feature extraction methods and NN-based feature extraction methods. Classical feature extraction methods, including selection-based methods like ACF and dimensionality reduction methods like PCA, represent different aspects of the data. These methods tend to perform well when dealing with small-scale data or when specific features need to be manually selected. However, they may face limitations in handling large-scale high-dimensional data and learning complex nonlinear relationships. NN-based methods, on the other hand, excel in extracting deep nonlinear features. Unsupervised learning models such as AE, VAE, and RBM demonstrate good performance in learning the latent representations of the data. Nevertheless, their training on large-scale data and sensitivity to hyperparameter selection can be challenging. CNN models, which move convolutional kernels in different dimensions, are suitable for feature extraction from vectors, matrices, and tensors but may require substantial data for effective training. TCN, designed specifically for time series data, captures long-term dependencies in sequences, but its performance may be influenced by sequence length and hyperparameters. The attention mechanism, as a signifier of varying attention levels to different parts of input data during the feature extraction process, offers flexible features but might come with higher computational costs. NNs like LSTM and GRU, employed for learning the mapping between inputs and outputs, may impact model performance due to training complexity and parameter adjustments. In practical applications, the selection and combination of these methods should be balanced based on the nature of the specific task, the characteristics of the data, and the availability of computational resources to achieve optimal feature extraction results.

3.3. Relationship Learning

Relationship learning refers to predictive methods that simulate the complex nonlinear relationships between input features and forecasted wind speed or wind power. Using or designing suitable prediction methods is crucial for accurate WSP/WPP. In reviewing the literature, AI-based relationship learning methods can be broadly categorized into two main types: single prediction models and hybrid prediction models.

3.3.1. Single Prediction Models

Single prediction models can be classified by complexity into simple nonlinear regression models, tree-based models, and DL models. Simple nonlinear regression models include support vector machine (SVM) [80], least squares SVM (LSSVM), support vector regression (SVR) [81,82], ELM [83], kernel ELM (KELM) [23,46], and various types of ANNs such as BPNN [84], radius basis function neural network (RBFNN), multilayer perceptron (MLP), wavelet neural network (WNN), and Elman neural network (ENN) [85]. Tree-based models encompass decision tree (DT), RF [86], gradient boosting decision tree (GBDT), gradient boosting regression tree (GBRT), extreme gradient boosting (XGBoost) [87], and light gradient boosting machine (LightGBM) [88]. With the advancement of DL technologies, deep neural networks (DNNs), including RNN [89], LSTM [90], BiLSTM [91], GRU [92], BiGRU, DBN, deep ELM (DELM), and Transformer have been widely applied in WSP and WPP due to their outstanding capability in handling complex nonlinear problems. Ding et al. [90] used CEEMD to decompose the non-stationary wind power time series into a series of relatively stationary components. Subsequently, these components are employed as the training set for the KELM prediction model, and the initial values and thresholds of KELM are optimized through WOA. Finally, the predicted output values of each component are aggregated to obtain the ultimate forecast of wind power. In the validation using data from the Shanghai wind farm in China, this approach demonstrates superior performance in terms of prediction accuracy and stability, with computational costs lower than those of other benchmark models. Table 3 summarizes more key findings and drawbacks of single prediction models.

The advantages of simple nonlinear regression models lie in their effective modeling of simple nonlinear relationships, such as SVM and ELM, which can efficiently capture nonlinear patterns in the data. However, their performance may be influenced by the selection of hyperparameters and may not be flexible enough for complex nonlinear problems. Tree-based models and their ensemble models, such as RF and GBDT, exhibit good interpretability and modeling capabilities for nonlinear relationships. However, when dealing with high-dimensional sparse data, these models may become overly complex and prone to overfitting. DNNs excel in handling complex nonlinear problems, especially for time series data, with capabilities such as RNN, LSTM, and Transformer that capture long-range dependencies. However, they often require substantial amounts of data for training, are sensitive to hyperparameters, and exhibit lower interpretability. When selecting models, a balance should be struck based on the nature of the specific problem, the characteristics of the data, and the requirement for model interpretability.

Table 3. Single prediction models.

Year	Type	Methods	Results	Drawbacks (Future Work)
2022	WSP	Neo4j + k-means clustering + gray wolf algorithm + SVM [93]	High accuracy, well-predictive stability, and acceptable time complexity	Further optimization
2022	WSP	VMD + partial least squares +improved atom search optimization + ELM [94]	Superior predictive performance to the other nine benchmark models	Not considering the impact of relevant environmental factors
2022	WSP	CEEMDAN + particle swarm optimization algorithm (PSO) + ENN [95]	Significantly improving wind speed prediction effectiveness and reducing errors	Exploring alternative decomposition and optimization algorithms
2022	WPP	CEEMD + whale optimization algorithm (WOA) + KELM [96]	The minimum values of *MAE*, *RMSE*, and *MAPE*: 0.2911%, 0.4305%, and 6.6%	
2022	WSP	CEEMDAN + genetic algorithm (GA) + BPNN [97]	Improvement in accuracy for the short-term and ultra-short-term prediction	Potential performance decline in ultra-short-term prediction using decomposition methods
2022	WSP	Multi-Task Lasso + MLP [98]	Significant effectiveness in feature selection and an increase in prediction accuracy of over 17%	
2022	WPP	Bayesian optimization + XGBoost [99]	Minimal prediction errors under conditions of extreme weather	Integration of additional meteorological elements for further improvement and testing
2023	WPP	skip-GRU [100]	Generating prediction intervals of higher quality than the baseline model	Involving multi-step interval forecasting considering numerical weather forecast information
2023	WPP	EEMD + PSO + LSTM [101]	Higher prediction accuracy and stability compared to other baseline models	
2022	WSP	EEMD + Attention + Transformer [102]	Achieving a new level of multi-step WSP on the NWTC-M2 dataset	Incomplete hyperparameter optimization, simple model structure, and a lack of in-depth error analysis

3.3.2. Hybrid Prediction Models

To overcome the limitations of individual prediction models and enhance the overall WSP/WPP performance, various hybrid prediction models have been proposed. These methods involve combining multiple single models in a sequential or parallel manner to enhance the model's learning ability for complex relationships. Additionally, by designing different error correction strategies, the output of prediction models can be adjusted to improve prediction accuracy. Sun et al. [103] integrated linear time series regression and nonlinear ML algorithms to design a composite WSP model that outperformed statistical models and univariate NN models in the 3–24 h prediction. Chen et al. [104] applied a multi-objective error regression method for error correction in prediction results, resulting in an improvement of over 26% in short-term WPP performance. Xiao et al. [105] proposed a time-correlated, statistics-based error correction algorithm, further enhancing the prediction accuracy of the short-term wind power hybrid prediction model. After error correction, the *RMSE* and mean absolute error (*MAE*) were reduced by 9.14% and 14.96%, respectively. More comprehensive reviews of hybrid prediction models refer to Table 4. Yin et al. [106] and Yang et al. [107], on the other hand, utilized Q-learning for model integration. As for [102] and [108], they, respectively, combined multiple models through weighting and data-adaptive censoring strategy. Both Xing et al. [101] and Duan et al. [109] employed data preprocessing methods; they first decomposed the raw data, independently predicted each component, and finally integrated the predictions through residual correction.

Table 4. Hybrid prediction models.

Article	Year	Type	Methods	Datasets and Results
[106]	2023	WSP	GRU + Broad Learning System (BLS)	Wind farm data from both countries had smaller prediction errors.
[107]	2023	WPP	SVM + LSTM + GRU + GBRT	Using a publicly available dataset from a 16 MW wind farm as a data source, *MSE*, *MAE*, and R^2 were 0.0244, 0.1185, and 0.9821, respectively.
[108]	2022	WSP	SVR + BiLSTM	The method of selecting the historical data sets of Lanzhou and Gaolan wind farms in Hexi Corridor has higher accuracy and stronger generalization ability.
[109]	2023	WSP	LSSVM + LSTM	With the open source data, the method has higher prediction accuracy.
[110]	2023	WPP	Seq2Seq + Attention + BiGRU	The dataset employed consists of measured data from a wind power station located in mainland China, and compared with other models, *MAE* decreased by 19.7%, 43.4%, 41.0%, and 46.2%, respectively.
[111]	2021	WSP	ELM + Outlier-Robust Extreme Learning Machine (ORELM) + DBN	Hourly wind speed data from the National Renewable Energy Laboratory, *RMSE*, *MAE*, *MAPE*, and symmetric *MAPE* reached 0.2047%, 0.1435%, 3.77%, and 3.74%, respectively.
[112]	2022	WPP	Group Method of Data Handling (GMDH) + Echo State Network (ESN) + ELM	Four group wind power datasets; the accuracy is superior to other five AI-based models.
[113]	2022	WSP	GRU + BiLSTM + DBN	*MAE* of the four original wind speed series collected from Xinjiang, China, at stations #1, #2, #3, and #4 are 0.0829 m/s, 0.0661 m/s, 0.0906 m/s, and 0.0803 m/s, respectively.
[114]	2022	WSP	MLP + RNN + SVM	Simulation results on actual large-scale short-term wind speed data validate the above attractive features of the proposed predictor.
[115]	2021	WSP	RNN + BPNN	Four wind speed series collected from a wind farm in Ningxia Hui Autonomous Region, China, outperformed other single models and traditional models.

In summary, the prediction methods for learning the relationship between input features and output wind speed or wind power can be categorized into single prediction models and hybrid prediction models. Single prediction models mainly include simple nonlinear regression models, tree-based models, and DL models. Among them, simple nonlinear regression models represented by SVM and ELM have the characteristics of simple model structure and fast training speed. However, when dealing with nonlinear relationships in complex large datasets, their accuracy is inferior to tree-based models and DL models. The goal of hybrid prediction models is to overcome the limitations of single prediction models and enhance the overall WSP/WPP performance. Approaches include combining single models and postprocessing methods such as error correction. Combining single models can be achieved through sequential (stacking) or parallel (weighted) methods to improve the accuracy and robustness of the prediction. Error correction involves using statistical methods or ML-based techniques to fit errors, thereby reducing the gap between predicted values and actual values. For small datasets and applications emphasizing training speed, simple nonlinear regression models, such as SVM and ELM, may be suitable choices. For large-scale and complex datasets, tree-based models and deep learning models often excel in capturing nonlinear relationships, offering higher accuracy. Hybrid models often outperform single models in terms of prediction accuracy; however, this comes at the cost of time.

In practical applications, the choice of models is influenced by the specific requirements of the application scenario and the characteristics of the data. Considering the effectiveness, scalability, and computational resource requirements of the models, a comprehensive evaluation is necessary based on the specific circumstances. For simple models, they typically perform well in scenarios dealing with small datasets where training speed is crucial, making them suitable for resource-constrained environments. These models are not only computationally lightweight but may also provide satisfactory performance for straightforward tasks. In contrast, tree-based models and deep learning models are generally more suitable for handling large-scale, complex datasets, possessing enhanced capabilities for capturing nonlinear relationships. However, these models may demand more computational resources, especially during the training phase. Given sufficient resources, they can deliver superior performance for more complex tasks. Regarding the scalability and computational resource requirements of models, for large-scale data and complex tasks, deep learning models are usually considered. These models, with support from hardware accelerators such as GPUs and TPUs, can train and infer more rapidly but require corresponding hardware investments. In terms of model combination, the choice between sequential or parallel strategies also requires a comprehensive consideration of scalability and computational resources. While parallel training may enhance efficiency, it may also increase the demand for computational resources. Therefore, when selecting combination strategies, a careful balance based on the available resources is needed to achieve optimal overall performance.

3.4. Parameter Optimization

Relationship learning refers to predictive methods that simulate the complex nonlinear relationships between input features and forecasted wind speed or wind power. Using or designing suitable prediction methods is crucial for accurate WSP/WPP. In reviewing the literature, AI-based relationship learning methods can be broadly categorized into two main types: single prediction models and hybrid prediction models.

The performance of different prediction models is largely influenced by the model structure and hyperparameters. Kernel-based models, such as SVM, LSSVM, and KELM, have performance dependent on the choice of the kernel function. In prediction models involving neural networks, parameters like weights, thresholds, hidden layer count, neurons per layer, learning rate, and optimizer type significantly impact model performance. Selecting appropriate parameters is crucial to enhance the predictive performance and generalization capability of the model. Traditional parameter optimization methods often rely on manual tuning, depending on expert experience and knowledge. As the complexity of problems increases and the parameter search space expands, an increasing number of studies are adopting intelligent algorithms to optimize prediction model parameters, proving their effectiveness in improving predictive performance. Table 5 summarizes some intelligent algorithms used in recent studies.

Hu et al. [93] applied the gray wolf algorithm (GWO) to optimize the kernel function parameters and penalty factor of SVM. In [23], the improved artificial bee colony algorithm (CABC) was utilized to optimize the key parameters C and λ of the KELM model, significantly enhancing the accuracy of WPP. Yang et al. [116] and Gao et al. [117] employed the differential evolution–grey wolf optimizer (DE-GWO) and fractional-order beetle swarm optimization (FO-BSO), respectively, to adjust the hyperparameters of LSSVM. In [118], an improved cuckoo search (ICS) was established to optimize the penalty factor and kernel function parameters of LSSVM. Additionally, Ding et al. [96] utilized the WOA to generate optimal initial parameters for KELM, contributing to the accurate short-term prediction of wind power. Through the improved seagull optimization algorithm (ISOA) in [119], the best parameters for each KELM network were precisely identified, enhancing the predictability of individual KELM models for wind speed subsequences after VMD decomposition. In [120], an innovative approach using improved hybrid differential evolution–Harris Hawks optimization (IHDEHHO) synchronized the optimization of internal parameters for PSR and KELM. Furthermore, Rayi et al. [121] optimized the parameters and associated weights of the deep mixed KELM using the sine cosine integrated water cycle algorithm (SCWCA) to enhance the adaptability of deep MKELM to outliers or noise in the data.

Zhang et al. [122] applied the improved sine and cosine algorithm (ISCA) to optimize the parameters of the BiLSTM model, achieving superior optimization results compared to the standard SCA. In [123], to ensure the accuracy and stability of WPP, the key parameters of the DELM model underwent optimization through multi-objective crisscross optimization (MOCSO). In [105], an improvement to the PSO was employed to optimize the optimal number of hidden neurons and the optimal learning rate for the LSTM model, significantly enhancing the short-term accuracy of WPP. Suo et al. [124] utilized the improved chimp optimization algorithm (IChOA) to optimize parameters for the BiGRU, proving its effectiveness in improving the predictive performance of BiGRU. Through swarm spider optimization (SSO), Wei et al. [125] optimized the number of hidden layer nodes for DBN, significantly improving the prediction performance of DBN for wind prediction. Zhu et al. [126], by introducing multi-objective elephant clan optimization (MOECO), generated the final wind power interval prediction, with optimization performance surpassing existing multi-objective optimization algorithms such as multi-objective elephant herding optimization and multi-objective PSO. Using adaptive differential evolution with an optional external archive (JADE), Wu et al. [127] efficiently optimized the parameter combination of the temporal fusion Transformer (TFT) model, ensuring the stability and reliability of WSP. Chen et al. [128], employing a multi-objective slime mold algorithm (MOSMA), conducted multi-objective parameter optimization for DAE and GRU stacking models, achieving performance with low error, low bias, and high qualification rate. In [129], introducing chaos sequences and Gaussian mutation strategy into the original sparrow algorithm, the improved sparrow algorithm (CSSOA) was used to adjust hyperparameters such as batch size, cell number, and learning rate for the LSTM network, thereby enhancing wind power prediction accuracy. Ewees et al. [130] employed the heap-based optimizer (HBO) to optimize the LSTM wind power prediction model, outperforming other metaheuristic optimization algorithms such as GA. However, like other optimization algorithms, HBO faces the challenges of losing the optimal solution and slow convergence speed. Among these optimization algorithms, GWO may excel in global search, while PSO stands out in handling continuous optimization problems with a relatively fast convergence speed. The weakness of the Cuckoo algorithm may lie in the need to adjust a large number of parameters, and when dealing with multi-objective problems, MOCSO and MOECO might be more competitive.

While intelligent algorithms have demonstrated significant effectiveness in optimizing wind forecasting parameters, their practical application requires a careful balance between their applicability and computational efficiency. Handling high-dimensional, large-scale datasets from hundreds of meteorological stations and wind farms still poses uncertainties regarding the performance of metaheuristic algorithms. This uncertainty may stem from the complexity and computational burden associated with these algorithms when dealing with high-dimensional data. Particularly, when addressing DNNs or hybrid models with numerous parameters, careful consideration must be given to the convergence speed and reliability of optimization results using intelligent algorithms. The training of

DNNs typically demands substantial data and intricate hyperparameter tuning, leading to challenges such as prolonged training times and potential overfitting. In real applications, algorithm scalability is crucial, especially for wind prediction tasks that require real-time responsiveness and efficiency. For the improvement and adjustment of existing models, it is recommended to start by optimizing the hyperparameters of intelligent algorithms to enhance their adaptability to high-dimensional datasets. Regarding DNNs, methods like transfer learning could be considered to leverage limited data during the training process better. Additionally, introducing advanced metaheuristic algorithms or deep learning techniques, such as adaptive learning rates and batch normalization, may contribute to improving the convergence speed and stability of optimization results.

Table 5. Intelligent algorithms.

Model Type	Models	Parameters	Optimization Methods
Kernel dependent models	SVM, LSSVM, KELM	Kernel function, etc.	GWO [93], CABC [23], DE-GWO [116], FO-BSO [117], ICS [118], WOA [96], ISOA [119], IHDEHHO [120], SCWCA [121]
NN-based models	ELM, BPNN, RNN, DBN, etc.	Weights, bias, learning rate, etc.	ISCA [122], MOCSO [123], PSO [105], IChOA [124], SSO [125], HBO [130], MOECO [126], JADE [127], MOSMA [128], CSSOA [129]

4. Performance Evolution Metrics

The classical metrics used to assess the performance of wind prediction can be mainly divided into two categories: deterministic prediction evaluation metrics and interval prediction evaluation metrics. Their definitions, equations, and evaluation criteria are detailed in Table 6 [131,132].

The evaluation of predictive performance involves two aspects: accuracy and robustness. Accuracy primarily focuses on the model's precision concerning observed values, while robustness concerns the model's stability when faced with different conditions, outliers, or uncertainties. The mentioned metrics mainly reflect predictive accuracy, and the indicators used to assess predictive robustness include mean absolute scaled error (MASE), median absolute deviation (MAD), and quantile loss. MASE improves upon *MAE* by comparing it with the corresponding *MAE* on training data, providing a measure of the model's robustness in various situations. MAD represents the median absolute deviation between observed and predicted values, measuring the model's sensitivity to outliers. Quantile loss evaluates the model's fitting to the entire distribution by comparing its predictions with actual values at different quantiles, thereby assessing its robustness. Among the metrics for evaluating predictive accuracy, *MSE* and *RMSE* involve the square of errors and are sensitive to error magnitudes. *MAE* calculates the average absolute value of errors, exhibiting lower sensitivity to outliers compared to *MSE*. *MAPE* is suitable for scenarios where percentage errors are crucial. The correlation coefficient (r) gauges the strength and direction of the linear relationship between actual and predicted values, while R^2 measures the proportion of variance explained by the model. *PICP*, *PINAW*, and *CWC* are commonly used evaluation metrics for interval prediction problems. The choice of metrics depends on the specific problem type and the focus on model performance.

Table 6. Performance evolution metrics.

Type	Index	Definition	Equations	Evaluation Criteria
Deterministic prediction evaluation metrics	MSE	Mean squared error	$MSE = \frac{1}{N}\sum_{i=1}^{N} \frac{f_i - \hat{f}_i}{f_i}$	❄
	$RMSE$	Root mean square error	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (f_i - \hat{f}_i)^2}$	❄
	MAE	Mean absolute error	$MAE = \frac{1}{N}\sum_{i=1}^{N}\left\lvert f_i - \hat{f}_i \right\rvert$	❄
	$MAPE$	Mean absolute percentage error	$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left\lvert \frac{f_i - \hat{f}_i}{f_i} \right\rvert \times 100\%$	❄
	r	Correlation coefficient	$r = \frac{\sum_{i=1}^{N} (f_i - \overline{f}_i)(\widehat{f}_i - \overline{\hat{f}}_i)}{\sqrt{\sum_{i=1}^{N} (f_i - \overline{f}_i)^2 (\widehat{f}_i - \overline{\hat{f}}_i)^2}}$	⏱
	R^2	Coefficient of determination	$R^2 = 1 - \frac{\sum_{i=1}^{N} (f_i - \hat{f}_i)^2}{\sum_{i=1}^{N} (f_i - \overline{f}_i)^2}$	⏱
Interval prediction evaluation metrics	$PICP$	Prediction interval coverage probability	$PICP = 1 - \frac{1}{N}\sum_{i=1}^{N} c_i, c_i = \begin{cases} 1, \hat{f}_i \in [L_i, U_i] \\ 0, \hat{f}_i \notin [L_i, U_i] \end{cases}$	⏱
	$PINAW$	Prediction intervals normalized average width	$PINAW = \frac{1}{N}\sum_{i=1}^{N} \frac{U_i - L_i}{f_{\max} - f_{\min}}$	❄
	CWC	Composite indacator	$CWC = \begin{cases} PINAW(1 + e^{(-\eta(PICP-\mu)),PICP<\mu} \\ PINAW, PICP \geq \mu \end{cases}$	❄

Note: The actual and predicted values are represented as f_i and $\hat{f}_i$, respectively. $\overline{f}_i$ and $\overline{\hat{f}}_i$ stand for the average values of the actual and predicted values, respectively. N is the length of the sequences. L_i and U_i denote the lower and upper bounds of the ith predicted interval. c_i equals to 1 when the actual value is within [L_i, U_i]. $f_{\max}$ and $f_{\min}$ denote the maximum and minimum values in sequences, respectively. μ symbolizes the preset PINC level, and η is equal to 50 based on recent literatures. ❄ represents the smaller the value, the better the prediction performance. ⏱ presents the closer the value is to 1, the better the prediction performance.

5. Challenges and Future Trends

In this section, challenges in AI-based wind prediction over the past three years are discussed, and future trends aimed at promoting prediction performance are presented.

5.1. Challenges over the Past Three Years

AI-based wind prediction faces three major challenges, including the following:

✓ Uncertainty in wind data

The intricate nature of uncertainty in wind data makes it challenging for predictive models to accurately learn the actual relationship between input features and wind speed/wind power output. Various types of uncertainty may exist in wind data, and these uncertainties may manifest diversely within datasets. Consequently, selecting appropriate preprocessing methods tailored to different types of uncertainty may pose a challenge. In practical applications, there is currently a lack of clear evidence indicating that a specific preprocessing method consistently outperforms others, likely due to variations in the types and complexity of uncertainty across different datasets.

✓ Comprehensive and efficient feature extraction

Insufficient data can lead to information loss and uncertainty during the feature extraction process. In practical applications, limited wind data may hinder feature extraction methods from fully capturing the complex relationships between wind speed and power, thereby constraining the performance of predictive models. The dynamic changes and diversity in wind fields further contribute to the complexity of features, a phenomenon commonly observed in real wind scenarios. While traditional methods boast lower computational complexity, they fall short in effectively handling the deep-seated, highly nonlinear features within wind data, potentially resulting in decreased model accuracy in practical wind prediction scenarios. On the contrary, feature extraction methods based on NNs exhibit superior performance when confronted with real-time, complex, and high-dimensional wind field data. However, the increased computational complexity associated with NN-based methods may pose challenges in terms of computational resources in practical applications, particularly in scenarios demanding real-time processing.

✓ Accurate learning of mapping relationship

In wind prediction, there are complex relationships among various factors, and modeling these intricate relationships requires flexible and robust methods of relationship learning. AI-based relationship learning methods can be categorized into single methods and hybrid methods. Single methods, such as simple nonlinear regression models, have a straightforward structure and low computational complexity but may lack the flexibility needed to handle complex nonlinear problems. On the other hand, tree-based models and deep learning models excel in addressing complex nonlinear problems. However, tree-based models might become overly complex when dealing with high-dimensional sparse data, leading to potential overfitting. DL models typically require substantial amounts of data for training, are sensitive to hyperparameters, and exhibit lower interpretability. Hybrid models can overcome the limitations of single models, but selecting and fine-tuning the combination of different models may require additional expertise and time. Furthermore, combining multiple models may lead to an increase in computational complexity, particularly in applications with high demands for real-time performance, presenting potential challenges. In practical applications, these challenges manifest as the complexity of models in handling real-time data, adapting to dynamic environments, and providing interpretable results. It requires a balance between model performance and computational efficiency.

5.2. Future Trends

Considering the above challenges, to further promote the performance of AI-based wind prediction, the following aspects can be focused on:

❄ Optimize data processing methods to adapt to specific geographical or meteorological conditions. Customized data processing methods may be more effective in particular environments and conditions. Furthermore, by developing new methods with flexibility and applicability and validating their performance in practical applications, we aim to gradually overcome the limitations of existing methods in handling wind data uncertainty.

❄ During feature extraction, it is essential to consider a comprehensive range of factors to ensure the thorough capture of various potential elements influencing wind speed and wind power. Simultaneously, attention must be paid to avoiding the introduction of redundant information to mitigate the risk of overfitting. In the context of long-term predictions, regularly updating the feature set is crucial to reflect potential time-varying influencing factors, contributing to maintaining the robustness of the model when facing dynamic environmental conditions.

❄ In enhancing the nonlinear fitting capability of the prediction model, designing appropriate hybrid strategies allows leveraging the strengths of single models to better adapt to complex relationships. However, when selecting and optimizing hybrid strategies, consideration must be given to the specific requirements of the application and constraints of computational resources, striking a balance between model performance and complexity.

❄ Due to the involvement of a large number of parameters in DL models and hybrid models, a combined approach of manual tuning and intelligent algorithms can be adopted. Manual tuning leverages professional knowledge and experience to adjust parameters, enhancing model performance. Simultaneously, intelligent algorithms can explore optimal solutions within the parameter space, improving optimization efficiency. The comprehensive application of these two methods allows for a more thorough exploration of parameter combinations, ensuring reliable model performance in various scenarios.

6. Conclusions

This paper systematically summarizes and organizes AI-based wind prediction research conducted from 2021 to 2023. Focusing on critical steps, including data preprocessing, feature extraction, relationship learning, and parameter optimization, delves into in-depth analysis and discussion of the key technologies and models involved in these crucial processes.

Data preprocessing methods are divided into outlier detection methods and decomposition-based methods, aimed at handling abnormal information in raw wind data and extracting implicit patterns in complex data. Feature extraction methods include traditional methods and NN-based methods, which are used to discover more effective features and eliminate redundant information. The relationship learning process simulates the precise mapping between input features and wind speed or wind power output, covering single prediction models and hybrid prediction models. Single models can be simple nonlinear regression models, tree-based models, and DL models, while hybrid methods include combining single models and incorporating postprocessing techniques. The specific method chosen depends on the scenario task and data characteristics. Parameter optimization primarily refers to optimizing hyperparameters in prediction models to improve their predictive performance by adjusting model structure and parameters. Currently, parameter optimization often adopts popular intelligent optimization algorithms, gradually replacing traditional manual tuning methods. In practical applications, computational cost is often a key factor. Particularly for predictive models intended to operate in large-scale or real-time environments, the computational cost directly influences the feasibility and practicality of the model. Therefore, balancing model performance and computational cost is paramount in the process of model design and optimization.

To further promote the performance in AI-based wind prediction, challenges in AI-based wind prediction over the past three years are discussed, and future trends aimed at promoting prediction performance are presented. This review provides guidance for researchers dedicated to promoting the effective application of AI technology in the field of wind prediction.

Author Contributions: Conceptualization, D.S. and X.T.; methodology, D.S.; software, X.T.; validation, D.S., X.T. and Q.H.; formal analysis, L.W.; investigation, M.D.; resources, J.Y.; data curation, L.W.; writing—original draft preparation, X.T.; writing—review and editing, S.E.; visualization, M.D.; supervision, Q.H.; project administration, L.W.; funding acquisition, J.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported under the framework of an international cooperation program managed by the National Natural Science Foundation of China under Grant 62211540397 and the National Research Foundation of Korea (NRF-2022K2A9A2A06045121), the Natural Science Foundation of Hunan Province (2021JJ30875), and the Natural Science Foundation of Changsha (kq2208288).

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Yi, Z.; Chen, Z.; Yin, K.; Wang, L.; Wang, K. Sensing as the key to the safety and sustainability of new energy storage devices. *Prot. Control Mod. Power Syst.* **2023**, *8*, 27. [CrossRef]
2. Dong, Y.; Ma, S.; Zhang, H.; Yang, G. Wind Power Prediction Based on Multi-class Autoregressive Moving Average Model with Logistic Function. *J. Mod. Power Syst. Clean Energy* **2022**, *10*, 1184–1193. [CrossRef]

3. Liao, W.; Wang, S.; Bak-Jensen, B.; Pillai, J.R.; Yang, Z.; Liu, K. Ultra-Short-Term Interval Prediction of Wind Power Based on Graph Neural Network and Improved Bootstrap Technique. *J. Mod. Power Syst. Clean Energy* **2023**, *11*, 1100–1114. [CrossRef]
4. Chen, R.; Zhang, Z.; Hu, J.; Zhao, L.; Li, C.; Zhang, X. Grouping-Based Optimal Design of Collector System Topology for a Large-Scale Offshore Wind Farm by Improved Simulated Annealing. *Prot. Control Mod. Power Syst.* **2024**, *9*, 94–111. [CrossRef]
5. Song, D.; Shen, G.; Huang, C.; Huang, Q.; Yang, J.; Dong, M.; Joo, Y.H.; Duić, N. Review on the Application of Artificial Intelligence Methods in the Control and Design of Offshore Wind Power Systems. *J. Mar. Sci. Eng.* **2024**, *12*, 424. [CrossRef]
6. Global Wind Energy Council. *GWEC Global Wind Report 2023*; Global Wind Energy Council: Bonn, Germany, 2023.
7. Song, D.; Tu, Y.; Wang, L.; Jin, F.; Li, Z.; Huang, C.; Xia, E.; Rizk-Allah, R.M.; Yang, J.; Su, M. Coordinated optimization on energy capture and torque fluctuation of wind turbines via variable weight NMPC with fuzzy regulator. *Appl. Energy* **2022**, *312*, 118821. [CrossRef]
8. Zhu, Y.; Zhang, Y.; Wei, Z. Hierarchical cluster coordination control strategy for large-scale wind power based on model predictive control and improved multi-time-scale active power dispatching. *J. Mod. Power Syst. Clean Energy* **2022**, *11*, 827–839. [CrossRef]
9. 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]
10. Tsai, W.-C.; Hong, C.-M.; Tu, C.-S.; Lin, W.-M.; Chen, C.-H. A review of modern wind power generation forecasting technologies. *Sustainability* **2023**, *15*, 10757. [CrossRef]
11. Duan, Z.; Liu, H.; Li, Y.; Nikitas, N. Time-variant post-processing method for long-term numerical wind speed forecasts based on multi-region recurrent graph network. *Energy* **2022**, *259*, 125021. [CrossRef]
12. Guan, S.; Wang, Y.; Liu, L.; Gao, J.; Xu, Z.; Kan, S. Ultra-short-term wind power prediction method based on FTI-VACA-XGB model. *Expert Syst. Appl.* **2024**, *235*, 121185. [CrossRef]
13. Ji, L.; Fu, C.; Ju, Z.; Shi, Y.; Wu, S.; Tao, L. Short-Term canyon wind speed prediction based on CNN—GRU transfer learning. *Atmosphere* **2022**, *13*, 813. [CrossRef]
14. Ali, N.; Calaf, M.; Cal, R.B. Clustering sparse sensor placement identification and deep learning based forecasting for wind turbine wakes. *J. Renew. Sustain. Energy* **2021**, *13*, 023307. [CrossRef]
15. Geibel, M.; Bangga, G. Data reduction and reconstruction of wind turbine wake employing data driven approaches. *Energies* **2022**, *15*, 3773. [CrossRef]
16. Huang, Y.; Zhang, B.; Pang, H.; Wang, B.; Lee, K.Y.; Xie, J.; Jin, Y. Spatio-temporal wind speed prediction based on Clayton Copula function with deep learning fusion. *Renew. Energy* **2022**, *192*, 526–536. [CrossRef]
17. Liu, R.; Song, Y.; Yuan, C.; Wang, D.; Xu, P.; Li, Y. GAN-Based Abrupt Weather Data Augmentation for Wind Turbine Power Day-Ahead Predictions. *Energies* **2023**, *16*, 7250. [CrossRef]
18. Wang, B.; Wang, T.; Yang, M.; Han, C.; Huang, D.; Gu, D. Ultra-Short-Term Prediction Method of Wind Power for Massive Wind Power Clusters Based on Feature Mining of Spatiotemporal Correlation. *Energies* **2023**, *16*, 2727. [CrossRef]
19. Wu, F.; Yang, M.; Shi, C. Short-Term Prediction of Wind Power Considering the Fusion of Multiple Spatial and Temporal Correlation Features. *Front. Energy Res.* **2022**, *10*, 878160. [CrossRef]
20. Yang, L.; Zheng, Z.; Zhang, Z. An improved mixture density network via wasserstein distance based adversarial learning for probabilistic wind speed predictions. *IEEE Trans. Sustain. Energy* **2021**, *13*, 755–766. [CrossRef]
21. Zhu, S.; Chen, X.; Luo, X.; Luo, K.; Wei, J.; Li, J.; Xiong, Y. Enhanced probabilistic spatiotemporal wind speed forecasting based on deep learning, quantile regression, and error correction. *J. Energy Eng.* **2022**, *148*, 04022004. [CrossRef]
22. Jiang, Z.; Che, J.; Li, N.; Tan, Q. Deterministic and probabilistic multi-time-scale forecasting of wind speed based on secondary decomposition, DFIGR and a hybrid deep learning method. *Expert Syst. Appl.* **2023**, *234*, 121051. [CrossRef]
23. Shan, J.-N.; Wang, H.-Z.; Pei, G.; Zhang, S.; Zhou, W.-H. Research on short-term power prediction of wind power generation based on WT-CABC-KELM. *Energy Rep.* **2022**, *8*, 800–809. [CrossRef]
24. Liu, G.; Wang, C.; Qin, H.; Fu, J.; Shen, Q. A novel hybrid machine learning model for wind speed probabilistic forecasting. *Energies* **2022**, *15*, 6942. [CrossRef]
25. Zhang, J.; Zhang, R.; Zhao, Y.; Qiu, J.; Bu, S.; Zhu, Y.; Li, G. Deterministic and Probabilistic Prediction of Wind Power Based on a Hybrid Intelligent Model. *Energies* **2023**, *16*, 4237. [CrossRef]
26. Alkabbani, H.; Ahmadian, A.; Zhu, Q.; Elkamel, A. Machine learning and metaheuristic methods for renewable power forecasting: A recent review. *Front. Chem. Eng.* **2021**, *3*, 665415. [CrossRef]
27. Yang, M.; Guo, Y.; Huang, Y. Wind power ultra-short-term prediction method based on NWP wind speed correction and double clustering division of transitional weather process. *Energy* **2023**, *282*, 128947. [CrossRef]
28. Tan, E.; Mentes, S.S.; Unal, E.; Unal, Y.; Efe, B.; Barutcu, B.; Onol, B.; Topcu, H.S.; Incecik, S. Short term wind energy resource prediction using WRF model for a location in western part of Turkey. *J. Renew. Sustain. Energy* **2021**, *13*, 013303. [CrossRef]
29. Zhou, Z.; Dai, Y.; Xiao, J.; Liu, M.; Zhang, J.; Zhang, M. Research on Short-Time Wind Speed Prediction in Mountainous Areas Based on Improved ARIMA Model. *Sustainability* **2022**, *14*, 15301. [CrossRef]
30. Liu, Y.; Wang, Y.; Wang, Q.; Zhang, K.; Qiang, W.; Wen, Q.H. Recent advances in data-driven prediction for wind power. *Front. Energy Res.* **2023**, *11*, 1204343. [CrossRef]
31. Ahsan, F.; Dana, N.H.; Sarker, S.K.; Li, L.; Muyeen, S.; Ali, M.F.; Tasneem, Z.; Hasan, M.M.; Abhi, S.H.; Islam, M.R. Data-driven next-generation smart grid towards sustainable energy evolution: Techniques and technology review. *Prot. Control Mod. Power Syst.* **2023**, *8*, 43. [CrossRef]

32. Hasan, M.N.; Toma, R.N.; Nahid, A.-A.; Islam, M.M.; Kim, J.-M. Electricity theft detection in smart grid systems: A CNN-LSTM based approach. *Energies* **2019**, *12*, 3310. [CrossRef]
33. Zidi, S.; Mihoub, A.; Qaisar, S.M.; Krichen, M.; Al-Haija, Q.A. Theft detection dataset for benchmarking and machine learning based classification in a smart grid environment. *J. King Saud Univ. Comput. Inf. Sci.* **2023**, *35*, 13–25. [CrossRef]
34. Niksa-Rynkiewicz, T.; Stomma, P.; Witkowska, A.; Rutkowska, D.; Słowik, A.; Cpałka, K.; Jaworek-Korjakowska, J.; Kolendo, P. An Intelligent Approach to Short-Term Wind Power Prediction Using Deep Neural Networks. *J. Artif. Intell. Soft Comput. Res.* **2023**, *13*, 197–210. [CrossRef]
35. Valdivia-Bautista, S.M.; Domínguez-Navarro, J.A.; Pérez-Cisneros, M.; Vega-Gómez, C.J.; Castillo-Téllez, B. Artificial Intelligence in Wind Speed Forecasting: A Review. *Energies* **2023**, *16*, 2457. [CrossRef]
36. Wu, Z.; Luo, G.; Yang, Z.; Guo, Y.; Li, K.; Xue, Y. A comprehensive review on deep learning approaches in wind forecasting applications. *CAAI Trans. Intell. Technol.* **2022**, *7*, 129–143. [CrossRef]
37. Feng, H.; Jin, Y.; Laima, S.; Han, F.; Xu, W.; Liu, Z. A ML-Based Wind Speed Prediction Model with Truncated Real-Time Decomposition and Multi-Resolution Data. *Appl. Sci.* **2022**, *12*, 9610. [CrossRef]
38. Li, L.-L.; Liu, Z.-F.; Tseng, M.-L.; Jantarakolica, K.; Lim, M.K. Using enhanced crow search algorithm optimization-extreme learning machine model to forecast short-term wind power. *Expert Syst. Appl.* **2021**, *184*, 115579. [CrossRef]
39. Hanifi, S.; Lotfian, S.; Zare-Behtash, H.; Cammarano, A. Offshore wind power forecasting—A new hyperparameter optimisation algorithm for deep learning models. *Energies* **2022**, *15*, 6919. [CrossRef]
40. Benti, N.E.; Chaka, M.D.; Semie, A.G. Forecasting Renewable Energy Generation with Machine learning and Deep Learning: Current Advances and Future Prospects. *Sustainability* **2023**, *15*, 7087. [CrossRef]
41. Qu, Z.; Li, J.; Hou, X.; Gui, J. A D-stacking dual-fusion, spatio-temporal graph deep neural network based on a multi-integrated overlay for short-term wind-farm cluster power multi-step prediction. *Energy* **2023**, *281*, 128289. [CrossRef]
42. Huang, C.-M.; Chen, S.-J.; Yang, S.-P.; Chen, H.-J. One-Day-Ahead Hourly Wind Power Forecasting Using Optimized Ensemble Prediction Methods. *Energies* **2023**, *16*, 2688. [CrossRef]
43. Khazaei, S.; Ehsan, M.; Soleymani, S.; Mohammadnezhad-Shourkaei, H. A high-accuracy hybrid method for short-term wind power forecasting. *Energy* **2022**, *238*, 122020. [CrossRef]
44. Fahim, P.; Vaezi, N.; Shahraki, A.; Khoshnevisan, M. An Integration of Genetic Feature Selector, Histogram-Based Outlier Score, and Deep Learning for Wind Turbine Power Prediction. *Energy Sources Part A Recovery Util. Environ. Eff.* **2022**, *44*, 9342–9365. [CrossRef]
45. He, Y.; Li, H.; Wang, S.; Yao, X. Uncertainty analysis of wind power probability density forecasting based on cubic spline interpolation and support vector quantile regression. *Neurocomputing* **2021**, *430*, 121–137. [CrossRef]
46. Fu, W.; Fu, Y.; Li, B.; Zhang, H.; Zhang, X.; Liu, J. A compound framework incorporating improved outlier detection and correction, VMD, weight-based stacked generalization with enhanced DESMA for multi-step short-term wind speed forecasting. *Appl. Energy* **2023**, *348*, 121587. [CrossRef]
47. Ammar, E.; Xydis, G. Wind speed forecasting using deep learning and preprocessing techniques. *Int. J. Green Energy* **2024**, *21*, 988–1016. [CrossRef]
48. Özen, C.; Deniz, A. Short-term wind speed forecast for Urla wind power plant: A hybrid approach that couples weather research and forecasting model, weather patterns and SCADA data with comprehensive data preprocessing. *Wind. Eng.* **2022**, *46*, 1526–1549. [CrossRef]
49. Li, J.; Song, Z.; Wang, X.; Wang, Y.; Jia, Y. A novel offshore wind farm typhoon wind speed prediction model based on PSO–Bi-LSTM improved by VMD. *Energy* **2022**, *251*, 123848. [CrossRef]
50. Ai, C.; He, S.; Fan, X.; Wang, W. Chaotic time series wind power prediction method based on OVMD-PE and improved multi-objective state transition algorithm. *Energy* **2023**, *278*, 127695. [CrossRef]
51. Shi, S.; Liu, Z.; Deng, X.; Chen, S.; Song, D. From Lidar Measurement to Rotor Effective Wind Speed Prediction: Empirical Mode Decomposition and Gated Recurrent Unit Solution. *Sensors* **2023**, *23*, 9379. [CrossRef]
52. Sun, H.; Song, T.; Li, Y.; Yang, K.; Xu, D.; Meng, F. EEMD-ConvLSTM: A model for short-term prediction of two-dimensional wind speed in the South China Sea. *Appl. Intell.* **2023**, *53*, 30186–30202. [CrossRef]
53. Xiong, J.; Peng, T.; Tao, Z.; Zhang, C.; Song, S.; Nazir, M.S. A dual-scale deep learning model based on ELM-BiLSTM and improved reptile search algorithm for wind power prediction. *Energy* **2023**, *266*, 126419. [CrossRef]
54. Wang, X.X.; Shen, X.P.; Ai, X.Y.; Li, S.J. Short-term wind speed forecasting based on a hybrid model of ICEEMDAN, MFE, LSTM and informer. *PLoS ONE* **2023**, *18*, e0289161.
55. Yang, Q.; Deng, C.; Chang, X. Ultra-short-term/short-term wind speed prediction based on improved singular spectrum analysis. *Renew. Energy* **2022**, *184*, 36–44. [CrossRef]
56. Zhang, W.; Lin, Z.; Liu, X. Short-term offshore wind power forecasting-A hybrid model based on Discrete Wavelet Transform (DWT), Seasonal Autoregressive Integrated Moving Average (SARIMA), and deep-learning-based Long Short-Term Memory (LSTM). *Renew. Energy* **2022**, *185*, 611–628. [CrossRef]
57. Zhang, Y.; Zhang, J.; Yu, L.; Pan, Z.; Feng, C.; Sun, Y.; Wang, F. A short-term wind energy hybrid optimal prediction system with denoising and novel error correction technique. *Energy* **2022**, *254*, 124378. [CrossRef]
58. Zhang, Y.; Zhang, W.; Guo, Z.; Zhang, S. An effective wind speed prediction model combining secondary decomposition and regularised extreme learning machine optimised by cuckoo search algorithm. *Wind Energy* **2022**, *25*, 1406–1433. [CrossRef]

59. Qiu, W.; Zhang, W.; Wang, G.; Guo, Z.; Zhao, J.; Ma, K. Combined wind speed forecasting model based on secondary decomposition and quantile regression closed-form continuous-time neural network. *Int. J. Green Energy* **2023**, 1–22. [CrossRef]
60. Hou, G.; Wang, J.; Fan, Y. Multistep short-term wind power forecasting model based on secondary decomposition, the kernel principal component analysis, an enhanced arithmetic optimization algorithm, and error correction. *Energy* **2024**, *286*, 129640. [CrossRef]
61. Emeksiz, C.; Tan, M. Multi-step wind speed forecasting and Hurst analysis using novel hybrid secondary decomposition approach. *Energy* **2022**, *238*, 121764. [CrossRef]
62. Ma, K.; Zhang, W.; Guo, Z.; Zhao, J.; Qiu, W. A hybrid forecasting model for very short-term wind speed prediction based on secondary decomposition and deep learning algorithms. *Earth Sci. Inform.* **2023**, *16*, 2421–2438. [CrossRef]
63. Zouaidia, K.; Rais, M.S.; Ghanemi, S. Weather forecasting based on hybrid decomposition methods and adaptive deep learning strategy. *Neural Comput. Appl.* **2023**, *35*, 11109–11124. [CrossRef]
64. Pei, M.; Ye, L.; Li, Y.; Luo, Y.; Song, X.; Yu, Y.; Zhao, Y. Short-term regional wind power forecasting based on spatial–temporal correlation and dynamic clustering model. *Energy Rep.* **2022**, *8*, 10786–10802. [CrossRef]
65. Hu, Y.; Guo, Y.; Fu, R. A novel wind speed forecasting combined model using variational mode decomposition, sparse auto-encoder and optimized fuzzy cognitive mapping network. *Energy* **2023**, *278*, 127926. [CrossRef]
66. Zhu, Y. Research on adaptive combined wind speed prediction for each season based on improved gray relational analysis. *Environ. Sci. Pollut. Res.* **2023**, *30*, 12317–12347. [CrossRef]
67. Wu, Y.-K.; Huang, C.-L.; Wu, S.-H.; Hong, J.-S.; Chang, H.-L. Deterministic and probabilistic wind power forecasts by considering various atmospheric models and feature engineering approaches. *IEEE Trans. Ind. Appl.* **2022**, *59*, 192–206. [CrossRef]
68. Zhang, J.; Zhao, Z.; Yan, J.; Cheng, P. Ultra-Short-Term Wind Power Forecasting Based on CGAN-CNN-LSTM Model Supported by Lidar. *Sensors* **2023**, *23*, 4369. [CrossRef]
69. Gong, M.; Yan, C.; Xu, W.; Zhao, Z.; Li, W.; Liu, Y.; Li, S. Short-term wind power forecasting model based on temporal convolutional network and Informer. *Energy* **2023**, *283*, 129171. [CrossRef]
70. Yu, C.; Yan, G.; Yu, C.; Mi, X. Attention mechanism is useful in spatio-temporal wind speed prediction: Evidence from China. *Appl. Soft Comput.* **2023**, *148*, 110864. [CrossRef]
71. Sarangi, S.; Dash, P.K.; Bisoi, R. Probabilistic prediction of wind speed using an integrated deep belief network optimized by a hybrid multi-objective particle swarm algorithm. *Eng. Appl. Artif. Intell.* **2023**, *126*, 107034. [CrossRef]
72. Kosana, V.; Teeparthi, K.; Madasthu, S. Hybrid convolutional BI-LSTM autoencoder framework for short-term wind speed prediction. *Neural Comput. Appl.* **2022**, *34*, 12653–12662. [CrossRef]
73. Ko, M.-S.; Lee, K.; Kim, J.-K.; Hong, C.W.; Dong, Z.Y.; Hur, K. Deep concatenated residual network with bidirectional LSTM for one-hour-ahead wind power forecasting. *IEEE Trans. Sustain. Energy* **2020**, *12*, 1321–1335. [CrossRef]
74. Huang, Q.; Wang, Y.; Yang, X.; Im, S.-K. Research on Wind Power Prediction Based on a Gated Transformer. *Appl. Sci.* **2023**, *13*, 8350. [CrossRef]
75. Pan, X.; Wang, L.; Wang, Z.; Huang, C. Short-term wind speed forecasting based on spatial-temporal graph transformer networks. *Energy* **2022**, *253*, 124095. [CrossRef]
76. Jiang, W.; Liu, B.; Liang, Y.; Gao, H.; Lin, P.; Zhang, D.; Hu, G. Applicability analysis of transformer to wind speed forecasting by a novel deep learning framework with multiple atmospheric variables. *Appl. Energy* **2024**, *353*, 122155. [CrossRef]
77. Zhu, Z.; Xu, Y.; Wu, J.; Liu, Y.; Guo, J.; Zang, H. Wind power probabilistic forecasting based on combined decomposition and deep learning quantile regression. *Front. Energy Res.* **2022**, *10*, 937240. [CrossRef]
78. Song, Y.; Tang, D.; Yu, J.; Yu, Z.; Li, X. Short-Term Forecasting Based on Graph Convolution Networks and Multiresolution Convolution Neural Networks for Wind Power. *IEEE Trans. Ind. Inform.* **2022**, *19*, 1691–1702. [CrossRef]
79. Gao, J.; Ye, X.; Lei, X.; Huang, B.; Wang, X.; Wang, L. A Multichannel-Based CNN and GRU Method for Short-Term Wind Power Prediction. *Electronics* **2023**, *12*, 4479. [CrossRef]
80. Li, Z.; Luo, X.; Liu, M.; Cao, X.; Du, S.; Sun, H. Wind power prediction based on EEMD-Tent-SSA-LS-SVM. *Energy Rep.* **2022**, *8*, 3234–3243. [CrossRef]
81. Fu, X.; Feng, Z.; Yao, X.; Liu, W. A Novel Twin Support Vector Regression Model for Wind Speed Time-Series Interval Prediction. *Energies* **2023**, *16*, 5656. [CrossRef]
82. Yuan, D.-D.; Li, M.; Li, H.-Y.; Lin, C.-J.; Ji, B.-X. Wind power prediction method: Support vector regression optimized by improved jellyfish search algorithm. *Energies* **2022**, *15*, 6404. [CrossRef]
83. Sun, W.; Wang, X. Improved chimpanzee algorithm based on CEEMDAN combination to optimize ELM short-term wind speed prediction. *Environ. Sci. Pollut. Res.* **2023**, *30*, 35115–35126. [CrossRef]
84. Jnr, E.O.-N.; Ziggah, Y.Y.; Rodrigues, M.J.; Relvas, S. A hybrid chaotic-based discrete wavelet transform and Aquila optimisation tuned-artificial neural network approach for wind speed prediction. *Results Eng.* **2022**, *14*, 100399.
85. Ding, L.; Bai, Y.; Liu, M.-D.; Fan, M.-H.; Yang, J. Predicting short wind speed with a hybrid model based on a piecewise error correction method and Elman neural network. *Energy* **2022**, *244*, 122630. [CrossRef]
86. Ponkumar, G.; Jayaprakash, S.; Kanagarathinam, K. Advanced Machine Learning Techniques for Accurate Very-Short-Term Wind Power Forecasting in Wind Energy Systems Using Historical Data Analysis. *Energies* **2023**, *16*, 5459. [CrossRef]
87. Guan, S.; Wang, Y.; Liu, L.; Gao, J.; Xu, Z.; Kan, S. Ultra-short-term wind power prediction method combining financial technology feature engineering and XGBoost algorithm. *Heliyon* **2023**, *9*, e16938. [CrossRef]

88. Liao, S.; Tian, X.; Liu, B.; Liu, T.; Su, H.; Zhou, B. Short-Term Wind Power Prediction Based on LightGBM and Meteorological Reanalysis. *Energies* **2022**, *15*, 6287. [CrossRef]
89. Shirzadi, N.; Nasiri, F.; Menon, R.P.; Monsalvete, P.; Kaifel, A.; Eicker, U. Smart Urban Wind Power Forecasting: Integrating Weibull Distribution, Recurrent Neural Networks, and Numerical Weather Prediction. *Energies* **2023**, *16*, 6208. [CrossRef]
90. Wei, J.; Wu, X.; Yang, T.; Jiao, R. Ultra-short-term forecasting of wind power based on multi-task learning and LSTM. *Int. J. Electr. Power Energy Syst.* **2023**, *149*, 109073. [CrossRef]
91. Liu, Z.; Liu, H. A novel hybrid model based on GA-VMD, sample entropy reconstruction and BiLSTM for wind speed prediction. *Measurement* **2023**, *222*, 113643. [CrossRef]
92. Zhao, Z.; Yun, S.; Jia, L.; Guo, J.; Meng, Y.; He, N.; Li, X.; Shi, J.; Yang, L. Hybrid VMD-CNN-GRU-based model for short-term forecasting of wind power considering spatio-temporal features. *Eng. Appl. Artif. Intell.* **2023**, *121*, 105982. [CrossRef]
93. Hu, H.; Li, Y.; Zhang, X.; Fang, M. A novel hybrid model for short-term prediction of wind speed. *Pattern Recognit.* **2022**, *127*, 108623. [CrossRef]
94. Hua, L.; Zhang, C.; Peng, T.; Ji, C.; Nazir, M.S. Integrated framework of extreme learning machine (ELM) based on improved atom search optimization for short-term wind speed prediction. *Energy Convers. Manag.* **2022**, *252*, 115102. [CrossRef]
95. Zhang, Y.; Chen, Y. Application of hybrid model based on CEEMDAN, SVD, PSO to wind energy prediction. *Environ. Sci. Pollut. Res.* **2022**, *29*, 22661–22674. [CrossRef] [PubMed]
96. Ding, Y.; Chen, Z.; Zhang, H.; Wang, X.; Guo, Y. A short-term wind power prediction model based on CEEMD and WOA-KELM. *Renew. Energy* **2022**, *189*, 188–198. [CrossRef]
97. Shang, Y.; Miao, L.; Shan, Y.; Gnyawali, K.R.; Zhang, J.; Kattel, G. A Hybrid Ultra-short-term and Short-term Wind Speed Forecasting Method based on CEEMDAN and GA-BPNN. *Weather. Forecast.* **2022**, *37*, 415–428. [CrossRef]
98. Wang, H.; Wang, J. Short-term wind speed prediction based on feature extraction with Multi-task Lasso and Multilayer Perceptron. *Energy Rep.* **2022**, *8*, 191–199. [CrossRef]
99. 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*, 905155. [CrossRef]
100. Quan, H.; Zhang, W.; Zhang, W.; Li, Z.; Zhou, T. An Interval Prediction Approach of Wind Power Based on Skip-GRU and Block-Bootstrap Techniques. *IEEE Trans. Ind. Appl.* **2023**, *59*, 4710–4719. [CrossRef]
101. Jiang, T.; Liu, Y. A short-term wind power prediction approach based on ensemble empirical mode decomposition and improved long short-term memory. *Comput. Electr. Eng.* **2023**, *110*, 108830. [CrossRef]
102. Wu, H.; Meng, K.; Fan, D.; Zhang, Z.; Liu, Q. Multistep short-term wind speed forecasting using transformer. *Energy* **2022**, *261*, 125231. [CrossRef]
103. Sun, F.; Jin, T. A hybrid approach to multi-step, short-term wind speed forecasting using correlated features. *Renew. Energy* **2022**, *186*, 742–754. [CrossRef]
104. Chen, J.; Liu, H.; Chen, C.; Duan, Z. Wind speed forecasting using multi-scale feature adaptive extraction ensemble model with error regression correction. *Expert Syst. Appl.* **2022**, *207*, 117358. [CrossRef]
105. Xiao, Z.; Tang, F.; Wang, M. Wind Power Short-Term Forecasting Method Based on LSTM and Multiple Error Correction. *Sustainability* **2023**, *15*, 3798. [CrossRef]
106. Zhang, Y.; Kong, X.; Wang, J.; Wang, S.; Zhao, Z.; Wang, F. A comprehensive wind speed prediction system based on intelligent optimized deep neural network and error analysis. *Eng. Appl. Artif. Intell.* **2024**, *128*, 107479. [CrossRef]
107. Xing, F.; Song, X.; Wang, Y.; Qin, C. A New Combined Prediction Model for Ultra-Short-Term Wind Power Based on Variational Mode Decomposition and Gradient Boosting Regression Tree. *Sustainability* **2023**, *15*, 11026. [CrossRef]
108. Wang, J.; Gao, D.; Zhuang, Z.; Wu, J. An optimized complementary prediction method based on data feature extraction for wind speed forecasting. *Sustain. Energy Technol. Assess.* **2022**, *52*, 102068. [CrossRef]
109. Ai, X.; Li, S.; Xu, H. Wind speed prediction model using ensemble empirical mode decomposition, least squares support vector machine and long short-term memory. *Front. Energy Res.* **2023**, *10*, 1043867. [CrossRef]
110. Zhou, G.; Hu, G.; Zhang, D.; Zhang, Y. A novel algorithm system for wind power prediction based on RANSAC data screening and Seq2Seq-Attention-BiGRU model. *Energy* **2023**, *283*, 128986. [CrossRef]
111. Liu, H.; Yang, R.; Wang, T.; Zhang, L. A hybrid neural network model for short-term wind speed forecasting based on decomposition, multi-learner ensemble, and adaptive multiple error corrections. *Renew. Energy* **2021**, *165*, 573–594. [CrossRef]
112. Yin, S.; Liu, H. Wind power prediction based on outlier correction, ensemble reinforcement learning, and residual correction. *Energy* **2022**, *250*, 123857. [CrossRef]
113. Yang, R.; Liu, H.; Nikitas, N.; Duan, Z.; Li, Y.; Li, Y. Short-term wind speed forecasting using deep reinforcement learning with improved multiple error correction approach. *Energy* **2022**, *239*, 122128. [CrossRef]
114. Sarp, A.O.; Mengüç, E.C.; Peker, M.; Güvenç, B.Ç. Data-adaptive censoring for short-term wind speed predictors based on MLP, RNN, and SVM. *IEEE Syst. J.* **2022**, *16*, 3625–3634. [CrossRef]
115. Duan, J.; Zuo, H.; Bai, Y.; Duan, J.; Chang, M.; Chen, B. Short-term wind speed forecasting using recurrent neural networks with error correction. *Energy* **2021**, *217*, 119397. [CrossRef]
116. Yang, Q.; Huang, G.; Li, T.; Xu, Y.; Pan, J. A novel short-term wind speed prediction method based on hybrid statistical-artificial intelligence model with empirical wavelet transform and hyperparameter optimization. *J. Wind. Eng. Ind. Aerodyn.* **2023**, *240*, 105499. [CrossRef]

117. Gao, Y.; Wang, B.; Chen, F.; Zhang, W.; Zhou, D.; Wu, F.; Chen, D. Multi-step wind speed prediction based on LSSVM combined with ESMD and fractional-order beetle swarm optimization. *Energy Rep.* **2023**, *9*, 6114–6134. [CrossRef]
118. Lu, P.; Ye, L.; Tang, Y.; Zhao, Y.; Zhong, W.; Qu, Y.; Zhai, B. Ultra-short-term combined prediction approach based on kernel function switch mechanism. *Renew. Energy* **2021**, *164*, 842–866. [CrossRef]
119. Chen, X.; Li, Y.; Zhang, Y.; Ye, X.; Xiong, X.; Zhang, F. A novel hybrid model based on an improved seagull optimization algorithm for short-term wind speed forecasting. *Processes* **2021**, *9*, 387. [CrossRef]
120. Fu, W.; Zhang, K.; Wang, K.; Wen, B.; Fang, P.; Zou, F. A hybrid approach for multi-step wind speed forecasting based on two-layer decomposition, improved hybrid DE-HHO optimization and KELM. *Renew. Energy* **2021**, *164*, 211–229. [CrossRef]
121. Rayi, V.K.; Mishra, S.; Naik, J.; Dash, P. Adaptive VMD based optimized deep learning mixed kernel ELM autoencoder for single and multistep wind power forecasting. *Energy* **2022**, *244*, 122585. [CrossRef]
122. Zhang, C.; Ma, H.; Hua, L.; Sun, W.; Nazir, M.S.; Peng, T. An evolutionary deep learning model based on TVFEMD, improved sine cosine algorithm, CNN and BiLSTM for wind speed prediction. *Energy* **2022**, *254*, 124250. [CrossRef]
123. Meng, A.; Zhu, Z.; Deng, W.; Ou, Z.; Lin, S.; Wang, C.; Xu, X.; Wang, X.; Yin, H.; Luo, J. A novel wind power prediction approach using multivariate variational mode decomposition and multi-objective crisscross optimization based deep extreme learning machine. *Energy* **2022**, *260*, 124957. [CrossRef]
124. Suo, L.; Peng, T.; Song, S.; Zhang, C.; Wang, Y.; Fu, Y.; Nazir, M.S. Wind speed prediction by a swarm intelligence based deep learning model via signal decomposition and parameter optimization using improved chimp optimization algorithm. *Energy* **2023**, *276*, 127526. [CrossRef]
125. Wei, Y.; Zhang, H.; Dai, J.; Zhu, R.; Qiu, L.; Dong, Y.; Fang, S. Deep Belief Network with Swarm Spider Optimization Method for Renewable Energy Power Forecasting. *Processes* **2023**, *11*, 1001. [CrossRef]
126. Zhu, Q.; Jiang, F.; Li, C. Time-varying interval prediction and decision-making for short-term wind power using convolutional gated recurrent unit and multi-objective elephant clan optimization. *Energy* **2023**, *271*, 127006. [CrossRef]
127. Wu, B.; Wang, L. Two-stage decomposition and temporal fusion transformers for interpretable wind speed forecasting. *Energy* **2024**, *288*, 129728. [CrossRef]
128. Chen, W.; Zhou, H.; Cheng, L.; Xia, M. Prediction of regional wind power generation using a multi-objective optimized deep learning model with temporal pattern attention. *Energy* **2023**, *278*, 127942. [CrossRef]
129. Wang, J.; Zhu, H.; Zhang, Y.; Cheng, F.; Zhou, C. A novel prediction model for wind power based on improved long short-term memory neural network. *Energy* **2023**, *265*, 126283. [CrossRef]
130. Ewees, A.A.; Al-qaness, M.A.; Abualigah, L.; Abd Elaziz, M. HBO-LSTM: Optimized long short term memory with heap-based optimizer for wind power forecasting. *Energy Convers. Manag.* **2022**, *268*, 116022. [CrossRef]
131. Han, Y.; Mi, L.; Shen, L.; Cai, C.; Liu, Y.; Li, K. A short-term wind speed interval prediction method based on WRF simulation and multivariate line regression for deep learning algorithms. *Energy Convers. Manag.* **2022**, *258*, 115540. [CrossRef]
132. Song, D.; Tan, X.; Deng, X.; Yang, J.; Dong, M.; Elkholy, M.; Talaat, M.; Joo, Y.H. Rotor equivalent wind speed prediction based on mechanism analysis and residual correction using Lidar measurements. *Energy Convers. Manag.* **2023**, *292*, 117385. [CrossRef]