iScience

CellPress OPEN ACCESS

Review

An overview of deterministic and probabilistic forecasting methods of wind energy

Yuying Xie,[2,3] Chaoshun Li,[1,2,*] Mengying Li,[3] Fangjie Liu,[2] and Meruyert Taukenova[2]

SUMMARY

In recent years, a variety of wind forecasting models have been developed, prompting necessity to review the abundant methods to gain insights of the state-of-the-art development status. However, existing literature reviews only focus on a subclass of methods, such as multi-objective optimization and machine learning methods while lacking the full particulars of wind forecasting field. Furthermore, the classification of wind forecasting methods is unclear and incomplete, especially considering the rapid development of this field. Therefore, this article aims to provide a systematic review of the existing deterministic and probabilistic wind forecasting methods, from the perspectives of data source, model evaluation framework, technical background, theoretical basis, and model performance. It is expected that this work will provide junior researchers with broad and detailed information on wind forecasting for their future development of more accurate and practical wind forecasting models.

INTRODUCTION

Overview of wind forecasting

In response to the increased global energy demand, intensified energy crisis and accelerated climate change, the share of renewable energy is growing exponentially in global energy market.[1] As a promising resource with abundant reserves and wide distribution, wind energy has attracted lots of attention both in academic research and in industrial applications. However, the inherent weather-dependent instability of wind resource[2] would adversely affect the security and reliability of the power grid.[3–7] Because large-scale electricity storage is still costly, effective wind energy prediction is essential to facilitate wind power grid integration,[8–10] further promoting the development of wind energy applications.[11] So far, numerous models for wind forecasting have been proposed,[12] serving different purpose, as shown in Figure 1. It should be noted that the wind forecasting referred to in this paper includes both wind speed prediction and wind power prediction.

The wind forecasting horizon varies from a few seconds to several days into the future,[13] and they can be grouped into four types, as shown in Table 1.

In terms of predictive spatial scope, the predications can be single turbine prediction (specific wind turbines), single farm prediction (certain wind farms), and clustered prediction (regional wind turbines). Compared with the prediction on individual wind turbines,[14] forecasts for particular wind farms are more prevalent.[15–17] Because of the expansion of installed wind power capacity and increased wind power integration, the utilization of clustered wind forecasting has been continuously promoted.[18,19]

According to the forecasting step size, the predication can be single-step or multi-step. The former predicts the next values x_{n+1} according to the historical time series of n observations $[x_1, x_2, \ldots, x_n]$, whereas the latter forecasts the next m values $[x_{n+1}, x_{n+2}, \ldots, x_{n+m}]$ of n observations.[20] Compared with single-step prediction, multi-step forecasting can provide more meaningful results in a longer period,[21,22] but its performance inevitably declines with the extension of prediction steps. In terms of optimization, wind forecasting can be a single-objective optimization problem (SOP) or a multi-objective optimization problem (MOP),[23] depending on whether these models are evaluated by single or multiple objectives. Earlier MOP was solved by converting multiple functions into a single objective function with constraints.[24] With the development of optimization techniques, a variety of multi-objective optimization algorithms with global optimization have been developed.[23]

[1]School of Civil and Hydraulic Engineering, Huazhong University of Science and Technology, Wuhan 430074, China

[2]China-EU Institute for Clean and Renewable Energy, Huazhong University of Science and Technology, Wuhan 430074, China

[3]Department of Mechanical Engineering and Research Institute for Smart Energy, the Hong Kong Polytechnic University, Kowloon, Hong Kong SAR, China

*Correspondence: csli@hust.edu.cn

https://doi.org/10.1016/j.isci.2022.105804

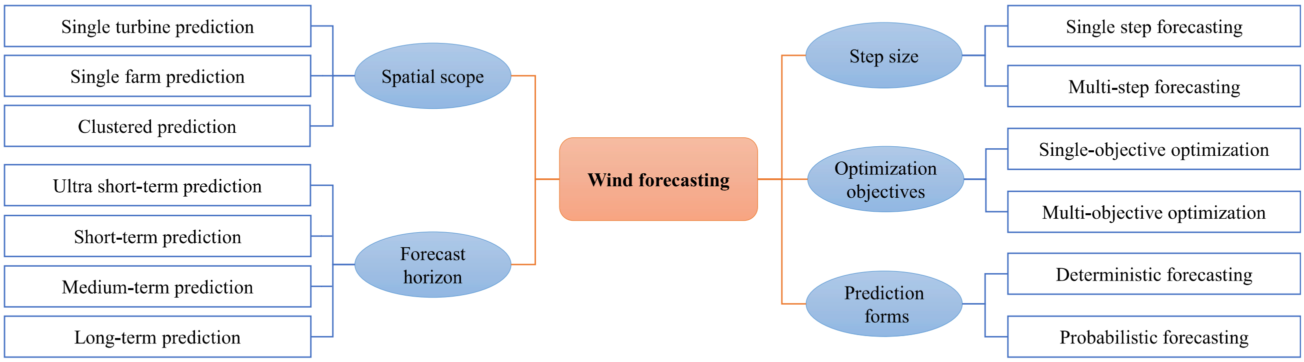

Figure 1. The classification of wind forecasting methods based on different criteria

Although wind forecasting methods can be classified in terms of forecast horizon, spatial scale, step size, and optimization objective, these criteria are sometimes overlapping and thus fail to systematically subdivide the existing prediction models. Therefore, based on the nature of predictive algorithms, the forecasting methods are most widely classified as deterministic and probabilistic forecasting methods,[25,26] which will be comprehensively reviewed in this work. The deterministic forecasting methods are those who output predicted deterministic point values corresponding for varies forecast horizons and spatial scale, whereas probabilistic forecasting method[27–31] can obtain the upper and lower boundaries of wind energy in the form of probability density or probability interval,[32,33] which in turn can provide supplemental information related to wind fluctuations to decision-makers.[34,35]

The objectives and contributions of this work

To summarize and compare different wind forecasting methods, a number of literature reviews exist in the literature. For example, Costa et al. (2008)[36] analyzed the history of short-term wind power prediction. Lei et al. (2009)[37] introduced the forecasts of wind speed and power generation. Al-Yahyai et al. (2010)[38] reviewed the utilization of the Numerical Weather Prediction (NWP) models in wind energy assessment. Soman et al. (2010)[13] reviewed wind forecasting methods in different time frames. Tascikaraoglu et al..(2014)[39] provided a survey on the combined approach of predicting short-term wind speed and power. Yan et al. (2015)[40] emphasized uncertainty analysis in wind power prediction. Marugán et al. (2018)[41] discussed the artificial neural network (ANN) applied to wind energy systems. Liu et al. (2019)[42] sorted out the intelligent predictors and auxiliary methods in the deterministic forecasting field. Tascikaraoglu et al. Liu et al. (2020)[23] discussed the application of multi-objective optimization technology in the field of wind forecasting. Yang et al. (2021)[43] provided a handbook on wind forecasting techniques. In addition, there are also literature reviews on renewable energy prediction, which cover wind forecasting. For instance, Zendehboudia et al. (2018)[44] offered a review from the perspective of the application of support vector machine (SVM) in predicting wind energy. Ren et al. (2015)[45] presented a literature review with a focus on the ensemble methods of wind power prediction. Pérez-Ortiz et al. (2016)[46] reviewed classifications and algorithms in renewable energy applications. Wang et al. (2019)[47] reviewed deep learning techniques applied to renewable energy forecasting. Lai et al. (2020)[48] discussed machine learning models in renewable energy forecasting.

Although the above studies summarized different wind forecasting methods, we found that (1) the majority of the existing reviews have been conducted focusing on only one specific aspect such as deep learning techniques, hybrid methods, or multi-objective optimization models while lacking the full particulars of wind forecasting field; (2) the subcategories of deterministic and probabilistic forecasting are not clearly identified, especially considering the diverse characteristics of various forecast models. For instance, the Numerical Weather Prediction (NWP) model is classified as the only physical model in some literature, whereas it is considered as a subcategory of physical models in other references; (3) with the continuous advancement of wind forecasting research, there is a lack of systematic summary of the existing methods in terms of data resources, forecasting mechanisms, and predictive results. To the best of our knowledge, the above research gaps have not been filled.

Therefore, this work aims to provide a practical reference for classifying, summarizing, and comparing different wind forecasting models. The main contributions of this work are fourfold: (1) A systematic

Table 1. The definition of four forecast horizons of wind forecasting

Forecast Horizon	Definitions
Ultrashort-term prediction	Conduct wind energy prediction only within a few seconds to 30 min ahead
Short-term prediction	Conduct wind energy prediction between 30 min and 6 h ahead
Medium-term prediction	Conduct wind energy prediction from 6 h to 1 day ahead
Long-term prediction	Conduct wind energy prediction which is more than 1 day ahead

classification and review of wind forecasting models has been provided from the perspective of deterministic and probabilistic forecasting; (2) An extensive survey of existing publications detailing the subcategories of wind forecasting methods has been conducted; (3) An in-depth investigation on the data sources, evaluation systems, forecasting models, merits and drawbacks of various methods has been performed; (4) The mainstream and possible development outlooks of wind forecasting methods have been explored, providing future directions for wind forecast-related research.

Research methodology

The wind forecasting publications are selected from Web of Science, ScienceDirect Search, IEEE Xplore, and Google Scholar databases. Both review and research publications are selected from a number of leading journals, such as Applied Energy, Renewable and Sustainable Energy Reviews, IEEE Transactions on Sustainable Energy, Energy Conversion and Management, IEEE Transactions on Power Systems, and Journal of Cleaner Production. With the keywords of ''wind energy prediction'', ''deterministic wind forecasting'', ''probabilistic wind forecasting'' and ''deep wind interval prediction'', over 300 publications are selected and reviewed. Among them, more than 200 are about deterministic wind forecasting whereas around 100 are about probabilistic wind forecasting. The cited references are published from 1998 to 2022 whereas most of them are published in the past decade. A visualization of the keywords from selected literature is shown in Figure 2.

The remainder of this paper is organized as follows. Section 2 provides an overview of data sources for wind prediction models. Section 3 firstly summaries commonly adopted evaluation metrics, and then reviews different categories of deterministic prediction methods. Section 4 presents the evaluation indicators frequently utilized in probabilistic forecasting methods, and then a comprehensive review of parametric and nonparametric probabilistic prediction methods. Section 5 discusses the mainstream and future trends in wind forecasting, and the limitation of this work. Finally, concluding remarks are given in Section 6.

INPUT DATA FOR WIND FORECASTING

High-quality input data plays a crucial role in promoting the development of wind forecasting models because data quality greatly affects model performance and effectiveness. The widely used data sources and processing techniques are summarized as follows.

Measurement of wind speed and wind power

Locally sensed wind speed and wind power data can serve as input and training targets for prediction models. These data is available from public websites such as National Renewable Energy Laboratory (NREL),[49] NASA's Prediction of Worldwide Energy Resource,[50] and National Institute of Wind Energy.[51] Locally sensed data can also be obtained via privately owned wind farms, such as wind fields in Fortaleza and Natal, Brazil,[52] Inner Mongolia, China,[53] etc. Wind data collected on websites are open access, which may facilitate data acquisition for numerical experiments and model validation. However, available data published on websites are relatively limited in terms of data volume and timescale. For example, NREL wind data are currently updated until 2012 and only wind speeds at specific altitudes are reported.[49] In contrast, data from operational wind farms are more abundant, with up-to-date wind data on request by researchers, though this data may not readily available because of confidentiality and security constraints.

NWP data

NWP data can provide forecasts of weather variables such as wind speed, wind direction, ambient temperature and momentum flux, which can be derived from NWP models, including the North American

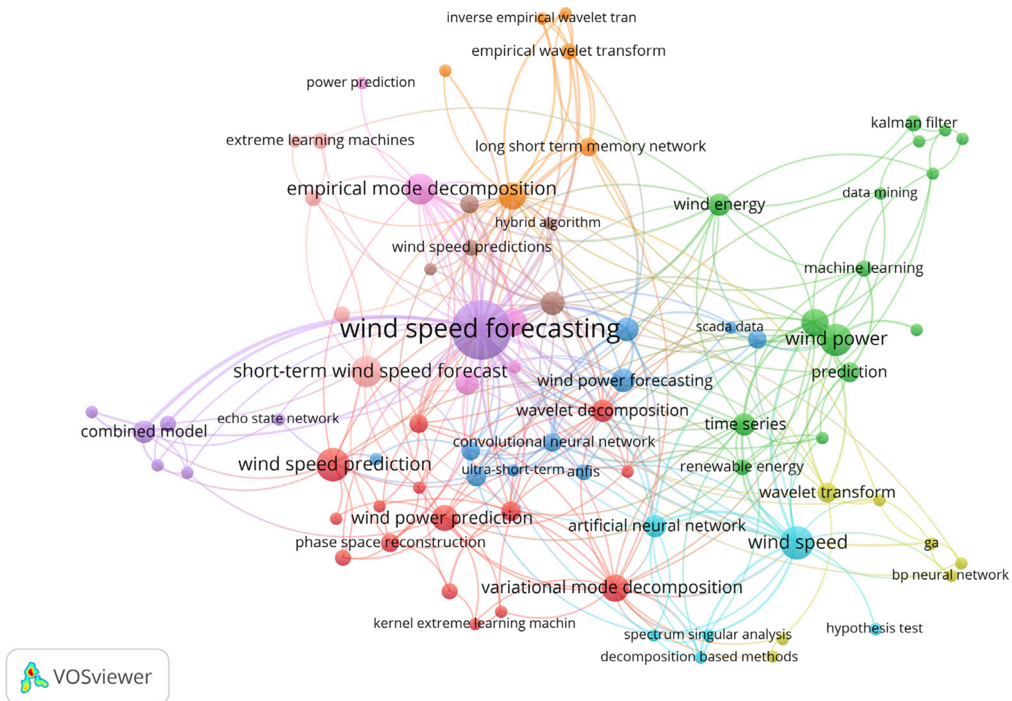

Figure 2. Keywords of the literature that are selected and cited in this work

Mesoscale,[54] the Global Forecast System,[55] and the Weather Research and Forecasting (WRF) system.[55] NWP data are also available on publicly websites, such as Wind Integration National Dataset (WIND).[49] In addition to providing wind forecasts, NWP data can also serve as auxiliary variables for prediction models.[13] For example, Buhan et al.[56] proposed an ANN/SVM model for wind power forecasts 48 h in advance, in which NWP data were used to correlate with measured data.

Exogenous data

Exogenous data can provide prediction models with exogenous variables, including meteorological data, geographic data and wind turbine data. Meteorological data can be collected directly from specific locations using instrumentation, where exogenous variables such as wind direction and air temperature are included. Geographic data, providing information on landforms and surface roughness, are available from the US Geological Survey[57] and the National Aeronautics and Space Administration.[58] Wind turbine data, which provide exogenous variables on wind turbine layout and specifications, can be obtained from publicly accessible sources such as the US Wind Turbine Database.[50] When considering the introduction of exogenous data as inputs for forecasting models, the determination of exogenous variables will depend primarily on their correlation with output variables. For example, Sánchez et al. (2006)[59] developed a statistical wind energy forecasting system, in which the mean hourly wind speed and direction were used as input data. Ye et al. (2017)[60] proposed a spatial correlation model that uses turbine diameter/distance and wind speed as model inputs, whereas other exogenous data (terrain, roughness, and atmospheric disturbance) are discarded.

Data processing techniques

Data processing techniques play a critical role in transforming raw data into smoother sequences,[61] thereby reducing the negative impact caused by incorrect or abnormal data. Commonly adopted data processing techniques include data decomposition, dimensionality deduction, and data denoising, further details of which are given in (Liu et al.).[62]

Data decomposition

The key to the decomposition technique is to decompose the original data into smoother sub-sequences using decomposition algorithms, and then to construct predictive sub-models accordingly.[63] There are

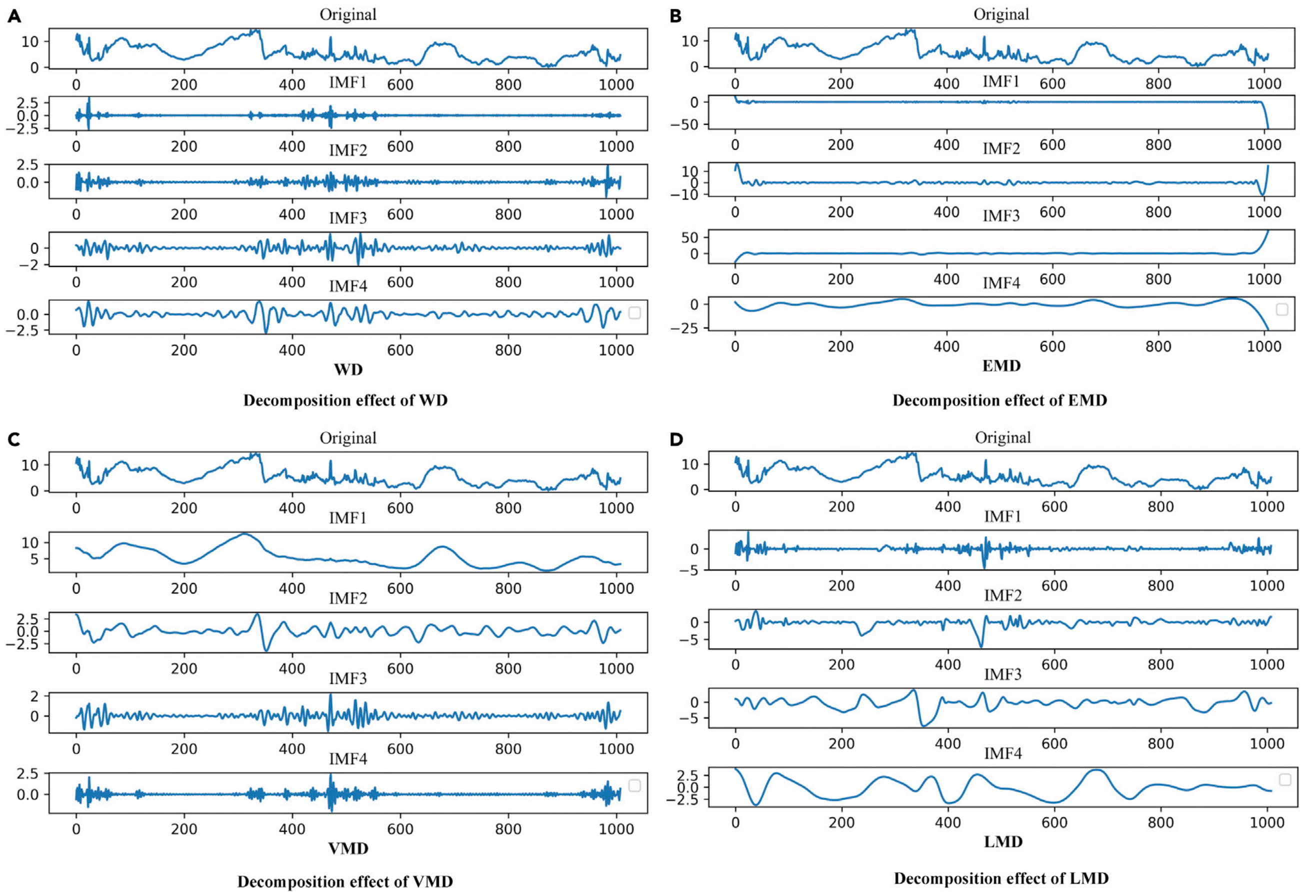

Figure 3. The decomposition effects of different decomposition algorithms

(A) Decomposition effect of WD, (B) Decomposition effect of EMD, (C) Decomposition effect of VMD and (D) Decomposition effect of LMD.

four commonly adopted decomposition algorithms, namely wavelet decomposition (WD), empirical mode decomposition (EMD), variational mode decomposition (VMD), and local mean decomposition (LMD). The decomposition effects of these four algorithms on the same dataset are shown in Figure 3, where the experimental data were retrieved from NREL, which is from an offshore wind field in California (longitude:−124.359, latitude: 41.814) during January 1 to 7, 2012 at 10-min intervals. Figure 3 shows that the decomposed components of all four decomposition algorithms exhibit smaller variation magnitudes when compared to the raw data sequence. However, the decompositions performance of different algorithms differs, especially when comparing the overall trends of all components.

To further improve the performance of data processing technique, a series of variants of these four decomposition algorithms have been derived. Furthermore, the secondary decomposition (SD) technique, which incorporates individual signal processing techniques, has been developed. A summary of various decomposition algorithms is given in Table 2.

Dimensionality deduction

The dimensionality deduction technique is to reduce the dimensionality of the input data, thus reducing the amount of data and the corresponding calculation complexity. There are two effective dimensionality reduction techniques: feature selection and feature extraction. Feature selection refers to the selection of variables from historical input data or decomposed sequences to eliminate redundant information, consisting of correlation analysis, clustering analysis, and phase space reconstruction (PSR). Feature extraction refers to the mapping of features of the original series or generating new features based on the original data. A variety of dimensionality deduction techniques are summarized in Table 3.

Table 2. Summary of data decomposition techniques

Algorithms	Mechanisms and Categories
WD	Decompose the time series into components with different frequency bands, widely adopted techniques include WD,[64,65] continuous WT,[66] discrete WT,[66] empirical wavelet transform (EWT),[67–69] wavelet domain denoising (WDD),[70] wavelet packet decomposition (WPD),[65,71,72] wavelet packet transform,[73] wavelet transform (WT).[74,75]
EMD	Decompose complex time series into finite intrinsic mode function (IMF). There are several algorithms that frequently introduced in the literature, including EMD,[65,76–83] complementary EEMD (CEEMD),[84–88] ensemble EMD (EEMD),[89–92] fast EEMD (FEEMD),[65,93–95] and the improved CEEMD with adaptive noise (ICEEMDAN).[96]
VMD	Decompose non-stationary series into several band-limited intrinsic mode function through VMD.[97–101]
LMD	Decompose the time series into the sum of production function components through LMD.[102]
SD	Combined with signal processing technologies for secondary decomposition, such as EMD-WPD,[103,104] FEEMD-VMD,[105] seasonal information extraction (SIE)-WD,[106] WD-VMD,[107] WPD-CEEMD with adaptive noise (CEEMDAN),[63,108]WPD-FEEMD,[109] VMD-EMD,[110] VMD-WPD.[111]

Data denoising

The data denoising technique is to eliminate noise interference by denoising input data, and there are three denoising techniques involved: singular spectrum analysis (SSA), wavelet threshold denoising (WTD), and decomposition-based denoising. SSA[123–126] removes noise by extracting the trend component from the highest frequency sublayer, whereas WTD[127] extracts noise by setting specific thresholds. The decomposition-based denoising[128–130] is implemented on the basis of decomposition algorithms, as shown in Figure 4. Specifically, original time series data are initially decomposed into sub-series by the decomposition algorithm, and high-frequency noise sequences with higher disorder and non-stationarity are subsequently eliminated by comparing the smoothing of sub-sequences. Finally, noise information that affects the prediction accuracy is removed, forming new inputs to forecasting models.

STATE-OF-THE-ART DETERMINISTIC FORECASTING METHOD

Evaluation metrics

The accuracy and effectiveness of deterministic forecasting models are often evaluated using quantitative indexes. The commonly adopted evaluation metrics are summarized in Table 4, where N denotes the number of samples, y_t denotes the observed wind data, o_t denotes the predicted value of wind speed/power, $\overline{y}$ and $\overline{o}$ denote the average of the actual and predicted data, respectively.

Table 3. The summary of dimensionality deduction techniques

Techniques	Summary		
Feature selection	✓ Correlation analysis refers to the selection of features by analyzing the correlation between features and outputs, and is frequently implemented based on autocorrelation function[112] and partial correlation function (PACF).[89,112]	✓ Clustering analysis refers to the selection of features by establishing clusters based on predictive targets, and commonly adopted cluster algorithms include dilation and erosion clustering,[16] DPK-medoids clustering,[16] fuzzy-c-means,[113] grid partitioning,[113] hierarchical agglomerative clustering,[114] hierarchical cluster (HC),[115] Kernel-based fuzzy c-means clustering,[116] K-means clustering,[114,117] spectral clustering.[114,118]	✓ PSR refers to performing feature selection by reconstructing the vector, and the principal technique involved is PSR.[119,120]
Feature extraction	Several techniques that are frequently used for the extraction of features, namely generalized principal component analysis,[121] Kernel principal component analysis,[119] and principal component analysis (PCA).[118,122]		

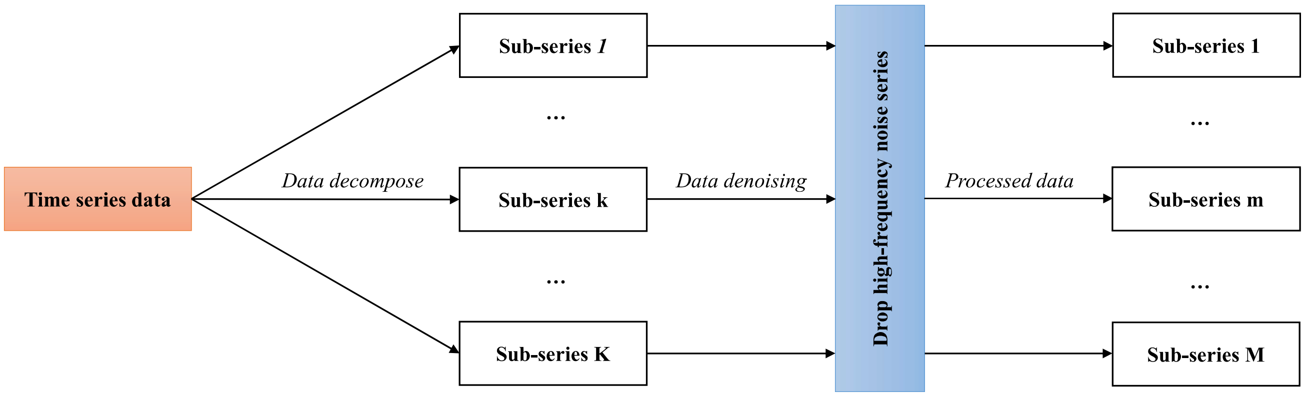

Figure 4. The process of decomposition-based denoising technique

Classification of deterministic forecasting method

The deterministic forecasting method can be divided into three categories: physical model, statistical model, and hybrid model,[137,138] as illustrated in Figure 5.

Physical model

Physical models estimate future wind speed or power generation based on meteorological data and physical information. Forecasts are derived from the mathematical solutions of the transient thermodynamic and hydrodynamic equations.[37,139] Physical models are composed of two subcategories: NWP-based models and spatial correlation models.

NWP-based models use NWP systems to produce wind speed or production power for the next 48 h at a 3-h resolution,[140] where the NWP systems makes predictions by representing topography using conservation equations and digital elevation model.[141] NWP-based models are generally operated every 6 or 12 h,[142] and model output statistics could be introduced to reduce residual errors. For example, Chen et al.[143] established an NWP model based on sampled NWP data provided by the Global Forecast System to produce wind speed, wind direction, temperature, humidity, and pressure forecasts, achieving better performance (MAE improved by approximately 9%–14%) than traditional ANN. More references related to NWP-based models can be found in (Landberg et al., Giebel et al.).[142,144]

The NWP-based model outperforms the time series model when the forecast horizon ranges from 3 to 6 h,[144] whereas its performance is limited when the forecast horizon is from 0 to 2 h.[62] Moreover, because of the complexity of the equations employed and the wide range of predictions, its accuracy is heavily dependent on the completeness of the initial conditions and the supercomputer operation.[145]

Spatial correlation models[59,60,146–148] exploit local and multi-position spatial correlations to predict wind speed and power production on time scales ranging from minutes to hours, and are supported by multi-topography investigations throughout the year. Alexiadis et al. (1999)[146] predicted wind speed and power generation for up to several hours using consequent averages at upwind sites, with a 20%–40% improvement in average error compared to persistence model. Sahin et al. (2000)[147] estimated regional wind speed based on functions transformed by wind speed and elevation changes, with an error less than 20%. Damousis et al. (2004)[149] used nearby wind speed and direction data to forecast wind speed and power generation for the next 30 min to 2 h, with a performance improvement of 10%–25% compared to persistent model. Barbounis et al. (2007)[150] conducted a 7-step wind speed prediction (from 15 min to 3 h) using wind data collected from remote locations, which outperformed other static neural and neuro-fuzzy models. Ye et al. (2017)[60] projected regional wind power using turbine diameter, turbine distance, and wind speed with an nRMSE of 4.73%.

Physical models perform well in medium- and long-term forecasts[151] but poorly in short-term forecasts, and thus have not been widely employed for short-term wind forecasting. It should be noted that in some cases, outputs from physical models may be used as auxiliary inputs to other models.

Table 4. Commonly adopted evaluation indexes for deterministic models (in alphabetical order)

Metrics	Equations	Implication	Features
Correlation coefficient (R)[131]	$R = \frac{\sum_{t=1}^{N}(y_t - \bar{y})(o_t - \bar{o})}{\sqrt{\sum_{t=1}^{N}(y_t - \bar{y})^2 (o_t - \bar{o})^2}}$	R varies between −1 and 1. The closer R is to one, the better the linear correlation between the predicted and actual values.	The correlation coefficient is invariant to scaling and shifting.
Coefficient of determination (R^2)[132]	$R^2 = 1 - \frac{\sum_{t=1}^{N}(y_t - o_t)^2}{\sum_{t=1}^{N}(y_t - \bar{y})^2}$	R^2 varies between zero and infinity. The best value of R^2 is one, which represents the best fit to the data.	The coefficient of determination cannot fully reflect the predictive capability of the model.
Directional change[133]	$DC = \frac{100}{N-1}\sum_{t=1}^{N-1} a_t$ $a_t = \begin{cases} 0, (o_{t+1} - y_t)(y_{t+1} - y_t) \leq 0 \\ 1, (o_{t+1} - y_t)(y_{t+1} - y_t) > 0 \end{cases}$	DC varies between zero and infinity, which can reflect the prediction movement directions or turning points of the forecasting model.	DC samples data points at their peaks and troughs in their movement.
Index of agreement (IA)[133]	$IA = 1 - \frac{\sum_{t=1}^{N}(o_t - y_t)^2}{\sum_{t=1}^{N}(\lvert o_t - \bar{y}\rvert + \lvert y_t - \bar{y}\rvert)^2}$	IA varies between zero and one. The agreement value of 1 indicates a perfect match, and 0 indicates no agreement at all.	IA can detect proportional differences in the observed and simulated means and variances.
Mean absolute error (MAE)[134]	$MAE = \frac{1}{N}\sum_{t=1}^{N}\lvert y_t - o_t\rvert$	MAE varies between zero and infinity, and a smaller MAE denotes higher prediction accuracy.	MAE is more sensitive to extreme values than other metrics.
Mean absolute percentage error (MAPE)[134,135]	$MAPE = \frac{1}{N}\sum_{t=1}^{N}\left\lvert\frac{y_t - o_t}{y_t}\right\rvert \times 100\%$	MAPE varies between 0 and 100%. 0%–10% denotes high accuracy, 10%–20% denotes good performance, 20%–50% denotes a reasonable prediction, 50%–100% denotes a poor result.	MAPE is expressed as a percentage, making the interpretation of results easier than other metrics.
Mean absolute scale error (MASE)[135]	$MASE = \frac{1}{N}\sum_{t=1}^{N}\left\lvert\frac{y_t - o_t}{\frac{1}{N-1}\sum_{t=1}^{N}\lvert y_t - y_{t-1}\rvert}\right\rvert$	MASE varies between zero and infinity, with values below one denoting a better prediction performance.	MASE has the advantage of being robust and stable.
Normalized mean absolute error (nMAE)[131]	$nMAE = \frac{\frac{1}{N}\sum_{t=1}^{N}\lvert y_t - o_t\rvert}{\bar{y}} \times 100\%$	nMAE varies between 0 and 100%. Smaller nMAE denotes higher forecasting accuracy, and nMAE equals to zero indicates a perfect model.	nMAE can provide a superior measure of predictive accuracy, and it avoids scale dependency.
Normalized root-mean-square error (nRMSE)[131]	$nRMSE = \frac{1}{\bar{y}_t}\sqrt{\frac{1}{N}\sum_{t=1}^{N}(y_t - o_t)^2} \times 100\%$	nRMSE varies between 0 and 100%. 0%–10% denotes excellent accuracy, 10%–20% denotes good accuracy, 20%–30% denotes logical accuracy, 30%–100% denotes relatively poor accuracy.	nRMSE is not always reliable for finding the best model, especially for small samples.
Root-mean-square error (RMSE)[55,134]	$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(y_t - o_t)^2}$	RMSE varies between zero and infinity. Smaller values denote higher forecasting accuracy, and values equal to zero indicate perfect models.	RMSE is pretty sensitive to outliers in the data.
Symmetric mean absolute percentage error (SMAPE)[136]	$SMAPE = \frac{1}{N}\sum_{t=1}^{N}\frac{\lvert y_t - o_t\rvert}{(\lvert y_t\rvert + \lvert o_t\rvert)/2} \times 100\%$	SMAPE varies between 0 and 100%. 0% denotes a perfect result, while a higher value denotes an inferior one.	SMAPE is symmetric, which can better avoid the bias caused by small values in real data.

Statistical model

Statistical models regress the relationship between input variables (e.g., past lagged values) and output variables (e.g., future predictions), which are commonly used in short-term wind predictions, including time series analysis-based models, Kalman filter models, and machine learning-based models.[152]

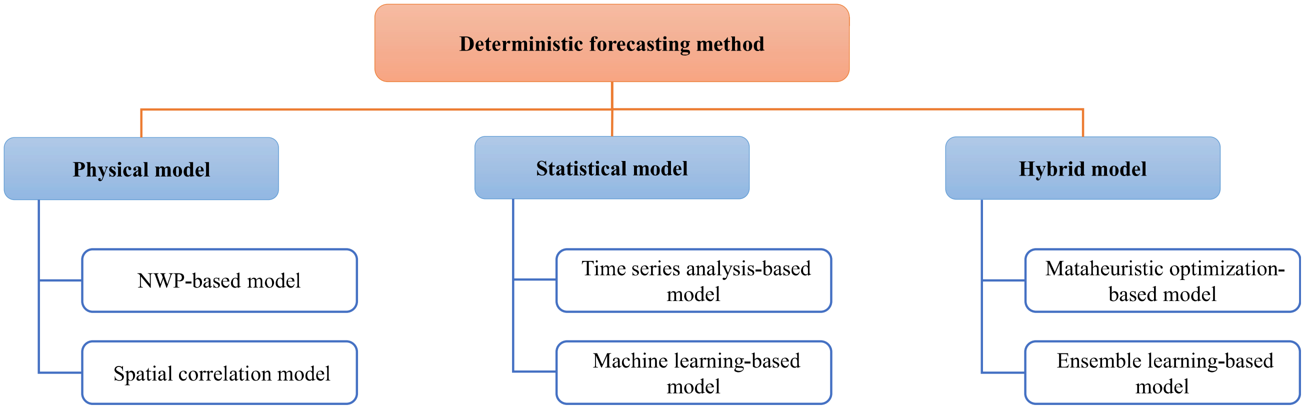

Figure 5. Classification of the deterministic forecasting method

Time series analysis-based model. The time series analysis-based model make predictions according to the autocorrelation of wind data. The predictors frequently used in the literature are autoregressive moving average model (ARMA), autoregressive integrated moving average model (ARIMA), ARIMA with the explanatory variable model (ARIMAX), fractional-ARIMA, seasonal ARIMA (SARIMA), and polynomial autoregressive model. Torres et al. (2005)[153] established an ARMA to achieve 10-h-ahead wind forecasts. Kavasseri et al. (2009)[154] proposed a fractional-ARIMA model to obtain one-day and two-day ahead wind forecasts. Erdem et al. (2011)[155] developed an ARMA to realize 1-h ahead wind speed forecast. Guo et al. (2010)[156] performed wind speed forecasting by introducing a SARIMA model. Shi et al. (2011)[157] proposed an autoregressive ARIMA based on direct and indirect approaches to predict power generation. Lydia et al. (2016)[158] predicted 1-h-ahead wind speed using exogenous variables (wind direction, wind shear, and temperature) based on an ARMA model. Karakuş et al. (2017)[159] established a PAR to forecast the wind speed/power for the next day, with nMAPE improved by about 4%–10%. In recent studies, time series analysis-based models have been used in combination with other models. For example, Camelo et al.[52] combined the ARIMAX with a machine learning-based model for monthly and hourly wind speed forecasting.

Kalman filter model. The Kalman filter (KF) model estimates the state of the process by minimizing errors while recursively combining observations with the latest projections, and thus approximate unknown variables contaminated by Gaussian noise.[160] Liu et al. (2012)[161] proposed a KF model initialized by ARIMA to predict wind speed. Poncela et al. (2013)[162] developed KF models using hourly historical power and NWP data for wind power prediction from 0 to 30 h. Chen et al. (2014)[163] conducted wind speed predictions based on the estimated states derived using KF. Zuluaga et al. (2015)[164] applied three Kalman filters to predict short-term wind speed. KF can also be adopted for data assimilation, especially for improving weather forecasts based on the NWP model. Crochet et al. (2004)[165] developed a KF procedure to improve the accuracy of 10-meter wind speed forecasts from the NWP model. Louka et al. (2008)[166] introduced KF as a post-processing method to remove systematic errors in NWP forecasts. Cassola et al. (2012)[167] employed KF to decrease the inherent flaws in NWP models with prediction horizons of 6, 12, 18, 24, and 36 h. Williams et al. (2013)[168] adapted KF to promote the observation capability of the NWP system. Stathopoulos et al. (2013)[169] utilized KF to reduce potential biases in regional atmospheric systems.

Machine learning-based model. Machine learning-based model makes predictions by learning the intrinsic relationships between input and output data rather than using mathematical methods for acquiring statistical information,[170]which can be further classified into shallow predictors-based method and deep learning-based predictors method.[42]

The shallow predictors-based method is established based on shallow neural networks, and there are four representative predictors involved, including ANN, extreme learning machine (ELM), SVM, and fuzzy logic model, as illustrated in Figure 6. For example, Cadenas et al. (2009)[171] and Fadare et al. (2010)[172] applied ANN to predict short-term wind speed with hourly data. Wang et al. (2018)[96] used the ELM model to predict wind speed with a 10-min prediction horizon. Li et al. (2021)[173] achieved short-term wind power prediction

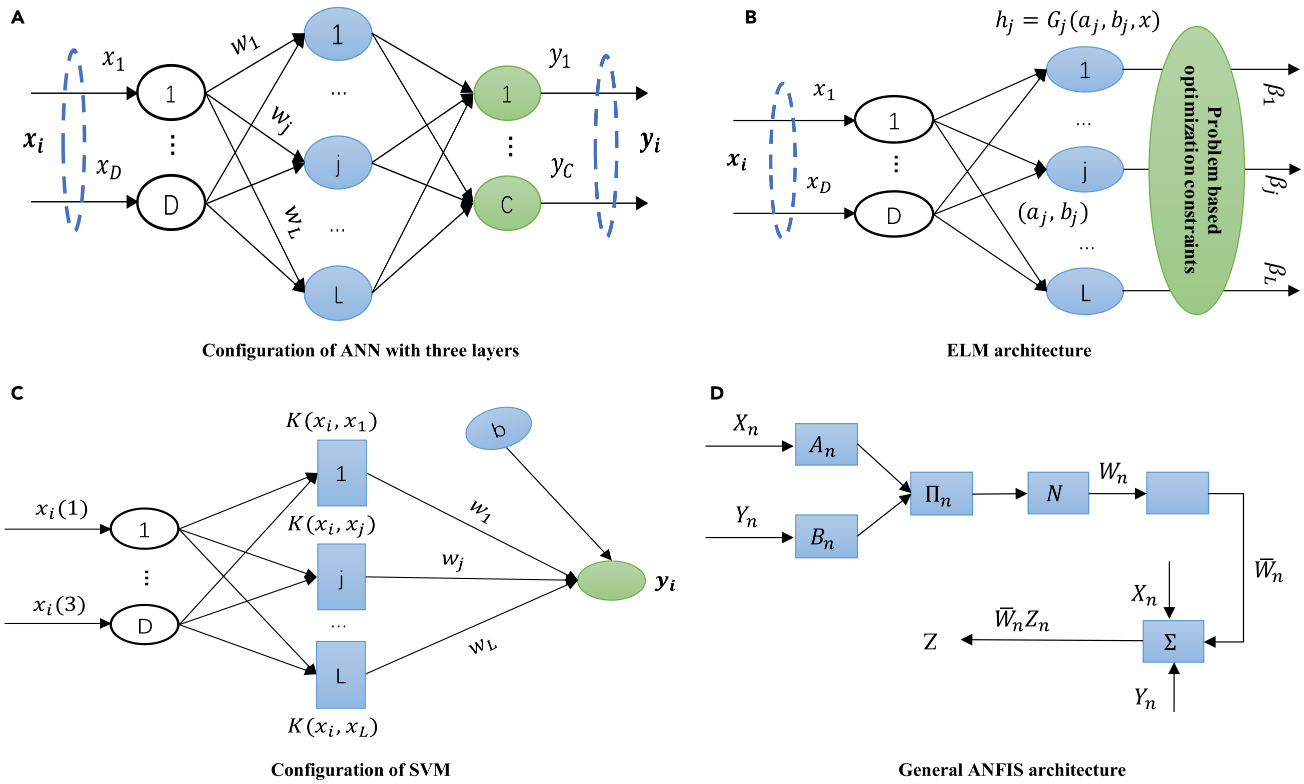

Figure 6. The schematic configurations of four representative shallow neural networks

(A) Configuration of ANN with three layers, (B) ELM architecture, (C) Configuration of SVM and (D) General ANFIS architecture.

using the ELM model. Hu et al. (2022)[174] proposed SVM-based wind speed prediction models based on daily wind speed data. Fuzzy logical models are developed through the combination of fuzzy logic and neural networks. Monfared et al. (2009)[175] applied fuzzy logic to ANN for the prediction of wind speed. Mohandes et al. (2011)[176] and Osório et al. (2015)[177] adopted an adaptive neuro-fuzzy inference system (ANFIS) for wind speed and wind power prediction. Ma et al. (2017)[178] implemented wind speed prediction using the generalized dynamic fuzzy neural network.

Deep learning-based predictors method introduces deep learning techniques with powerful learning capabilities.[179] The predictors including autoencoder (AE), restricted Boltzmann machine (RBM), deep neural network (DNN), convolutional neural networks (CNN), and transfer learning (TL) have been widely adopted, as shown in Figure 7. Owing to the superior performance of deep predictors over shallow predictors, the application of deep learning-based predictors method for wind forecasting has been greatly promoted. Dalto et al. (2015)[180] conducted ultra-short-term wind speed prediction using measured wind data with a temporal resolution of 10 min. Liu et al. (2015)[181] developed a CNN-based model for ultra-short-term wind power forecasts. Santhosh et al. (2019),[182] Khodayar et al. (2019),[183] and He et al. (2022)[184] applied DBN for short-term, 1–24 h ahead, and ultra-short-term wind speed predictions, respectively. Bourakadi et al. (2022)[185] proposed an AE model to predict short-term wind power using wind speed and power data at 1-h intervals. It is worth mentioning that CNNs are often employed as data pre-processing tools for other models, owing to their excellent feature extraction capability. For example, Ji et al. (2022)[186] established a CNN-gate recurrent unit (GRU) model for wind speed prediction, where CNN was used to extract characteristic input vectors.

TL is a novel machine learning approach for tackling issues in domains with distinct but relevant tasks,[186] thus enabling wind prediction based on forecasts of nearby locations. Machine learning techniques (especially DNNs) are generally adopted as predictors. For example, Hu et al. (2016)[187] and Liu et al. (2021)[188] developed DNN-based models through transfer learning to provide short-term predictions on wind speed and wind power.

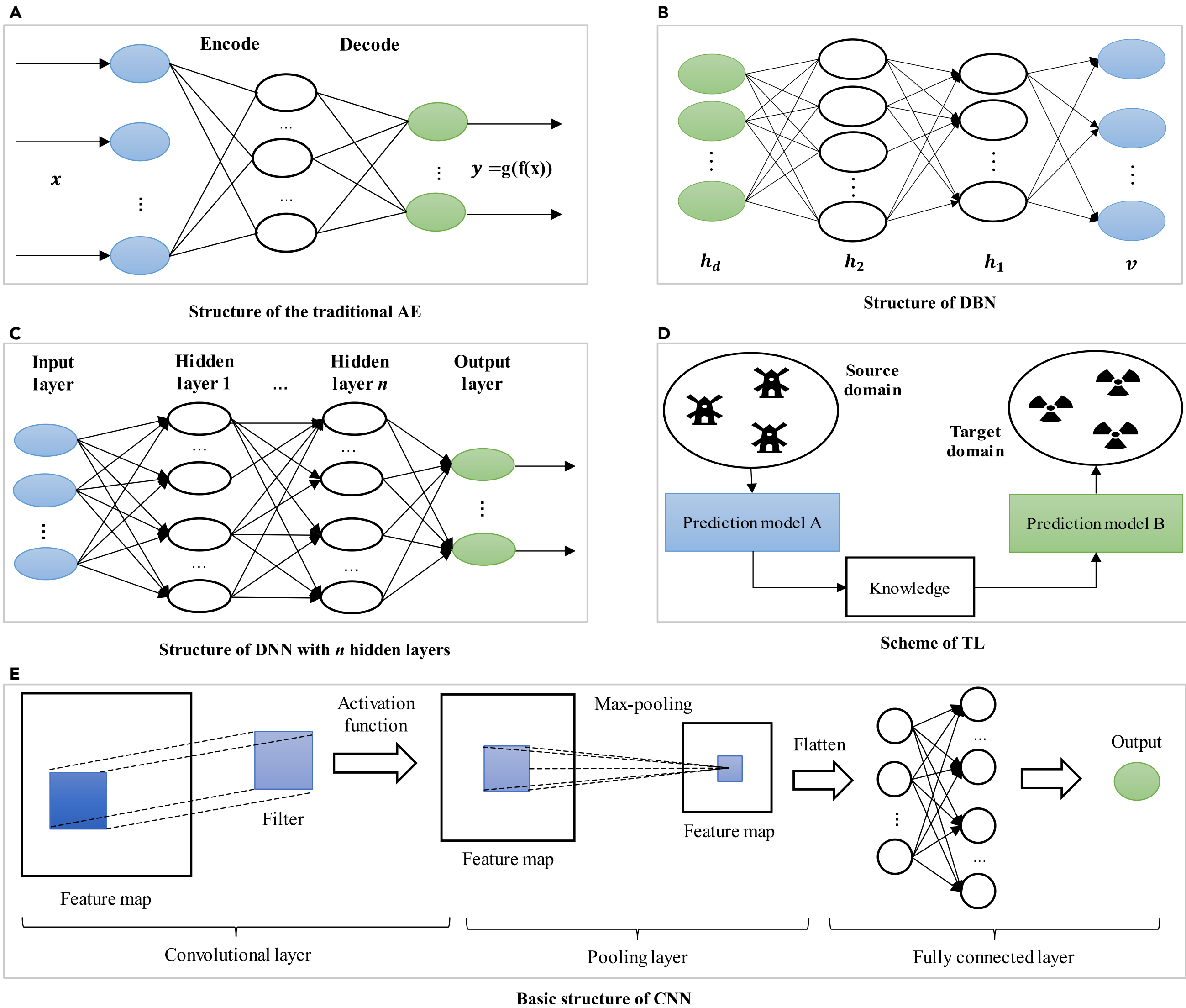

Figure 7. The structures and schemes of five categories of deep learning technologies

(A) Structure of the traditional AE, (B) Structure of DBN, (C) Structure of DNN with *n* hidden layers, (D) Scheme of TL and (E) Basic structure of CNN.

Dedicated to improving performance of machine learning-based models, conventional predictors such as ANN and DNN have been improved by improving the internal structure of predictors or by integrating other effective mechanisms. More details of machine learning-based models are summarized in Table 5.

Hybrid model

Hybrid models perform predictions by combining the advantages of multiple prediction models, which can be divided into metaheuristic optimization-based model and ensemble learning-based model.

Metaheuristic optimization-based model. The metaheuristic optimization-based model is to introduce metaheuristic algorithms for model construction. Earlier models were mainly based on traditional metaheuristic optimization algorithms, such as Particle Swarm Optimization (PSO), whose performance can be improved by two effective means. One is to apply new metaheuristic algorithms to the wind forecasting field, which can be achieved by introducing metaheuristic algorithms applied in other domains or by designing novel optimization algorithms. Another is to upgrade existing algorithms to enhance the overall performance, which can be achieved by using operators from other algorithms, adding mechanisms from other domains, or modifying optimization procedures of the existing algorithms.

Table 5. Variants of the conventional shallow and deep learning-based prediction methods

Predictors	Variants and references
Shallow predictors-based method	
ANN	Adaptive linear element,[189] back propagation neural network (BPNN),[189,190] Elman neural network (ENN),[133] echo state network (ESN),[191] generalized regression neural network (GRNN),[192] multi-layer perceptron (MLP),[193,194] wavelet neural network (WNN),[125] radial basis function neural network (RBFNN)[189,192]
ELM	Hysteretic ELM (HELM),[53] kernel ELM (KELM),[195] online sequential outlier robust ELM,[111] outlier robust ELM,[196] regularized ELM (RELM)[197]
SVM	Core vector machine,[119] ε-SVM,[198] evolutionary SVM,[199] Least square SVM (LSSVM),[73] reduced support vector machine,[122] v-SVM[198]
Deep learning-based predictors method	
AE	Rough stacked autoencoder,[200] rough stacked denoising autoencoder,[200] stacked denoising autoencoder,[201] staked independent recurrent autoencoder,[202] stacked sparse autoencoder[203]
CNN	Convolutional LSTM (ConvLSTM),[204] convolutional SVM[205]
DNN	Bidirectional GRU,[206] gate recurrent unit (GRU),[207,208] long short-term memory (LSTM),[86,209–212] minimal gated memory,[213] recurrent neural network (RNN),[214] stacked ELM (SELM),[215] stacked RNN[208]

Adopting metaheuristic optimization algorithms have two essential functions: model parameter optimization and coefficients determination. From the perspective of optimizing parameters within the models, most algorithms are adopted to update the weights and biases of given prediction models. In some cases, these algorithms can be used to determine internal parameters of the model, such as the number of layers or the size of neurons of neural networks. Determining the optimal coefficients can be further classified as two types of functions: those introduce decomposition algorithms for data processing, and those establish predictive systems by combining individual models. The former introduces optimization algorithms to search for optimal coefficients to superpose forecasting results of each decomposed component, whereas the latter uses optimization algorithms to determine the best solution for combining individual models. Moreover, some models are developed by introducing both decomposition algorithms and combined prediction systems, where metaheuristic optimization algorithms are utilized to determine the optimal prediction model for each sub-series. A summary of relevant metaheuristic optimization-based models is presented in Table 6.

Ensemble learning-based model. Ensemble learning-based models are combinations of different predictive models with various existing technologies (e.g., data processing and optimization strategies), which generally outperform traditional forecasting models. Table 7 presents representative ensemble models and an exhaustive comparison among them. The model improvement range ($I_{min} \sim I_{max}$) is calculated using the evaluation metric (e.g., MAE, MAPE, NMAE) of the proposed ensemble model, and the minimal and maximal metric values of the benchmark models.[19,252,253]

$$I_{min} = \frac{V_{b,min} - V_p}{V_{b,min}} \quad \text{(Equation 1)}$$

$$I_{max} = \frac{V_{b,max} - V_p}{V_{b,max}} \quad \text{(Equation 2)}$$

where $V_{b,min}$ and $V_{b,max}$ denote the minimal and maximal metric values of the benchmark models, V_p denotes the metric value of the proposed ensemble model.

Comparison of deterministic forecasting models

The comparison of different subcategories of deterministic forecasting methods is summarized in Table 8.

STATE-OF-THE-ART PROBABILISTIC FORECASTING METHOD

Evaluation metrics for probabilistic models

A variety of evaluation metrics have been proposed to evaluate the performances of probabilistic forecasting models based on distribution, quantile, or intervals.

Table 6. The summary of the metaheuristic optimization-based models (in alphabetical order)

Algorithms	Mechanisms	Reference models	Optimization variables
Artificial bee colony (ABC)	Deriving from the honey harvesting mechanism of bees, individuals are updated according to the greedy strategy to identify individuals to be abandoned.	RVM-ABC[216] BPNN-Improved ABC (IABC)[217]	Kernel parameters of RVM models. Weight and threshold values of BPNN.
Atomic search algorithm	Deriving from the atoms displacement in a molecular system, individuals are updated by sequentially updating acceleration, velocity, and position.	SVM-Improved atomic search algorithm[218]	Parameters of SVM.
Bald eagle search	Deriving from the intelligent social behavior of bald eagle fishing, individuals are updated by dividing the movement into three phases.	VMD-Combined models-Multi-objective bald eagle search[219]	Optimal weights for sub-models, which is composed of ENN, ARIMA, LSSVM, GMDH, ELM, BPNN, and LSTM.
Bat algorithm	Deriving from the echolocation of bats, individuals are updated based on the speed, position, and loudness.	GRNN-Improved bat algorithm[192]	Weights between layers and thresholds of the hidden layer of neural networks.
Charged system search (CSS)[220]	Deriving from the electrostatics law in Coulomb and Gaussian physics, individuals are updated based on fitness value and separation distance.	Hybrid model-CSS[220]	Weight coefficients of individual models, including ANN, ANFIS, and LSSVM.
Clonal selection algorithm	Deriving from the cell proliferation and differentiation after contact with the antigen, individuals are updated by adjusting affinity.	WNN-Improved clonal selection algorithm[221]	Free parameters of WNN.
Coral reef optimization	Deriving from the colony behavior of coral polyps, individuals are updated by modeling and simulating different processes.	ELM-coral reef optimization[222,223]	Optimal set of meteorological variables from the WRF.
Cuckoo search	Deriving from the brood parasitism behavior of certain cuckoo species, individuals are updated based on random updates.	GMDH-Multi-objective cuckoo search optimization[224]	Partial description coefficients of GMDH networks.
Differential evolution	Proposing for solving real number optimization problem, individuals are updated based on mutual cooperation and competition.	SVR-Hybrid improved cuckoo search[225]	Parameters of penalty function and kernel function in SVR.
Dragonfly algorithm	Deriving from the behavior of dragonflies in search of prey, individuals are updated by constantly changing the position vector X and direction vector ΔX.	Immune selection multi-objective optimization dragonfly algorithm[226] SVM-Improved dragonfly algorithm[25]	Weights of four components, which is composed of ARIMA, BPNN, ENN, and ELM. Parameters of SVM.
Extremal optimization (EO)	Deriving from the self-organizing key evolution model, individuals are updated based on species with minimal fitness values.	EnsemLSTM-EO[227]	Parameters of top-layer SVRM.
Firefly algorithm	Deriving from the flickering behavior of fireflies, individuals are updated by moving particle to better position ones.	MLP-firefly algorithm[228]	Parameters of MLP.

(*Continued on next page*)

Table 6. *Continued*

Algorithms	Mechanisms	Reference models	Optimization variables
Flower-pollination algorithm	Deriving from the pollination process of flowering plants, individuals are updated based on the conversion probability.	CEEMDAN-Combined model-flower-pollination algorithm[61]	Optimal weight coefficients of the combined model based on BPNN, RBFNN, GRNN, WNN, and ENN.
Genetic algorithm (GA)	Deriving from the biological evolution process of natural selection and genetics mechanism, individuals are updated by repeatedly modifying the group composed of individual solutions.	EEMD-GA-LSTM[229]	Appropriate IMFs as the features set for LSTM training.
Gray wolf optimization (GWO)	Deriving from the predation activities of gray wolves, individuals are updated by allowing individuals to move to the best three individuals in the group.	VMD-GWO-ELM[230]	Parameters of PSR and the number of hidden neurons in ELM for each component.
		CEEMDAN-Multi-objective GWO-KNEA[231]	Parameters of KNEA.
		KELM-Improved GWO (IGWO)[232]	Parameters of KELM.
		Improved multi-objective GWO[233]	Parameter of ANN models (SVM, GRNN, BPNN, ANFIS, LSTM, and NARNN).
Harris hawks optimization (HHO)	Deriving from the collaborative behavior and chasing style of Harris hawks, individuals are updated based on search phrases, search transitions, and developmental stages.	PSR-KELM-HHO[234]	Parameters of PSR and KELM.
		KELM-Improved HHO[235]	Parameters of KELM.
Imperialist competitive algorithm	Deriving from the imperialist colonial competition mechanism, individuals are updated through assimilation and mutation.	ANN-Imperialist competitive algorithm[74]	Parameters of ANN.
Marine predator algorithm	Deriving from the natural behaviors of prey and predator to solve complex optimization and engineering problems, individuals are updated based on stochastic strategy on three stages.	ANFIS-Improved MPA[236]	Optimal weights' values between Layer 4 and Layer 5 of ANFIS.
Mayfly Algorithm (MA)	Deriving from the group behavior of mayflies, individuals are updated by randomly scattering in d-dimensional space.	VMD-combined models-Multi-objective MA[237]	Optimal weight coefficients for integrating the forecasting values of the sub-series.
Multi-verse optimization	Deriving from the parallel universes, individuals are updated through the detection of white holes and black holes, and the mining of wormholes.	Combined models-Multi-objective multi-verse optimization[238]	Combination strategy of member models, consisting of BPNN, ELM, and biLSTM.
PSO	Deriving from the way birds and other groups cooperate to search for food, individuals are updated based on the experiences of particles.	ANN/NNS-PSO[239]	State vectors and internal parameters of ANN/NNS.
		ANFIS-PSO[240]	Parameters of membership functions in ANFIS.
		LSSVM-PSO[241]	Parameter of LSSVM.
		ESN-multi-objective PSO[242]	Parameter of ESN.
Salp swarm optimizer (SSO)	Deriving from the foraging activities of bee colonies in the ocean, individuals are updated by the leader leading the crowd and the crowd following the leader.	EEMD-Multi-objective SSO (MOSSO)-Combined system[243]	Optimal weight of individual forecasting systems, namely BP, RVFL, ENN, and GRNN.

(*Continued on next page*)

Table 6. *Continued*

Algorithms	Mechanisms	Reference models	Optimization variables
Tabu search algorithm	Deriving from the merit-seeking features of humans with memory functions, individuals are updated by comparing a series of search directions.	BPNN-Tabu search algorithm[244]	Parameters of BPNN.
Whale optimization algorithm (WOA)	Deriving from the behavior of whales surrounding their prey, individuals are updated by selecting the optimal individual or randomly choosing individuals.	Elman NN-MOWOA[133] MLP-WOA[245]	Weights and thresholds of Elman NN. Parameters of MLP.

Distribution evaluation

Skewness and kurtosis[254] are two metrics used in parametric methods to evaluate patterns of distributions. Specifically, skewness (*s*) is used to measure distribution symmetry, whereas kurtosis (*k*) evaluates the height and sharpness of the central peak in the distribution:

$$s = E(\varepsilon^3) / \sigma^3 \quad \text{(Equation 3)}$$

$$k = E(\varepsilon^4) / \sigma^4 \quad \text{(Equation 4)}$$

where σ and ε denote the SD and prediction error, and E denotes the expectation operator.

Quantile evaluation

Average proportion deviation (PD) and scoring rule (S)[255] are two important indicators in nonparametric methods, which can evaluate the reliability and sharpness of predicted quantiles. An average proportion closer to zero signifies more reliable results whereas higher scoring values signify superior performance. The expressions of average proportion deviation and scoring rule are as follows.

$$PD_Q^\alpha = \frac{1}{T}\sum\nolimits_{t=1}^{T} \eta_t - \alpha \quad \text{(Equation 5)}$$

$$S_{Q,t} = \sum\nolimits_{i=1}^{r} \left(1\left\{y_t \le \hat{q}_t^{\alpha_i}\right\} - \alpha_i\right)\left(y_t - \hat{q}_t^{\alpha_i}\right) \quad \text{(Equation 6)}$$

$$S_Q = \frac{1}{T}\sum\nolimits_{t=1}^{T} S_{Q,t} \quad \text{(Equation 7)}$$

where T denotes the number of forecasting samples, η_t denotes the quantile indicator, α denotes the nominal proportion, $\hat{q}_t^\alpha$ denotes the estimated quantile.

Interval evaluation

Interval results from probabilistic forecasting models are primarily evaluated according to five metrics, including prediction interval coverage probability (PICP), prediction interval normalized averaged width (PINAW)/prediction interval normalized root-mean-square width (PINRW), average coverage error (ACE), coverage width criterion (CWC), and interval normalized average deviation (INAD).

PICP[256] and PINRW/PINAW[257–260] are two types of fundamental evaluation indexes, where PICP evaluates the total probability of the actual value falling within the interval, whereas PINAW/PINRW evaluates the width of the interval. When PINAW/PINRW reaches high levels, the PICP can be very large (perhaps even 100%), which reduces the reliability of the predictive results. However, when the PINAW/PINRW is too small, the resulting PICP may not match the expected prediction interval nominal confidence (PINC). In this case, the discrepancy between PICP and PINC can be evaluated using average coverage error (ACE)[261] index. The formulas for PICP, PINAW/PINRW, and ACE are,

$$\text{PICP} = \left(\frac{1}{n}\sum_{i=1}^{n} c_i\right) \times 100\% \quad \text{(Equation 8)}$$

Table 7. Summary of representative ensemble learning-based models (in alphabetical order)

Model	Data source	Input	Output	Interval	Step	Benchmark	Improvement
ARIMA-ANN[246]	Isla deCedros, Cerro de la Virgen, Holbox	Wind speed	Wind speed	1 h	Singlestep	ARIMA, ANN	0.5057–1.2555 (MAE)
ARIMA-ANN[52]	Fortaleza and Natal, Brazil	Wind speed, pressure, temperature, precipitation	Wind speed	1 month, 1 h	Singlestep	ARIMA, ARIMAX, HW, ANN	1 month: −0.2%–2.76%, 1 h: 6.74%–18.81% (MAPE)
ARIMAX-ANN[52]							1 month: 0.45%–2.26%, 1 h: 7.58%–20.15% (MAPE)
CEEMDAN-SSA-ENN[247]	Sichuan, China	Wind speed	Wind speed	1 h	3-step-ahead	Persistence, ENN, ARIMA, EMD-RAIMA	1-step: −1.32%–51.82%, 2-step: 5.92%–85.27%, 3-step: 11.26%–99.63% (MAPE)
CNN-GRU[248]	Inner Mongolia, China	Wind speed	Wind speed	15 min	3-step-ahead	MLP, random forest regression, SVR, causal convolutional network, GRU	1-step: 0.22%–3.83%, 2-step: 0.17%–4.24%, 3-step: 0.22%–4.79% (MAPE)
EEMD-PACF-LSTM-GPR[89]	Zhangjiakou, China	Wind speed	Wind speed	5 min	Singlestep	ARIMA, BPNN, LSTM, GPR, EEMD-LSTM, EEMD-GPR	0.19%–15.55% (MAPE)
EEMD-SSA-ENN[247]	Sichuan, China	Wind speed	Wind speed	1 h	3-step-ahead	Persistence, ENN, ARIMA, EMD-RAIMA	1-step: 1.69%–54.83%, 2-step: 5.61%–85.77%, 3-step: 9.67%–84.25% (MAPE)
ELM-ICEEMDAN-ARIMA[96]	Zhangye regions, China	Wind speed	Wind speed	10 min	Singlestep	ELM-EMD-ARIMA, ELM-EEMD-ARIMA, ELM-CEEMDAN-ARIMA	0.3746%–4.0833% (MAPE)
EMD-SSA-ENN[247]	Sichuan, China	Wind speed	Wind speed	1 h	3-step-ahead	Persistence, ENN, ARIMA, EMD-RAIMA	1-step: −1.77%–51.37%, 2-step: 3.17%–83.04%, 3-step: 5.40%–87.92% (MAPE)
ESN-PSR[249]	Jiangsu, China	Wind power	Wind power	10 min	3-step-ahead	ANN, add-weighted one-rank local-region, gray model, PSO-based nonlinear gray Bernoulli model	1-step: 0.02%–1.63%, 2-step: 0.03%–1.67%, 3-step: −0.08%–2.06% (NMAE)
EWT-LSTM-ENN[68]	China	Wind speed	Wind speed	1 h	3-step-ahead	ARIMA, BPNN, GRNN, LSTM, ENN, EWT-ENN, WPD-LSTM-ENN, EMD-LSTM-ENN	1-step: 0.35%–6.09%, 2-step: 1.10%–7.22%, 3-step: 0.34%–7.21% (MAPE)
HC-VMD-GA-BPNN[115]	Changma, China	Wind speed	Wind speed	10 min	Singlestep	BPNN, RBF, Elman, GA-BP, HC-GA-BP	5.92%–12.70% (MAPE)

(Continued on next page)

Table 7. ***Continued***

Model	Data source	Input	Output	Interval	Step	Benchmark	Improvement
Holt-Winters (HW)-ANN[52]	Fortaleza and Natal, Brazil	Wind speed	Wind speed	1 month, 1 h	Singlestep	ARIMA, ARIMAX, HW, ANN	1 month: −1.52%–0.74%, 1 h: −0.8%–12.87% (MAPE)
LSSVM-PSR[249]	Jiangsu, China	Wind power	Wind power	10 min	3-step-ahead	ANN, add-weighted one-rank local-region, gray model, PSO-based nonlinear gray Bernoulli model	1-step: 0.04%–1.65%, 2-step: 0.06%–1.64%, 3-step: 0.45%–2.59% (NMAE)
LSTM-HELM-DE[53]	Inner Mongolia, China	Wind speed	Wind speed	10 min, 1 h	Singlestep	ARIMA, ANN, SVR, ELM, LSTM	0.16002%–1.41506% (MAPE)
LSTM-SVR-EO[227]	Inner Mongolia, China	Wind speed	Wind speed	10 min, 1 h	Singlestep	ARIMA, SVR, KNN, ANN and gradient boosting regression tree	10 min: 0.0745%–1.0900%, 1 h: 0.6468%–3.6227% (MAPE)
PACF-EMD-feed-forward neural network[128]	Zhangye, China.	Wind speed	Wind speed	1 month, 24h	12-step (monthly), 10-step (daily)	FNN, EMD-FNN	1 month: 3.10%–3.76%, 24h: 3.89%–12.55%
RELM-PSR[249]	Jiangsu, China	Wind power	Wind power	10 min	3-step-ahead	ANN, add-weighted one-rank local-region, gray model, PSO-based nonlinear gray Bernoulli model	1-step: 0.39%–2.00%, 2-step: 0.37%–1.95%, 3-step: 0.40%–2.54% (NMAE)
SARIMA-PCA-balancing-ANN[250]	Macau and Petrolina, Brazil	Temperature, humidity, pressure, direction, wind speed	Wind speed	6 h	4-step-ahead	SARIMA, SARIMA-Wavelet, NN	1-step: 0.25–0.92,2-step: 0.11–0.90, 3-step: 0.18–1.05, 4-step: 0.25–1.10 (MAE)
SIE-WD-GA-SVR[106]	Gansu, China	Wind speed	Wind speed	1 h	Singlestep	BPNN, SVR, SIE–BPNN, SIE–SVR	0.0250%–0.5488% (MAPE)
Time varying filter-based EMD-fuzzy entropy-SSA-KELM-HHO and GWO-ConvLSTM[251]	Sotavento Galicia, Beresford	Wind speed	Wind speed	10 min for Sotavento Galicia, 1 h for Beresford	3-step-ahead	SVR, KELM, LSTM, ConvLSTM, EMD-KELM, EMD-ConvLSTM, CEEMDAN-KELM, CEEMDAN-SSAPSR-KELM, CEEMDAN-SSAPSR-FS-KELM-MHHOGWO	1-step: 1.6336%–20.4800%, 2-step: 0.1780%–37.1696%, 3-step: 0.6657%–51.4949% (MAPE)
VMD-SSA-LSTM-ELM[124]	China	Wind speed	Wind speed	1 h	5-step-ahead	ARIMA, LSTM, ELM, VMD-ELM, VMD-LSTM-ELM, EMD-SSA-LSTM-ELM, WPD-LSTM-ELM	1-step: −0.93%–3.61%, 2-step: −0.62%–6.47%, 3-step: −0.20%–9.11%, 4-step: 0.15%–11.24%, 5-step: 0.28%–13.02% (MAPE)

(*Continued on next page*)

Table 7. *Continued*

Model	Data source	Input	Output	Interval	Step	Benchmark	Improvement
VMD-SSA-PSR-GWO-sine cosine algorithm-PSR-ELM[230]	Sotavento Galicia, Inner Mongolia	Wind speed	Wind speed	10 min for Sotavento Galicia, 30 min for Inner Mongolia	Singlestep	GS-SVM, GS-BP, GS-ELM, IHGWOSCA-PSR-ELM, SSA-IHGWOSCA-PSR-ELM, OVMD-SSA-CC-PSR-ELM, EMD-IHGWOSCA-PSR-ELM, OVMD-IHGWOSCA-PSR-ELM, EMD-SSA-IHGWOSCA-PSR-ELM	−0.3577%–18.4725% (MAPE)
					3-step-ahead		1-step: −0.3577%–18.4725%, 2-step: −0.6747%–32.3794%, 3-step: −0.1985%–31.4509% (MAPE)
WDD-WPD-ARMA-EMD-ELM-Outlier Correction Method[70]	/	Wind speed	Wind speed	20 min	5-step-ahead	ARIMA, BPNN, Elman, ELM, WPD-ELM	1-step: 0.02%–10.36%, 2-step: 3.00%–10.75%, 3-step: 4.76%–10.46%, 4-step: 3.06%–7.77%, 5-step: 2.59%–8.90% (MAPE)
WTD-RNN-ANFIS[127]	/	Wind speed	Wind speed	15 min	3-step-ahead	ANN, SVM, RNN, WTD-ANN, WTD-SVM, WTD-RNN	1-step: 0.0473–0.3116, 2-step: 0.0811–0.3872, 3-step: 0.7118–0.3386 (MAPE)

Table 8. Comparisons between different categories of deterministic forecasting methods

Categories	Inputs	Horizons	Merits	Drawbacks
Physical model				
NWP model	Forecasted meteorological values (wind direction, ambient temperature, surrounding humidity, etc.)	Medium term, long term	It has superior performance than time series models within the time horizon ranging from 3 to 6 h.	Limited effect during forecast period between 0 and 2 h, and the accuracy depends on initial conditions.
Spatial correlation model	Wind data, exogenous data (wind direction, turbine diameter/distance, etc.)	Short term, medium term	It fully considers physical information in nearby regions.	Suffering from a high demand for historical data in the simulation of complex spatiotemporal changes.
Statistical model				
Time series analysis-based model	Wind data, exogenous data (wind direction, wind shear, temperature, etc.)	Short term, medium term, long term	It can provide accurate predictions for mean wind speed by simple calculations.	Frequently used for ultra-short-term forecasting rather than long-term forecasting.
Kalman filter model	Wind data, NWP data	Medium term, long term	It has been extensively utilized in complicated physical models for meteorological purposes.	Given its linear form and the discontinuity of wind series in certain cases, its application in the field of wind forecasting may be erroneous.
Machine learning-based model	Wind data, exogenous data such as temperature	Ultra-short term, short term	It has a strong learning ability and can better adapt to complex nonlinear relationships.	Restricted to short-term forecasts as it may not perceive hidden information with long-term dependence.
Hybrid model	Wind data, exogenous data (humidity, pressure, direction, etc.)	Ultra-short/short/medium/long term	It can achieve superior predictive performance by combining a variety of techniques.	The trade-off between forecast accuracy and predictive efficiency remains to be retrieved.

$$c_i = \begin{cases} 0 & y_i \notin (L_i, U_i) \\ 1 & y_i \in (L_i, U_i) \end{cases} \quad \text{(Equation 9)}$$

$$\text{PINAW} = \frac{1}{nR}\sum_{i=1}^{n}\left(U_i - L_i\right) \times 100\% \quad \text{(Equation 10)}$$

$$\text{PINRW} = \frac{1}{R}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(U_i - L_i)^2} \times 100\% \quad \text{(Equation 11)}$$

$$\text{ACE} = \text{PICP} - \text{PINC} \quad \text{(Equation 12)}$$

where n denotes the number of samples, U_i and L_i denote the upper bound and lower bound, y_i denotes the observed wind speed/power value, R denotes the target variable range of maximum and minimum values.

In view of the contradiction between interval coverage and interval width, a comprehensive index named CWC[257–260] was developed and further improved to better assess the overall performance of prediction models. The original and improved CWC expressions are as follows.

$$CWC_{original} = \begin{cases} \text{PINAW}\ (\text{PICP} \geq \text{PINC}) \\ \text{PINAW} + \exp(\eta(\text{PINC} - PICP))(\text{PICP} < \text{PINC}) \end{cases} \quad \text{(Equation 13)}$$

$$CWC_{improved} = \begin{cases} \eta_1 \cdot \text{PINAW}\ (\text{PICP} \geq \text{PINC}) \\ (0.1 + \eta_1 \cdot \text{PINAW})\left[1 + \exp(\eta_2(\text{PINC} - PICP))\right](\text{PICP} < \text{PINC}) \end{cases} \quad \text{(Equation 14)}$$

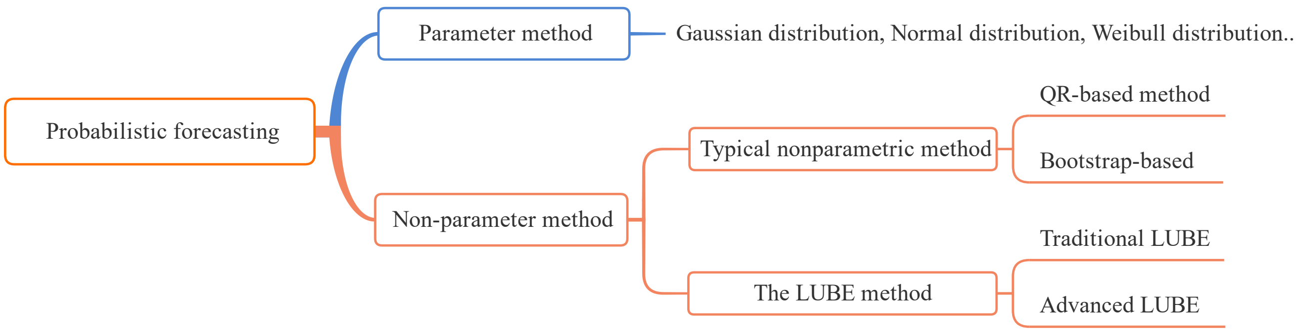

Figure 8. The classification of the probabilistic method

where η_1 is used to magnify PINAW and η and η_2 are used to magnify the difference between PICP and PINC.

INAD[256] describes the deviation of observed values from the prediction interval, reflecting the reliability of the prediction model to a certain extent, with smaller INAD indicating lower deviations. The formulas of INAD are,

$$\mathrm{INAD} = \frac{1}{n}\sum_{i=1}^{n}\varepsilon_i \qquad \text{(Equation 15)}$$

$$\varepsilon_i = \begin{cases} \dfrac{L_i - y_i}{y_i}, y_i < L_i \\ 0, y_i \in (L_i, U_i) \\ \dfrac{y_i - U_i}{y_i}, y_i > U_i \end{cases} \qquad \text{(Equation 16)}$$

where ε_i denotes the deviation value, n denotes the number of samples, U_i and L_i denote the upper and lower bounds, and y_i denotes the observed wind speed/power.

Classification of probabilistic forecasting method

Probabilistic forecasting methods estimate wind speed/power in the form of probability density or probability interval,[262–264]which is composed of parametric and nonparametric methods[265,266] as shown in Figure 8.

Parametric method

The parametric method is implemented by assuming that prediction errors or wind fluctuations obey a certain probability distribution, which can directly output probability intervals under different confidence levels, or obtain residual distributions of wind speed/power through the probability density function. A variety of distribution functions have been proposed, because a particular distribution function cannot be adapted to all forecasting scenarios. The representative distribution functions and the corresponding literature are summarized in Table 9, where ρ and υ respectively denote wind power and wind speed. Note that the parametric method cannot determine the accuracy of the presumed distributions and therefore cannot guarantee the validity of the predictive results.

Nonparametric method

The nonparametric method is essentially a data-driven method that requires a large number of samples and densities or quantiles for estimation, which can be divided into typical nonparametric methods and LUBE methods.

Typical nonparametric method. Typical nonparametric methods train NNs by minimizing error-based cost functions, of which quantile regression (QR) and bootstrap are two commonly used techniques. Bremnes et al. (2004)[283] conducted probabilistic wind power forecasts with 24–47 h forecast

Table 9. Summary of representative distributions in parametric methods (in alphabetical order)

Distribution/References	Variable	Horizon	Output	Result
Beta distribution[267]	ρ	1h, 24h	Distribution of wind power error	The curve is symmetric.
Beta distribution[268]	ρ	10 min	Interval of wind power	The coverage reaches 95% with narrower prediction interval.
Beta distribution[269]	ρ	6h, 48h	Interval of wind power	Intervals of the expected actual power may be derived for various confidence levels based on probability distributions.
Cauchy distribution[270]	ρ	15 min	Interval of wind power	It can achieve more accurate modeling of the true wind power output confidence intervals.
Gamma distribution[271]	υ	6 h	Distribution of wind speed	It provides an adequate and unified description almost everywhere.
Gaussian distribution[254]	ρ	1–12h	Distribution of wind power error	Its distribution is approximately symmetric with a kurtosis of exact 3.
Gaussian distribution[272]	υ/ρ	12h	Distributions of wind speed/power errors	Distributions of wind speed are systematically, while that for wind power are typically not symmetric with a large peak.
Gumbel distribution[273]	υ	24 h	Distribution of wind speed	Gumbel distributions based on the maximum likelihood estimates has superiority on modeling.
Kappa distribution[274]	υ	1 h	Distribution of wind speed	It can obtain the best performances for one-component parametric models.
Logit-normal distribution[275]	ρ	10 min	Distribution of wind power	It shows superiority over classical assumptions about the shape of predictive densities.
Lorenz disturbance distribution[276]	υ	10 min	Interval of wind speed	Its interval estimations have great prediction accuracy at 95 and 80% confidence level.
Mixed distribution[277]	ρ	1 h	Distribution of wind power error	It is a far better approximation to the actual data, and is a good approximation (especially within the limit on outliers).
Normal distribution[278]	ρ	10 min	Distribution of wind power error	The error distribution is likely symmetric
Rayleigh distribution[279]	υ	1 min	Distribution of wind speed	Its accuracy is lower than Weibull predictions are in overall.
Versatile distribution[280]	ρ	/	Distribution of wind power error	It can well represent the distribution of wind power forecast errors for all forecast timescales and magnitudes.
Weibull distribution[281]	υ	1h	Distribution of wind speed/power	It better fit to the measured probability distributions, and returns smaller error values in calculating the power density.
Weibull distribution[282]	υ	10 min	Distribution of wind power	The maximum likelihood method provides more accurate estimation of Weibull parameters.

horizon using local QR. Wan et al. (2013)[284] employed several bootstrap methods to predict wind power intervals, with the pairs bootstrap shows the optimal performance. Zhang et al. (2016)[249] performed a probabilistic interval prediction using quantile regression averaging. Wan et al. (2017)[255] developed a multi-step probabilistic forecasting model using direct QR to estimate wind power from 10 min to 3 h. Ji et al. (2019)[285] constructed prediction intervals using the bootstrap method after obtaining deterministic results. Peng et al. (2021)[286] predicted wind power 15 min in advance via the QR method. Although the above models can provide superior predictions than benchmark models, the resulting forecast intervals focus on minimizing errors rather than optimizing interval characteristics because they are trained based on error-based cost functions.

The LUBE method. The LUBE[287] method can directly generate upper and lower bounds of the prediction interval (PI) based on machine learning. Depending on the predictors and optimizers introduced, LUBE methods can be divided into traditional LUBE method and advanced LUBE method.

Traditional LUBE model is established based on shallow neural networks, which shares a similar working principle as ANN, as illustrated in Figure 9. Prediction models are first constructed from shallow neural networks with two output neurons, and then single- or multi-objective optimization functions are designed by considering interval characteristics (interval coverage, interval width, etc.). Finally, the devised LUBE model is trained iteratively, and the final predictive interval can be obtained after satisfying the convergence conditions.

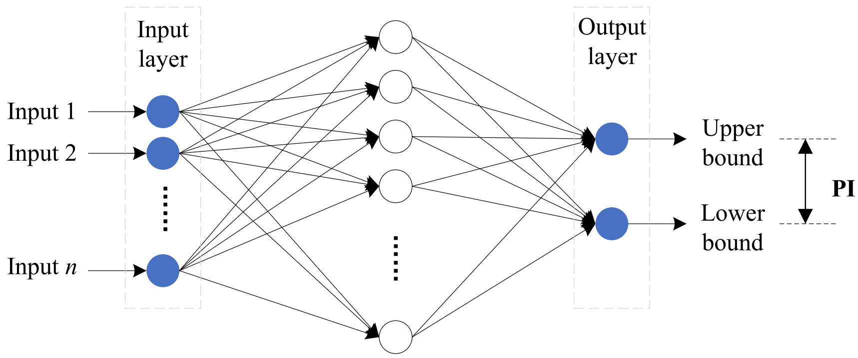

Figure 9. The network structure of the predictor in the ANN-based LUBE model

Although traditional LUBE methods address the problems existing in typical nonparametric probabilistic methods, there remain several obvious limitations. First, the learning capability of shallow machine learning models is insufficient, so that the performance of the traditional LUBE models is limited. Specifically, ANN lacks the ability to model or abstract data, SVM is unable to handle large amounts of data, and ELM possesses weak general approximation characteristics. Second, gradient descent (GD) algorithms are standard training techniques for NNs, but are rarely introduced in traditional LUBE models, because the designed loss functions are essentially nonlinear and non-differential. In contrast, heuristic algorithms, such as CSS, BA, and PSO, are widely introduced for parameter optimization. Considering the considerable number of parameters to be tuned within the LUBE framework, the computational cost is relatively expensive, which limits the efficiency of model training.

To solve the above issues, advanced LUBE models using DNNs as predictors have been proposed, which are implemented through two means. One approach is to improve loss functions in traditional LUBE models based on PI characteristics, thereby improving optimization effects of heuristic algorithms. Another approach is to construct linearly differentiable loss functions and then introduce optimization algorithms with higher learning rates and training efficiency.

The traditional and advanced LUBE models proposed in the most recent literature are summarized chronologically in Table 10, where η denotes PINC, ρ and υ denote wind power and wind speed.

Comparison of probabilistic forecasting models

The comparisons of different probabilistic forecasting methods are summarized in Table 11.

DISCUSSIONS

The mainstream of wind forecasting methods

Deterministic forecasting methods have undergone a long period of development compared to probabilistic forecasting methods, and thus have a comprehensive knowledge system that has contributed significantly to the field of wind forecasting.[25] Among the various deterministic forecasting methods, the development of hybrid models has received great attention,[309] especially those that combine deep predictors with practical techniques such as data decomposition.[2]

In contrast, probabilistic forecasting method is still at a rapid development stage, but there is a consensus that its application can trigger higher economic benefits than deterministic forecasting method.[33] Among the probabilistic wind forecasting method, nonparametric methods (known as ''distribution-free'') make more contributions than parametric methods.[265] Because the demand for wind forecasting increases in practical engineering applications, probabilistic wind forecasting methods will inevitably become the mainstream of wind forecasting in the future.

Future trend of wind forecasting

Superior adaptiveness of wind forecasting models

Effective prediction models should be adaptive to different types of wind farms rather than concentrating on a specific wind farm. More specifically, although strong compliancy to a particular wind farm can

Table 10. Summary of the work using LUBE method in chronological order

Year	Predictor	Variable	Optimization	Horizon	Functions	Results
Traditional LUBE methods						
2014	ANN[24]	ρ	PSO	24 h	CWC	η = 90%: 90%–92.30% (PICP), 0.1605–0.7257 (PINAW/CWC)
2015	ELM[288]	ρ	quantum-behaved PSO	1 h	Deviation	η = 90%: 90.30% (PICP),0.3025 (PINAW) η = 80%: 81.62% (PICP),0.2017 (PINAW)
2015	MLP[289]	v	NSGA-II	24 h	PICP, width	$\eta \in$ (90%,100%): 91.40%–99.70% (PICP), 0.452–0.739 (width)
2015	BPNN[290]	v	CS	1 h	Interval errors	η = 95%: 0.3433–0.5062 (CWC) η = 90%: 0.2347–0.3403 (CWC) η = 85%: 0.2015–0.2520 (CWC) η = 80%: 0.1688–1.0585 (CWC)
2016	feedforward NN[291]	v	BA	15 min	PINAW, PICP	η = 85%: 93.49% (PICP), 0.3296 (PINAW)
2016	ANN[292]	ρ	ABC	30 min	CWC	η = 90%: 90.09% (PICP), 0.1949 (RPIW) η = 95%: 95.94% (PICP), 0.2345 (RPIW) η = 99%: 99.38% (PICP), 0.2577 (RPIW)
2016	SVM[261]	v	MODE	30 min	ACE, Winkler score	η = 90%: 90.77–92.56% (PICP), 0.77–2.56 (ACE) η = 95%: 94.05%–96.13% (PICP), −0.95–1.13 (ACE) η = 99%: 96.43%–99.7% (PICP), −2.57–0.7 (ACE)
2017	BPNN[293]	ρ	PSO	1 h	CWC	η = 80%: 0.0983 (PINAW), 0.0379 (ACE) η = 90%: 0.1943 (PINAW), 0.0664 (ACE)
2017	KELM[294]	ρ	IABC	5–15 min	Errors, PINAW, INAD	η = 90%, 5-min-ahead: 90.75% (PICP), 0.176 (PINAW), 0.005 (INAD) η = 90%, 10-min-ahead: 90.448% (PICP), 0.224 (PINAW), 0.01 (INAD) η = 90%, 15-min ahead: 90.305% (PICP), 0.229 (PINAW), 0.016 (INAD)
2018	ANN[295]	ρ	CSS	1 h, 24 h	CWC	1-h-ahead CWC: 0.00580–0.1097 (η = 50%), 0.0745–0.1384 (η = 60%), 0.0902–0.1759 (η = 70%), 0.1105–0.2243 (η = 80%), 0.1443–0.2990 (η = 90%) 24-h-ahead CWC: 0.1097 (η = 50%), 0.1384 (η = 60%), 0.1759 (η = 70%), 0.2243 (η = 80%), 0.2923 (η = 90%)
2018	LSSVM[296]	v	MOALO	1 h	PICP, PINAW	η = 95%: 0.48 (CWC) η = 90%: 0.359–0.403 (CWC) η = 85%: 0.30–0.353 (CWC)
2019	LSSVM[297]	ρ	MOSSO	10–60 min	PICP, PINAW	10-min-ahead CWC: 0.276 (η = 95%), 0.232 (η = 90%), 0.200 (η = 85%) 30-min-ahead CWC: 0.290 (η = 95%), 0.245 (η = 90%), 0.150 (η = 85%) 60-min-ahead CWC: 0.435 (η = 95%), 0.332 (η = 90%), 0.303 (η = 85%)
2021	BPNN[298]	ρ	Electromagnetism-like	90 min	PICP, mean interval bound	From 1 to 6 steps with 90-min interval: PICP: 98.88%, 98.5%, 99.75%, 98.75%, 98.88%, 99.1% Degree: 0.7712, 0.6875, 0.7775, 0.7338, 0.7288, 0.7462
2022	FLN[299]	ρ	WOA	10 min	PICP, PINRW, deviation	η = 90%: 94.80% (PICP), 0.1399 (PINRW) η = 80%: 85.40% (PICP), 0.1041 (PINRW)

(*Continued on next page*)

Table 10. *Continued*

Year	Predictor	Variable	Optimization	Horizon	Functions	Results
Advanced LUBE methods						
2018	ENN[300]	ρ	DA	1 h	NCWC	η = 90%: 93.81%–94.24% (PICP), 0.620–0.664 (PINAW/CWC)
2021	LSTM[301]	ρ	NSGA-II	15 min,1 h	PINAW, PIEE	15-min-ahead: 0.00978 (PINAW), 0.121 (PIEE) 1-h-ahead: 0.00899 (PINAW), 0.1647 (PIEE)
2020	BiLSTM[302]	υ	Adam	30 min	$f_1(W,b)$ $f_2(W,b)$	η = 90%: 93.44% (PICP), 0.0183 (PINAW), 0.8139 (CWC)
2020	LSTM[303]	ρ	GD	10 min	$f_1(W,b)$ $f_2(W,b)$	η = 90%: 95% (PICP), 0.18 (PINRW), 0.01 (INAD), 1.03 (CWC)
2020	GRU[304]	υ	GD	10 min	$UB_{con_{PM}}$ $LB_{con_{PM}}$	η = 90%: 96.40% (PICP), 0.0703 (PINRW),1.0703 (CWC)
2020	GRU[305]	ρ	Adam	10 min	f_{cost}	η = 90%: 94.770% (PICP), 0.0659 (PINRW), 1.4165 (CWC)
2020	GRU[306]	υ	Adam	20 min	L_n and U_n	η = 90%: 93.47%–94.72 (PICP), 0.0711–0.0919 (PINRW), 0.0090–0.0130 (INAD), 0.5186–0.6331 (CWC)
2021	TCN[307]	υ	Adam	15 min	Y_i	η = 90%: 92.50% (PICP), 0.0906 (PINRW), 0.3212 (CWC)
2022	GRU[308]	ρ	SGD	10 min	SCWC	η = 90%: 97.54% (PINCP), 16.62 (PINCW)

guarantee the superiority of predictive results, the prediction performance will be significantly reduced when applying to distinct wind farms. Therefore, the exploration of adaptive wind forecasting models is of great importance, especially considering that designing individual models for each of the different wind fields is time-consuming and labor-intensive.[42] It is expected that evaluations of model performances will focus on both model validity and adaptability in the future. In this regard, the spatial correlation between different wind farms can be analyzed,[43] and technologies such as transfer learning[188] and reinforcement learning[310] can be introduced. In addition, adaptive model hyperparameters adjustment can be achieved through automatic optimization algorithms[311] or adaptive adjustment strategies.[305,306]

Longer forecasting horizon of wind forecasting methods

The development of wind energy will inevitably increase the predictability requirements for wind forecasting methods, and thus long-term forecasts will play a critical role in the reformed electricity market.[13] To date, numerous long-term forecasting models have been successfully applied to wind forecasting, such as 72-h ahead[312] and 1 to 10 days ahead.[264] It is expected that wind forecasts with more forward-looking information, especially multi-step ahead forecasts, will receive higher attention, although their performance will decrease as the number of steps increases.[20] Furthermore, although most of the existing wind energy prediction methods have been developed by analyzing the relationship between historical sequences and output variables, the application of forecast data (such as NWP data[13]) with forward-looking wind information will be further strengthened, thus improving the predictability of future medium-term and long-term forecasts.

Combining classical and advanced knowledge

The combination of classical knowledge (e.g., metaheuristic algorithms) and advanced techniques (e.g., machine learning) will remain valuable for wind forecasting. Meanwhile, applications of novel technologies will continue to evolve,[43] mainly through three promising techniques: (1) Simplifying the structure of existing predictors or reintroducing mechanisms to existing predictors. For example, Li et al. (2022)[313] improved the Kalman filter using PSO for wind power prediction. Liu et al. (2022)[314] combined the graph convolutional neural with GRU to capture spatial influence in wind forecasting. (2) Improving current optimization algorithms by combining different optimization algorithms or developing novel optimization algorithms. For example, Tuerxun et al. (2022)[315] developed the modified tuna swarm optimization algorithm to achieve ultra-short-term wind speed prediction. Wei et al. (2022)[316] proposed IGWO with the back-propagation neural network to predict short-term wind power. (3) Integrating techniques from other domains, such as attention mechanisms,[286,317] Clayton Copula function,[318] and Granular Computing,[319] into the field of wind forecasting.

Table 11. The comparison between different categories of probabilistic forecasting models

Categories	Input	Predictors	Outputs	Horizons	Features
Parametric method	Wind data, exogenous data	Distribution functions	Interval/ distribution	Ultra-short/short/ medium/long term	It cannot guarantee the accuracy of the hypothetical distribution, thus failing to ensure the validity of results.
Typical nonparametric method	Wind data, exogenous data	QR/bootstrap	Probability quantile/interval	Ultra-short/short/ medium/long term	This method focuses only on minimizing errors, while ignoring the interval characteristics.
Traditional LUBE method	Wind data, exogenous data	Shallow neural networks	Interval	Mainly for ultra-short-term and short-term	The adopted shallow predictors have insufficient learning capability, and parameter tuning is achieved by heuristic algorithms.
Advanced LUBE method	Wind data, exogenous data	Deep neural networks	Interval	Mainly for ultra-short-term and short-term	DNNs and GD algorithms can be used by improving or reconfiguring loss functions, but the model structure may be relatively complex.

The remaining limitation of this work

Overall, this work fills the research gaps that existed in previous studies, but there remain some limitations that warrant continued improvement in the future. On one hand, even if we have examined the existing literature as comprehensively as possible, relevant studies are not exhaustively enumerated in this work, especially given that wind forecasting models are rapidly evolving and frequently updated. On the other hand, because of the wide range of models involved, comparisons of all the different types of models through experimental results are not presented, thus failing to visualize the quantitative effects of different forecasting models.

CONCLUSIONS

This paper provides a systematic review of wind forecasting methods. Both deterministic and probabilistic forecasting methods are thoroughly surveyed from the perspectives of technical background, theoretical basis, and model performance. The deterministic forecasting methods have been classified into physical model, statistical model and hybrid model. In particular, the structures and variants of machine learning models have been compared, and the ensemble models have been summarized in terms of model input, time horizon and prediction effects. In addition, a comprehensive investigation of the probabilistic prediction models, which cover parametric methods and nonparametric methods, is provided. Representative distribution functions have been compared according to time scales and prediction results. Specifically, the increasingly popular LUBE models have been thoroughly investigated. Then, the mainstream and promising development trends of wind forecasting methods are discussed, and the outlook is provided from the aspects of model adaptability, model predictability, and the hybridization of classical and advanced techniques. In summary, this work offers a broad and insightful overview of wind forecasting research, which enables reliable integration of renewable energy resources. It is expected that junior researchers can benefit from our work to gain an all-round understanding of wind energy forecasting, thus identifying the models that best suits their research objectives.

ACKNOWLEDGMENTS

This work is partially supported by the National Natural Science Foundation of China (No. 51879111, No. 51679095), by the Applied Fundamental Frontier Project of Wuhan Science and Technology Bureau (No. 2018010401011269), by the Hubei Provincial Natural Science Foundation of China (No. 2019CFA068) and the Key Research and Development Program of Hubei Province of China (No. 2021BAA193).

AUTHOR CONTRIBUTIONS

Conceptualization, Y.X. and C.L.; Methodology, Y.X., C.L., and M.L.; Investigation, F.L.; Writing-Original Draft, Y.X.; Writing–Review and Editing, Y.X., C.L., M.L., and M.T.; Supervision, C.L. and M.L.; Funding Acquisition, C.L.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Wang, J., Song, Y., Liu, F., and Hou, R. (2016). Analysis and application of forecasting models in wind power integration: a review of multi-step-ahead wind speed forecasting models. Renew. Sustain. Energy Rev. *60*, 960–981.
2. Ma, Z., Chen, H., Wang, J., Yang, X., Yan, R., Jia, J., and Xu, W. (2020). Application of hybrid model based on double decomposition, error correction and deep learning in short-term wind speed prediction. Energy Convers. Manag. *205*, 112345.
3. Akçay, H., and Filik, T. (2017). Short-term wind speed forecasting by spectral analysis from long-term observations with missing values. Appl. Energy *191*, 653–662.
4. Naik, J., Bisoi, R., and Dash, P. (2018). Prediction interval forecasting of wind speed and wind power using modes decomposition based low rank multi-kernel ridge regression. Renew. Energy *129*, 357–383.
5. Doherty, R., and O'malley, M. (2005). A new approach to quantify reserve demand in systems with significant installed wind capacity. IEEE Trans. Power Syst. *20*, 587–595.
6. Schmall, J., Huang, S.-H., Li, Y., Billo, J., Conto, J., and Zhang, Y. (2015). Voltage Stability of Large-Scale Wind Plants Integrated in Weak Networks: An ERCOT Case Study (IEEE), pp. 1–5.
7. Ancell, G., and Clarke, J. (2008). Integration of Large Scale Wind Generation in the New Zealand Power System and Electricity Market (IEEE), pp. 1–6.
8. Qian, Z., Pei, Y., Zareipour, H., and Chen, N. (2019). A review and discussion of decomposition-based hybrid models for wind energy forecasting applications. Appl. Energy *235*, 939–953.
9. Hill, D.C., McMillan, D., Bell, K.R.W., and Infield, D. (2012). Application of auto-regressive models to UK wind speed data for power system impact studies. IEEE Trans. Sustain. Energy *3*, 134–141.
10. Chang, G., Lu, H., Chang, Y., and Lee, Y. (2017). An improved neural network-based approach for short-term wind speed and power forecast. Renew. Energy *105*, 301–311.
11. Yang, W., Wang, J., Lu, H., Niu, T., and Du, P. (2019). Hybrid wind energy forecasting and analysis system based on divide and conquer scheme: a case study in China. J. Clean. Prod. *222*, 942–959.
12. Zhu, X., and Genton, M.G. (2012). Short-term wind speed forecasting for power system operations. Int. Stat. Rev. *80*, 2–23.
13. Soman, S.S., Zareipour, H., Malik, O., and Mandal, P. (2010). A Review of Wind Power and Wind Speed Forecasting Methods with Different Time Horizons (IEEE), pp. 1–8.
14. Kusiak, A., and Li, W. (2010). Short-term prediction of wind power with a clustering approach. Renew. Energy *35*, 2362–2369.
15. Amjady, N., Keynia, F., and Zareipour, H. (2011). Wind power prediction by a new forecast engine composed of modified hybrid neural network and enhanced particle swarm optimization. IEEE Trans. Sustain. Energy *2*, 265–276.
16. Hao, Y., Dong, L., Liao, X., Liang, J., Wang, L., and Wang, B. (2019). A novel clustering algorithm based on mathematical morphology for wind power generation prediction. Renew. Energy *136*, 572–585.
17. Dong, L., Wang, L., Khahro, S.F., Gao, S., and Liao, X. (2016). Wind power day-ahead prediction with cluster analysis of NWP. Renew. Sustain. Energy Rev. *60*, 1206–1212.
18. Tasnim, S., Rahman, A., Oo, A.M.T., and Haque, M.E. (2018). Wind power prediction in new stations based on knowledge of existing Stations: a cluster based multi source domain adaptation approach. Knowl. Base. Syst. *145*, 15–24.
19. Lang, J., Peng, X., Li, W., Cai, T., Gan, Z., Duan, S., and Li, C. (2021). A novel two-stage interval prediction method based on minimal gated memory network for clustered wind power forecasting. Wind Energy *24*, 450–464.
20. Ben Taieb, S., Bontempi, G., Atiya, A.F., and Sorjamaa, A. (2012). A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition. Expert Syst. Appl. *39*, 7067–7083.
21. Niu, T., Wang, J., Zhang, K., and Du, P. (2018). Multi-step-ahead wind speed forecasting based on optimal feature selection and a modified bat algorithm with the cognition strategy. Renew. Energy *118*, 213–229.
22. Li, C., Xiao, Z., Xia, X., Zou, W., and Zhang, C. (2018). A hybrid model based on synchronous optimisation for multi-step short-term wind speed forecasting. Appl. Energy *215*, 131–144.
23. Liu, H., Li, Y., Duan, Z., and Chen, C. (2020). A review on multi-objective optimization framework in wind energy forecasting techniques and applications. Energy Convers. Manag. *224*, 113324.
24. Quan, H., Srinivasan, D., and Khosravi, A. (2014). Short-term load and wind power forecasting using neural network-based prediction intervals. IEEE Trans. Neural Netw. Learn. Syst. *25*, 303–315.
25. Li, L.-L., Zhao, X., Tseng, M.-L., and Tan, R.R. (2020). Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm. J. Clean. Prod. *242*, 118447.
26. Xiang, L., Li, J., Hu, A., and Zhang, Y. (2020). Deterministic and probabilistic multi-step forecasting for short-term wind speed based on secondary decomposition and a deep learning method. Energy Convers. Manag. *220*, 113098.
27. Pinson, P. (2006). Estimation of the Uncertainty in Wind Power Forecasting (École Nationale Supérieure des Mines de Paris).
28. Juban, J., Siebert, N., and Kariniotakis, G.N. (2007). Probabilistic Short-Term Wind Power Forecasting for the Optimal Management of Wind Generation (IEEE), pp. 683–688.
29. Yang, M., Zhu, S., Liu, M., and Lee, W.-J. (2014). One parametric approach for short-term JPDF forecast of wind generation. IEEE Trans. Ind. Applicat. *50*, 2837–2843.
30. Yang, M., Lin, Y., and Han, X. (2016). Probabilistic wind generation forecast based on sparse Bayesian classification and Dempster–Shafer theory. IEEE Trans. Ind. Applicat. *52*, 1998–2005.
31. Lin, Y., Yang, M., Wan, C., Wang, J., and Song, Y. (2019). A multi-model combination approach for probabilistic wind power forecasting. IEEE Trans. Sustain. Energy *10*, 226–237.
32. He, Y., Yan, Y., and Xu, Q. (2019). Wind and solar power probability density prediction via fuzzy information granulation and support vector quantile regression. Int. J. Electr. Power Energy Syst. *113*, 515–527.
33. Haque, A.U., Nehrir, M.H., and Mandal, P. (2014). A hybrid intelligent model for deterministic and quantile regression approach for probabilistic wind power forecasting. IEEE Trans. Power Syst. *29*, 1663–1672.
34. Yu, X., Zhang, W., Zang, H., and Yang, H. (2018). Wind power interval forecasting based on confidence interval optimization. Energies *11*, 3336.
35. Wu, Y.-K., Wu, Y.-C., Hong, J.-S., Phan, L.H., and Quoc, D.P. (2020). Probabilistic Forecast of Wind Power Generation with Data Processing and Numerical Weather Predictions (IEEE), pp. 1–11.
36. Costa, A., Crespo, A., Navarro, J., Lizcano, G., Madsen, H., and Feitosa, E. (2008). A review on the young history of the wind power short-term prediction. Renew. Sustain. Energy Rev. *12*, 1725–1744.

37. Lei, M., Shiyan, L., Chuanwen, J., Hongling, L., and Yan, Z. (2009). A review on the forecasting of wind speed and generated power. Renew. Sustain. Energy Rev. *13*, 915–920.

38. Al-Yahyai, S., Charabi, Y., and Gastli, A. (2010). Review of the use of numerical weather prediction (NWP) models for wind energy assessment. Renew. Sustain. Energy Rev. *14*, 3192–3198.

39. Tascikaraoglu, A., and Uzunoglu, M. (2014). A review of combined approaches for prediction of short-term wind speed and power. Renew. Sustain. Energy Rev. *34*, 243–254.

40. Yan, J., Liu, Y., Han, S., Wang, Y., and Feng, S. (2015). Reviews on uncertainty analysis of wind power forecasting. Renew. Sustain. Energy Rev. *52*, 1322–1330.

41. Marugán, A.P., Márquez, F.P.G., Perez, J.M.P., and Ruiz-Hernández, D. (2018). A survey of artificial neural network in wind energy systems. Appl. Energy *228*, 1822–1836.

42. Liu, H., Chen, C., Lv, X., Wu, X., and Liu, M. (2019). Deterministic wind energy forecasting: a review of intelligent predictors and auxiliary methods. Energy Convers. Manag. *195*, 328–345.

43. Yang, B., Zhong, L., Wang, J., Shu, H., Zhang, X., Yu, T., and Sun, L. (2021). State-of-the-art one-stop handbook on wind forecasting technologies: an overview of classifications, methodologies, and analysis. J. Clean. Prod. *283*, 124628.

44. Zendehboudi, A., Baseer, M.A., and Saidur, R. (2018). Application of support vector machine models for forecasting solar and wind energy resources: a review. J. Clean. Prod. *199*, 272–285.

45. Ren, Y., Suganthan, P., and Srikanth, N. (2015). Ensemble methods for wind and solar power forecasting—a state-of-the-art review. Renew. Sustain. Energy Rev. *50*, 82–91.

46. Pérez-Ortiz, M., Jiménez-Fernández, S., Gutiérrez, P., Alexandre, E., Hervás-Martínez, C., and Salcedo-Sanz, S. (2016). A review of classification problems and algorithms in renewable energy applications. Energies *9*, 607.

47. Wang, H., Lei, Z., Zhang, X., Zhou, B., and Peng, J. (2019). A review of deep learning for renewable energy forecasting. Energy Convers. Manag. *198*, 111799. https://doi.org/10.1016/j.enconman.2019.111799.

48. Lai, J.P., Chang, Y.M., Chen, C.H., and Pai, P.F. (2020). A survey of machine learning models in renewable energy predictions. Appl. Sci. *10*, 5975. https://doi.org/10.3390/app10175975.

49. Draxl, C., Clifton, A., Hodge, B.M., and McCaa, J. (2015). The wind integration national dataset (WIND) toolkit. Appl. Energy *151*, 355–366. https://doi.org/10.1016/j.apenergy.2015.03.121.

50. Rand, J.T., Kramer, L.A., Garrity, C.P., Hoen, B.D., Diffendorfer, J.E., Hunt, H.E., and Spears, M. (2020). A continuously updated, geospatially rectified database of utility-scale wind turbines in the United States. Sci. Data *7*, 15. https://doi.org/10.1038/s41597-020-0353-6.

51. Yadav, G.R., Muneender, E., and Santhosh, M. (2021). Wind speed prediction using hybrid long short-term memory neural network based approach. Adv. Mater. Sci. Eng. *21–23*, 1–6.

52. Camelo, H.d.N., Lucio, P.S., Leal Junior, J.B.V., Carvalho, P.C.M.d., and Santos, D.v.G.d. (2018). Innovative hybrid models for forecasting time series applied in wind generation based on the combination of time series models with artificial neural networks. Energy *151*, 347–357. https://doi.org/10.1016/j.energy.2018.03.077.

53. Hu, Y.L., and Chen, L. (2018). A nonlinear hybrid wind speed forecasting model using LSTM network, hysteretic ELM and Differential Evolution algorithm. Energy Convers. Manag. *173*, 123–142. https://doi.org/10.1016/j.enconman.2018.07.070.

54. Chavez, R., Gomez, H., Herbert, J.F., Romo, A., and Probst, O. (2013). Mesoscale modeling and remote sensing for wind energy applications. Rev. Mexic. Fisica *59*, 114–129.

55. Zhao, P., Wang, J., Xia, J., Dai, Y., Sheng, Y., and Yue, J. (2012). Performance evaluation and accuracy enhancement of a day-ahead wind power forecasting system in China. Renew. Energy *43*, 234–241. https://doi.org/10.1016/j.renene.2011.11.051.

56. Buhan, S., Ozkazanc, Y., and Cadirci, I. (2016). Wind pattern recognition and reference wind mast data correlations with NWP for improved wind- electric power forecasts. IEEE Trans. Ind. Inform. *12*, 991–1004. https://doi.org/10.1109/TII.2016.2543004.

57. Cheng, F.Y., Hsu, Y.C., Lin, P.L., and Lin, T.H. (2013). Investigation of the effects of different land use and land cover patterns on Mesoscale meteorological simulations in the taiwan area. J. Appl. Meteorol. Climatol. *52*, 570–587. https://doi.org/10.1175/JAMC-D-12-0109.1.

58. Liu, W.T., Tang, W., and Polito, P.S. (1998). NASA scatterometer provides global ocean-surface wind fields with more structures than numerical weather prediction. Geophys. Res. Lett. *25*, 761–764. https://doi.org/10.1029/98GL00544.

59. Sánchez, I. (2006). Short-term prediction of wind energy production. Int. J. Forecast. *22*, 43–56. https://doi.org/10.1016/j.ijforecast.2005.05.003.

60. Ye, L., Zhao, Y., Zeng, C., and Zhang, C. (2017). Short-term wind power prediction based on spatial model. Renew. Energy *101*, 1067–1074. https://doi.org/10.1016/j.renene.2016.09.069.

61. Zhang, W., Qu, Z., Zhang, K., Mao, W., Ma, Y., and Fan, X. (2017). A combined model based on CEEMDAN and modified flower pollination algorithm for wind speed forecasting. Energy Convers. Manag. *136*, 439–451. https://doi.org/10.1016/j.enconman.2017.01.022.

62. Liu, H., and Chen, C. (2019). Data processing strategies in wind energy forecasting models and applications: a comprehensive review. Appl. Energy *249*, 392–408. https://doi.org/10.1016/j.apenergy.2019.04.188.

63. Liu, H., Mi, X., and Li, Y. (2018). Comparison of two new intelligent wind speed forecasting approaches based on wavelet packet decomposition, complete ensemble empirical mode decomposition with adaptive noise and artificial neural networks. Energy Convers. Manag. *155*, 188–200. https://doi.org/10.1016/j.enconman.2017.10.085.

64. Kiplangat, D.C., Asokan, K., and Kumar, K.S. (2016). Improved week-ahead predictions of wind speed using simple linear models with wavelet decomposition. Renew. Energy *93*, 38–44. https://doi.org/10.1016/j.renene.2016.02.054.

65. Liu, H., Tian, H.-q., and Li, Y.-f. (2015). Four wind speed multi-step forecasting models using extreme learning machines and signal decomposing algorithms. Energy Convers. Manag. *100*, 16–22. https://doi.org/10.1016/j.enconman.2015.04.057.

66. Liu, H., Tian, H.-q., Chen, C., and Li, Y.-f. (2013). An experimental investigation of two Wavelet-MLP hybrid frameworks for wind speed prediction using GA and PSO optimization. Int. J. Electr. Power Energy Syst. *52*, 161–173. https://doi.org/10.1016/j.ijepes.2013.03.034.

67. Liu, H., Wu, H., and Li, Y. (2018). Smart wind speed forecasting using EWT decomposition, GWO evolutionary optimization, RELM learning and IEWT reconstruction. Energy Convers. Manag. *161*, 266–283. https://doi.org/10.1016/j.enconman.2018.02.006.

68. Liu, H., Mi, X.-w., and Li, Y.-f. (2018). Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network. Energy Convers. Manag. *156*, 498–514. https://doi.org/10.1016/j.enconman.2017.11.053.

69. Li, Y., Wu, H., and Liu, H. (2018). Multi-step wind speed forecasting using EWT decomposition, LSTM principal computing, RELM subordinate computing and IEWT reconstruction. Energy Convers. Manag. *167*, 203–219. https://doi.org/10.1016/j.enconman.2018.04.082.

70. Mi, X.-w., Liu, H., and Li, Y.-f. (2017). Wind speed forecasting method using wavelet, extreme learning machine and outlier correction algorithm. Energy Convers. Manag. *151*, 709–722. https://doi.org/10.1016/j.enconman.2017.09.034.

71. Liu, H., Tian, H.-q., Pan, D.-f., and Li, Y.-f. (2013). Forecasting models for wind speed using wavelet, wavelet packet, time series and Artificial Neural Networks. Appl. Energy

107, 191–208. https://doi.org/10.1016/j.apenergy.2013.02.002.

72. Meng, A., Ge, J., Yin, H., and Chen, S. (2016). Wind speed forecasting based on wavelet packet decomposition and artificial neural networks trained by crisscross optimization algorithm. Energy Convers. Manag. *114*, 75–88. https://doi.org/10.1016/j.enconman.2016.02.013.

73. Wang, J.-Z., Wang, Y., and Jiang, P. (2015). The study and application of a novel hybrid forecasting model – a case study of wind speed forecasting in China. Appl. Energy *143*, 472–488. https://doi.org/10.1016/j.apenergy.2015.01.038.

74. Aghajani, A., Kazemzadeh, R., and Ebrahimi, A. (2016). A novel hybrid approach for predicting wind farm power production based on wavelet transform, hybrid neural networks and imperialist competitive algorithm. Energy Convers. Manag. *121*, 232–240. https://doi.org/10.1016/j.enconman.2016.05.024.

75. Catalão, J., Pousinho, H.M.I., and Mendes, V.M.F. (2011). Short-term wind power forecasting in Portugal by neural networks and wavelet transform. Renew. Energy *36*, 1245–1251. https://doi.org/10.1016/j.renene.2010.09.016.

76. Ren, Y., Suganthan, P.N., and Srikanth, N. (2015). A comparative study of empirical mode decomposition-based short-term wind speed forecasting methods. IEEE Trans. Sustain. Energy *6*, 236–244. https://doi.org/10.1109/TSTE.2014.2365580.

77. An, X., Jiang, D., Zhao, M., and Liu, C. (2012). Short-term prediction of wind power using EMD and chaotic theory. Commun. Nonlinear Sci. Numer. Simulat. *17*, 1036–1042. https://doi.org/10.1016/j.cnsns.2011.06.003.

78. Amjady, N., and Abedinia, O. (2017). Short term wind power prediction based on improved kriging interpolation, empirical mode decomposition, and closed-loop forecasting engine. Sustainability *9*, 2104. https://doi.org/10.3390/su9112104.

79. Zhang, C., Wei, H., Zhao, J., Liu, T., Zhu, T., and Zhang, K. (2016). Short-term wind speed forecasting using empirical mode decomposition and feature selection. Renew. Energy *96*, 727–737. https://doi.org/10.1016/j.renene.2016.05.023.

80. Liu, H., Chen, C., Tian, H.-q., and Li, Y.-f. (2012). A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks. Renew. Energy *48*, 545–556. https://doi.org/10.1016/j.renene.2012.06.012.

81. Niu, D., Liang, Y., and Hong, W.-C. (2017). Wind speed forecasting based on EMD and GRNN optimized by FOA. Energies *10*, 2001. https://doi.org/10.3390/en10122001.

82. Naik, J., Satapathy, P., and Dash, P.K. (2018). Short-term wind speed and wind power prediction using hybrid empirical mode decomposition and kernel ridge regression. Appl. Soft Comput. *70*, 1167–1188. https://doi.org/10.1016/j.asoc.2017.12.010.

83. Fei, S.-w. (2016). A hybrid model of EMD and multiple-kernel RVR algorithm for wind speed prediction. Int. J. Electr. Power Energy Syst. *78*, 910–915. https://doi.org/10.1016/j.ijepes.2015.11.116.

84. Yang, Z., and Wang, J. (2018). A combination forecasting approach applied in multistep wind speed forecasting based on a data processing strategy and an optimized artificial intelligence algorithm. Appl. Energy *230*, 1108–1125. https://doi.org/10.1016/j.apenergy.2018.09.037.

85. Yang, Z., and Wang, J. (2018). A hybrid forecasting approach applied in wind speed forecasting based on a data processing strategy and an optimized artificial intelligence algorithm. Energy *160*, 87–100. https://doi.org/10.1016/j.energy.2018.07.005.

86. Liu, Z., Hara, R., and Kita, H. (2021). Hybrid forecasting system based on data area division and deep learning neural network for short-term wind speed forecasting. Energy Convers. Manag. *238*, 114136. https://doi.org/10.1016/j.enconman.2021.114136.

87. da Silva, R.G., Ribeiro, M.H.D.M., Moreno, S.R., Mariani, V.C., and Coelho, L.d.S. (2021). A novel decomposition-ensemble learning framework for multi-step ahead wind energy forecasting. Energy *216*, 119174. https://doi.org/10.1016/j.energy.2020.119174.

88. Wang, J., and Yang, Z. (2021). Ultra-short-term wind speed forecasting using an optimized artificial intelligence algorithm. Renew. Energy *171*, 1418–1435. https://doi.org/10.1016/j.renene.2021.03.020.

89. Huang, Y., Liu, S., and Yang, L. (2018). Wind speed forecasting method using EEMD and the combination forecasting method based on GPR and LSTM. Sustainability *10*, 3693. https://doi.org/10.3390/su10103693.

90. Wang, Y., Wang, S., and Zhang, N. (2013). A Novel Wind Speed Forecasting Method Based on Ensemble Empirical Mode Decomposition and GA-BP Neural Network (IEEE). https://doi.org/10.1109/PESMG.2013.6672195.

91. Wang, S., Zhang, N., Wu, L., and Wang, Y. (2016). Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method. Renew. Energy *94*, 629–636. https://doi.org/10.1016/j.renene.2016.03.103.

92. Bokde, N., Feijóo, A., and Kulat, K. (2018). Analysis of differencing and decomposition preprocessing methods for wind speed prediction. Appl. Soft Comput. *71*, 926–938. https://doi.org/10.1016/j.asoc.2018.07.041.

93. Sun, W., and Liu, M. (2016). Wind speed forecasting using FEEMD echo state networks with RELM in Hebei, China. Energy Convers. Manag. *114*, 197–208. https://doi.org/10.1016/j.enconman.2016.02.022.

94. Sun, W., and Wang, Y. (2018). Short-term wind speed forecasting based on fast ensemble empirical mode decomposition, phase space reconstruction, sample entropy and improved back-propagation neural network. Energy Convers. Manag. *157*, 1–12. https://doi.org/10.1016/j.enconman.2017.11.067.

95. Liu, H., Tian, H.-q., and Li, Y.-f. (2015). Comparison of new hybrid FEEMD-MLP, FEEMD-ANFIS, Wavelet Packet-MLP and Wavelet Packet-ANFIS for wind speed predictions. Energy Convers. Manag. *89*, 1–11. https://doi.org/10.1016/j.enconman.2014.09.060.

96. Wang, L., Li, X., and Bai, Y. (2018). Short-term wind speed prediction using an extreme learning machine model with error correction. Energy Convers. Manag. *162*, 239–250. https://doi.org/10.1016/j.enconman.2018.02.015.

97. Han, L., Jing, H., Zhang, R., and Gao, Z. (2019). Wind power forecast based on improved Long Short Term Memory network. Energy *189*, 116300. https://doi.org/10.1016/j.energy.2019.116300.

98. Wu, Q., and Lin, H. (2019). Short-term wind speed forecasting based on hybrid variational mode decomposition and least squares support vector machine optimized by bat algorithm model. Sustainability *11*, 652.

99. Abdoos, A.A. (2016). A new intelligent method based on combination of VMD and ELM for short term wind power forecasting. Neurocomputing *203*, 111–120. https://doi.org/10.1016/j.neucom.2016.03.054.

100. Moreno, S.R., Mariani, V.C., and Coelho, L.d.S. (2021). Hybrid multi-stage decomposition with parametric model applied to wind speed forecasting in Brazilian Northeast. Renew. Energy *164*, 1508–1526. https://doi.org/10.1016/j.renene.2020.10.126.

101. Duan, J., Wang, P., Ma, W., Fang, S., and Hou, Z. (2022). A novel hybrid model based on nonlinear weighted combination for short-term wind power forecasting. Int. J. Electr. Power Energy Syst. *134*, 107452. https://doi.org/10.1016/j.ijepes.2021.107452.

102. Tian, Z. (2020). Short-term wind speed prediction based on LMD and improved FA optimized combined kernel function LSSVM. Eng. Appl. Artif. Intell. *91*, 103573. https://doi.org/10.1016/j.engappai.2020.103573.

103. Yin, H., Dong, Z., Chen, Y., Ge, J., Lai, L.L., Vaccaro, A., and Meng, A. (2017). An effective secondary decomposition approach for wind power forecasting using extreme learning machine trained by crisscross optimization. Energy Convers. Manag. *150*, 108–121. https://doi.org/10.1016/j.enconman.2017.08.014.

104. Liu, H., Mi, X., and Li, Y. (2018). An experimental investigation of three new hybrid wind speed forecasting models using multi-decomposing strategy and ELM

algorithm. Renew. Energy *123*, 694–705. https://doi.org/10.1016/j.renene.2018.02.092.

105. Wang, D., Luo, H., Grunder, O., Lin, Y., and Guo, H. (2017). Multi-step ahead electricity price forecasting using a hybrid model based on two-layer decomposition technique and BP neural network optimized by firefly algorithm. Appl. Energy *190*, 390–407. https://doi.org/10.1016/j.apenergy.2016.12.134.

106. Wang, J., and Li, Y. (2017). Short-term wind speed prediction using signal preprocessing technique and evolutionary support vector regression. Neural Process. Lett. *48*, 1043–1061.

107. Liu, H., Duan, Z., Han, F.-z., and Li, Y.-f. (2018). Big multi-step wind speed forecasting model based on secondary decomposition, ensemble method and error correction algorithm. Energy Convers. Manag. *156*, 525–541. https://doi.org/10.1016/j.enconman.2017.11.049.

108. Du, P., Wang, J., Yang, W., and Niu, T. (2019). A novel hybrid model for short-term wind power forecasting. Appl. Soft Comput. *80*, 93–106. https://doi.org/10.1016/j.asoc.2019.03.035.

109. Liu, H., Tian, H.-q., Liang, X.-f., and Li, Y.-f. (2015). Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks. Appl. Energy *157*, 183–194. https://doi.org/10.1016/j.apenergy.2015.08.014.

110. Tian, Z. (2021). Modes decomposition forecasting approach for ultra-short-term wind speed. Appl. Soft Comput. *105*, 107303. https://doi.org/10.1016/j.asoc.2021.107303.

111. Zhang, D., Peng, X., Pan, K., and Liu, Y. (2019). A novel wind speed forecasting based on hybrid decomposition and online sequential outlier robust extreme learning machine. Energy Convers. Manag. *180*, 338–357. https://doi.org/10.1016/j.enconman.2018.10.089.

112. Liu, D., Niu, D., Wang, H., and Fan, L. (2014). Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm. Renew. Energy *62*, 592–597. https://doi.org/10.1016/j.renene.2013.08.011.

113. Adedeji, P.A., Akinlabi, S., Madushele, N., and Olatunji, O.O. (2020). Wind turbine power output very short-term forecast: a comparative study of data clustering techniques in a PSO-ANFIS model. J. Clean. Prod. *254*, 120135. https://doi.org/10.1016/j.jclepro.2020.120135.

114. Wang, Y., Liu, Y., Li, L., Infield, D., and Han, S. (2018). Short-term wind power forecasting based on clustering pre-calculated CFD method. Energies *11*, 854.

115. Zhang, Y., Pan, G., Chen, B., Han, J., Zhao, Y., and Zhang, C. (2020). Short-term wind speed prediction model based on GA-ANN improved by VMD. Renew. Energy *156*, 1373–1388. https://doi.org/10.1016/j.renene.2019.12.047.

116. He, Q., Wang, J., and Lu, H. (2018). A hybrid system for short-term wind speed forecasting. Appl. Energy *226*, 756–771. https://doi.org/10.1016/j.apenergy.2018.06.053.

117. Azimi, R., Ghofrani, M., and Ghayekhloo, M. (2016). A hybrid wind power forecasting model based on data mining and wavelets analysis. Energy Convers. Manag. *127*, 208–225. https://doi.org/10.1016/j.enconman.2016.09.002.

118. Liu, D., Wang, J., and Wang, H. (2015). Short-term wind speed forecasting based on spectral clustering and optimised echo state networks. Renew. Energy *78*, 599–608. https://doi.org/10.1016/j.renene.2015.01.022.

119. Sun, S., Qiao, H., Wei, Y., and Wang, S. (2017). A new dynamic integrated approach for wind speed forecasting. Appl. Energy *197*, 151–162. https://doi.org/10.1016/j.apenergy.2017.04.008.

120. Santamaría-Bonfil, G., Reyes-Ballesteros, A., and Gershenson, C. (2016). Wind speed forecasting for wind farms: a method based on support vector regression. Renew. Energy *85*, 790–809. https://doi.org/10.1016/j.renene.2015.07.004.

121. Hu, Q., Su, P., Yu, D., and Liu, J. (2014). Pattern-based wind speed prediction based on generalized principal component analysis. IEEE Trans. Sustain. Energy *5*, 866–874. https://doi.org/10.1109/TSTE.2013.2295402.

122. Kong, X., Liu, X., Shi, R., and Lee, K.Y. (2015). Wind speed prediction using reduced support vector machines with feature selection. Neurocomputing *169*, 449–456. https://doi.org/10.1016/j.neucom.2014.09.090.

123. Jiang, P., and Li, C. (2018). Research and application of an innovative combined model based on a modified optimization algorithm for wind speed forecasting. Measurement *124*, 395–412. https://doi.org/10.1016/j.measurement.2018.04.014.

124. Liu, H., Mi, X., and Li, Y. (2018). Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, LSTM network and ELM. Energy Convers. Manag. *159*, 54–64. https://doi.org/10.1016/j.enconman.2018.01.010.

125. Wang, J., Heng, J., Xiao, L., and Wang, C. (2017). Research and application of a combined model based on multi-objective optimization for multi-step ahead wind speed forecasting. Energy *125*, 591–613. https://doi.org/10.1016/j.energy.2017.02.150.

126. Rodrigues Moreno, S., Gomes da Silva, R., Cocco Mariani, V., and dos Santos Coelho, L. (2020). Multi-step wind speed forecasting based on hybrid multi-stage decomposition model and long short-term memory neural network. Energy Convers. Manag. *213*, 112869. https://doi.org/10.1016/j.enconman.2020.112869.

127. Cheng, L., Zang, H., Ding, T., Sun, R., Wang, M., Wei, Z., and Sun, G. (2018). Ensemble recurrent neural network based probabilistic wind speed forecasting approach. Energies *11*, 1958. https://doi.org/10.3390/en11081958.

128. Guo, Z., Zhao, W., Lu, H., and Wang, J. (2012). Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model. Renew. Energy *37*, 241–249. https://doi.org/10.1016/j.renene.2011.06.023.

129. Niu, X., and Wang, J. (2019). A combined model based on data preprocessing strategy and multi-objective optimization algorithm for short-term wind speed forecasting. Appl. Energy *241*, 519–539. https://doi.org/10.1016/j.apenergy.2019.03.097.

130. Santhosh, M., Venkaiah, C., and Vinod Kumar, D.M. (2018). Ensemble empirical mode decomposition based adaptive wavelet neural network method for wind speed prediction. Energy Convers. Manag. *168*, 482–493. https://doi.org/10.1016/j.enconman.2018.04.099.

131. Bounoua, Z., Ouazzani Chahidi, L., and Mechaqrane, A. (2021). Estimation of daily global solar radiation using empirical and machine-learning methods: a case study of five Moroccan locations. Sustain. Mater. Technol. *28*, e00261. https://doi.org/10.1016/j.susmat.2021.e00261.

132. Rushdi, M.A., Yoshida, S., Watanabe, K., and Ohya, Y. (2021). Machine learning approaches for thermal updraft prediction in wind solar tower systems. Renew. Energy *177*, 1001–1013. https://doi.org/10.1016/j.renene.2021.06.033.

133. Wang, J., Du, P., Niu, T., and Yang, W. (2017). A novel hybrid system based on a new proposed algorithm-Multi-Objective Whale Optimization Algorithm for wind speed forecasting. Appl. Energy *208*, 344–360. https://doi.org/10.1016/j.apenergy.2017.10.031.

134. Ding, J., Chen, G., and Yuan, K. (2020). Short-term wind power prediction based on improved grey wolf optimization algorithm for extreme learning machine. Processes *8*, 109. https://doi.org/10.3390/pr8010109.

135. Janković, R., Mihajlović, I., Štrbac, N., and Amelio, A. (2021). Machine learning models for ecological footprint prediction based on energy parameters. Neural Comput. Appl. *33*, 7073–7087. https://doi.org/10.1007/s00521-020-05476-4.

136. Gupta, D., Natarajan, N., and Berlin, M. (2022). Short-term wind speed prediction using hybrid machine learning techniques. Environ. Sci. Pollut. Res. Int. *29*, 50909–50927. https://doi.org/10.1007/s11356-021-15221-6.

137. Tian, Z. (2021). A state-of-the-art review on wind power deterministic prediction. Wind Eng. *45*, 1374–1392.

138. Fan, H., Zhang, X., Mei, S., Chen, K., and Chen, X. (2020). M2gsnet: multi-modal multi-task graph spatiotemporal network for ultra-short-term wind farm cluster power prediction. Appl. Sci. *10*, 7915.

139. Siebert, N. (2008). Development of Methods for Regional Wind Power Forecasting.

140. Huang, N., Wu, Y., Cai, G., Zhu, H., Yu, C., Jiang, L., Zhang, Y., Zhang, J., and Xing, E. (2019). Short-term wind speed forecast with low loss of information based on feature generation of OSVD. IEEE Access *7*, 81027–81046. https://doi.org/10.1109/ACCESS.2019.2922662.

141. Lazić, L., Pejanović, G., and Živković, M. (2010). Wind forecasts for wind power generation using the Eta model. Renew. Energy *35*, 1236–1243. https://doi.org/10.1016/j.renene.2009.10.028.

142. Landberg, L., Giebel, G., Nielsen, H.A., Nielsen, T., and Madsen, H. (2003). Short-term prediction—an overview. Wind Energy 6, 273–280.

143. Chen, N., Qian, Z., Nabney, I.T., and Meng, X. (2014). Wind power forecasts using Gaussian processes and numerical weather prediction. IEEE Trans. Power Syst. *29*, 656–665. https://doi.org/10.1109/TPWRS.2013.2282366.

144. Giebel, G. (2003). State-of-the-Art on Methods and Software Tools for Short-Term Prediction of Wind Energy Production.

145. Xiang, L., Deng, Z., and Hu, A. (2019). Forecasting short-term wind speed based on IEWT-LSSVM model optimized by bird swarm algorithm. IEEE Access *7*, 59333–59345. https://doi.org/10.1109/ACCESS.2019.2914251.

146. Alexiadis, M.C., Dokopoulos, P.S., and Sahsamanoglou, H.S. (1999). Wind speed and power forecasting based on spatial correlation models. IEEE Trans. Energy Conver. *14*, 836–842. https://doi.org/10.1109/60.790962.

147. Sahin, A.D., and Sen, Z. (2000). Wind energy directional spatial correlation functions and application for prediction. Wind Eng. *24*, 223–231. https://doi.org/10.1260/0309524001495576.

148. Yong-ning, Z. (2012). Spatial Model for Short Term Wind Power Prediction Considering Wake Effects (Power System Protection and Control).

149. Damousis, I.G., Alexiadis, M.C., Theocharis, J.B., and Dokopoulos, P.S. (2004). A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation. IEEE Trans. Energy Convers. *19*, 352–361. https://doi.org/10.1109/TEC.2003.821865.

150. Barbounis, T.G., and Theocharis, J.B. (2007). Locally recurrent neural networks for wind speed prediction using spatial correlation. Inform. Sci. *177*, 5775–5797. https://doi.org/10.1016/j.ins.2007.05.024.

151. Sharifian, A., Ghadi, M.J., Ghavidel, S., Li, L., and Zhang, J. (2018). A new method based on Type-2 fuzzy neural network for accurate wind power forecasting under uncertain data. Renew. Energy *120*, 220–230. https://doi.org/10.1016/j.renene.2017.12.023.

152. Ouyang, T., Zha, X., and Qin, L. (2013). A survey of wind power ramp forecasting. Energy Power Eng. *5*, 368–372. https://doi.org/10.4236/epe.2013.54B071.

153. Torres, J.L., García, A., De Blas, M., and De Francisco, A. (2005). Forecast of hourly average wind speed with ARMA models in Navarre (Spain). Sol. Energy *79*, 65–77. https://doi.org/10.1016/j.solener.2004.09.013.

154. Kavasseri, R.G., and Seetharaman, K. (2009). Day-ahead wind speed forecasting using f-ARIMA models. Renew. Energy *34*, 1388–1393. https://doi.org/10.1016/j.renene.2008.09.006.

155. Erdem, E., and Shi, J. (2011). ARMA based approaches for forecasting the tuple of wind speed and direction. Appl. Energy *88*, 1405–1414. https://doi.org/10.1016/j.apenergy.2010.10.031.

156. Guo, Z., Zhao, J., Zhang, W., and Wang, J. (2011). A corrected hybrid approach for wind speed prediction in Hexi Corridor of China. Energy *36*, 1668–1679. https://doi.org/10.1016/j.energy.2010.12.063.

157. Shi, J., Qu, X., and Zeng, S. (2011). Short-term wind power generation forecasting: direct versus indirect arima-based approaches. Int. J. Green Energy *8*, 100–112. https://doi.org/10.1080/15435075.2011.546755.

158. Lydia, M., Suresh Kumar, S., Immanuel Selvakumar, A., and Edwin Prem Kumar, G. (2016). Linear and non-linear autoregressive models for short-term wind speed forecasting. Energy Convers. Manag. *112*, 115–124. https://doi.org/10.1016/j.enconman.2016.01.007.

159. Karakuş, O., Kuruoğlu, E.E., and Altınkaya, M.A. (2017). One-day ahead wind speed/power prediction based on polynomial autoregressive model. IET Renew. Power Gener. *11*, 1430–1439. https://doi.org/10.1049/iet-rpg.2016.0972.

160. Malmberg, A., Holst, U., and Holst, J. (2005). Forecasting near-surface ocean winds with Kalman filter techniques. Ocean Eng. *32*, 273–291. https://doi.org/10.1016/j.oceaneng.2004.08.005.

161. Liu, H., Tian, H.-q., and Li, Y.-f. (2012). Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction. Appl. Energy *98*, 415–424. https://doi.org/10.1016/j.apenergy.2012.04.001.

162. Poncela, M., Poncela, P., and Perán, J.R. (2013). Automatic tuning of Kalman filters by maximum likelihood methods for wind energy forecasting. Appl. Energy *108*, 349–362. https://doi.org/10.1016/j.apenergy.2013.03.041.

163. Chen, K., and Yu, J. (2014). Short-term wind speed prediction using an unscented Kalman filter based state-space support vector regression approach. Appl. Energy *113*, 690–705. https://doi.org/10.1016/j.apenergy.2013.08.025.

164. Zuluaga, C.D., Álvarez, M.A., and Giraldo, E. (2015). Short-term wind speed prediction based on robust Kalman filtering: an experimental comparison. Appl. Energy *156*, 321–330. https://doi.org/10.1016/j.apenergy.2015.07.043.

165. Crochet, P. (2004). Adaptive Kalman filtering of 2-metre temperature and 10-metre wind-speed forecasts in Iceland. Meteorol. Appl. *11*, 173–187. https://doi.org/10.1017/S1350482704001252.

166. Louka, P., Galanis, G., Siebert, N., Kariniotakis, G., Katsafados, P., Pytharoulis, I., and Kallos, G. (2008). Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering. J. Wind Eng. Ind. Aerod. *96*, 2348–2362. https://doi.org/10.1016/j.jweia.2008.03.013.

167. Cassola, F., and Burlando, M. (2012). Wind speed and wind energy forecast through Kalman filtering of Numerical Weather Prediction model output. Appl. Energy *99*, 154–166. https://doi.org/10.1016/j.apenergy.2012.03.054.

168. Williams, J.L., III, Maxwell, R.M., and Monache, L.D. (2013). Development and verification of a new wind speed forecasting system using an ensemble Kalman filter data assimilation technique in a fully coupled hydrologic and atmospheric model. J. Adv. Model. Earth Syst. *5*, 785–800. https://doi.org/10.1002/jame.20051.

169. Stathopoulos, C., Kaperoni, A., Galanis, G., and Kallos, G. (2013). Wind power prediction based on numerical and statistical models. J. Wind Eng. Ind. Aerod. *112*, 25–38. https://doi.org/10.1016/j.jweia.2012.09.004.

170. Foley, A.M., Leahy, P.G., Marvuglia, A., and McKeogh, E.J. (2012). Current methods and advances in forecasting of wind power generation. Renew. Energy *37*, 1–8. https://doi.org/10.1016/j.renene.2011.05.033.

171. Cadenas, E., and Rivera, W. (2009). Short term wind speed forecasting in La Venta, Oaxaca, México, using artificial neural networks. Renew. Energy *34*, 274–278. https://doi.org/10.1016/j.renene.2008.03.014.

172. Fadare, D.A. (2010). The application of artificial neural networks to mapping of wind speed profile for energy application in Nigeria. Appl. Energy *87*, 934–942. https://doi.org/10.1016/j.apenergy.2009.09.005.

173. Li, L.-L., Liu, Z.-F., Tseng, M.-L., Jantarakolica, K., and Lim, M.K. (2021). Using enhanced crow search algorithm optimization-extreme learning machine model to forecast short-term wind power. Expert Syst. Appl. *184*, 115579. https://doi.org/10.1016/j.eswa.2021.115579.

174. Hu, H., Li, Y., Zhang, X., and Fang, M. (2022). A novel hybrid model for short-term

prediction of wind speed. Pattern Recogn. *127*, 108623. https://doi.org/10.1016/j.patcog.2022.108623.

175. Monfared, M., Rastegar, H., and Kojabadi, H.M. (2009). A new strategy for wind speed forecasting using artificial intelligent methods. Renew. Energy *34*, 845–848. https://doi.org/10.1016/j.renene.2008.04.017.

176. Mohandes, M., Rehman, S., and Rahman, S.M. (2011). Estimation of wind speed profile using adaptive neuro-fuzzy inference system (ANFIS). Appl. Energy *88*, 4024–4032. https://doi.org/10.1016/j.apenergy.2011.04.015.

177. Osório, G., Matias, J.C.O., and Catalão, J. (2015). Short-term wind power forecasting using adaptive neuro-fuzzy inference system combined with evolutionary particle swarm optimization, wavelet transform and mutual information. Renew. Energy *75*, 301–307. https://doi.org/10.1016/j.renene.2014.09.058.

178. Ma, X., Jin, Y., and Dong, Q. (2017). A generalized dynamic fuzzy neural network based on singular spectrum analysis optimized by brain storm optimization for short-term wind speed forecasting. Appl. Soft Comput. *54*, 296–312. https://doi.org/10.1016/j.asoc.2017.01.033.

179. Bengio, Y. (2009). Learning Deep Architectures for AI. Foundations and Trends® in Machine Learning, *2*, pp. 1–127. https://doi.org/10.1561/2200000006.

180. Dalto, M., Matuško, J., and Vašak, M. (2015). Deep Neural Networks for Ultra-short-term Wind Forecasting (IEEE), pp. 1657–1663.

181. Liu, J., Shi, Q., Han, R., and Yang, J. (2021). A hybrid GA–PSO–CNN model for ultra-short-term wind power forecasting. Energies *14*, 6500. https://doi.org/10.3390/en14206500.

182. Santhosh, M., Venkaiah, C., and Kumar, D.V. (2019). Short-term wind speed forecasting approach using ensemble empirical mode decomposition and deep Boltzmann machine. Sustain. Energy Grids Netw. *19*, 100242. https://doi.org/10.1016/j.segan.2019.100242.

183. Khodayar, M., Wang, J., and Manthouri, M. (2019). Interval deep generative neural network for wind speed forecasting. IEEE Trans. Smart Grid *10*, 3974–3989. https://doi.org/10.1109/TSG.2018.2847223.

184. Jiajun, H., Chuanjin, Y., Yongle, L., and Huoyue, X. (2020). Ultra-short term wind prediction with wavelet transform, deep belief network and ensemble learning. Energy Convers. Manag. *205*, 112418. https://doi.org/10.1016/j.enconman.2019.112418.

185. El Bourakadi, D., Yahyaouy, A., and Boumhidi, J. (2022). Improved extreme learning machine with AutoEncoder and particle swarm optimization for short-term wind power prediction. Neural Comput. Appl. *34*, 4643–4659. https://doi.org/10.1007/s00521-021-06619-x.

186. Ji, L., Fu, C., Ju, Z., Shi, Y., Wu, S., and Tao, L. (2022). Short-term canyon wind speed prediction based on CNN—GRU transfer learning. Atmosphere *13*. https://doi.org/10.3390/atmos13050813.

187. Hu, Q., Zhang, R., and Zhou, Y. (2016). Transfer learning for short-term wind speed prediction with deep neural networks. Renew. Energy *85*, 83–95. https://doi.org/10.1016/j.renene.2015.06.034.

188. Liu, X., Cao, Z., and Zhang, Z. (2021). Short-term predictions of multiple wind turbine power outputs based on deep neural networks with transfer learning. Energy *217*, 119356. https://doi.org/10.1016/j.energy.2020.119356.

189. Li, G., and Shi, J. (2010). On comparing three artificial neural networks for wind speed forecasting. Appl. Energy *87*, 2313–2320. https://doi.org/10.1016/j.apenergy.2009.12.013.

190. Ren, C., An, N., Wang, J., Li, L., Hu, B., and Shang, D. (2014). Optimal parameters selection for BP neural network based on particle swarm optimization: a case study of wind speed forecasting. Knowl. Base. Syst. *56*, 226–239. https://doi.org/10.1016/j.knosys.2013.11.015.

191. Chitsazan, M.A., Sami Fadali, M., and Trzynadlowski, A.M. (2019). Wind speed and wind direction forecasting using echo state network with nonlinear functions. Renew. Energy *131*, 879–889. https://doi.org/10.1016/j.renene.2018.07.060.

192. Xiao, L., Qian, F., and Shao, W. (2017). Multi-step wind speed forecasting based on a hybrid forecasting architecture and an improved bat algorithm. Energy Convers. Manag. *143*, 410–430. https://doi.org/10.1016/j.enconman.2017.04.012.

193. Mohandes, M.A., Rehman, S., and Halawani, T.O. (1998). A neural networks approach for wind speed prediction. Renew. Energy *13*, 345–354. https://doi.org/10.1016/S0960-1481(98)00001-9.

194. Navas, R.K.B., and Prakash, S. (2021). A novel ultra-short term wind power forecasting intelligence system based on hybrid neural network. Mater. Today Proc. *47*, 1145–1148. https://doi.org/10.1016/j.matpr.2021.07.336.

195. Wang, H., Sun, J., Sun, J., and Wang, J. (2017). Using random forests to select optimal input variables for short-term wind speed forecasting models. Energies *10*, 1522. https://doi.org/10.3390/en10101522.

196. Liu, H., Duan, Z., Li, Y., and Lu, H. (2018). A novel ensemble model of different mother wavelets for wind speed multi-step forecasting. Appl. Energy *228*, 1783–1800. https://doi.org/10.1016/j.apenergy.2018.07.050.

197. Zhou, J., Sun, N., Jia, B., and Peng, T. (2018). A novel decomposition-optimization model for short-term wind speed forecasting. Energies *11*, 1752. https://doi.org/10.3390/en11071752.

198. Jiang, P., Wang, Y., and Wang, J. (2017). Short-term wind speed forecasting using a hybrid model. Energy *119*, 561–577. https://doi.org/10.1016/j.energy.2016.10.040.

199. Salcedo-Sanz, S., Ortiz-Garcı'a, E.G., Pérez-Bellido, Á.M., Portilla-Figueras, A., and Prieto, L. (2011). Short term wind speed prediction based on evolutionary support vector regression algorithms. Expert Syst. Appl. *38*, 4052–4057. https://doi.org/10.1016/j.eswa.2010.09.067.

200. Khodayar, M., Kaynak, O., and Khodayar, M.E. (2017). Rough deep neural architecture for short-term wind speed forecasting. IEEE Trans. Ind. Inform. *13*, 2770–2779. https://doi.org/10.1109/TII.2017.2730846.

201. Sergio, A.T., and Ludermir, T.B. (2015). Deep Learning for Wind Speed Forecasting in Northeastern Region of Brazil (IEEE), pp. 322–327.

202. Wang, S., Wang, L., Hu, H., and Zeng, Y.-R. (2021). Effective wind power prediction using novel deep learning network: stacked independently recurrent autoencoder. Renew. Energy *164*, 642–655. https://doi.org/10.1016/j.renene.2020.09.108.

203. Qureshi, A.S., Khan, A., Zameer, A., and Usman, A. (2017). Wind power prediction using deep neural network based meta regression and transfer learning. Appl. Soft Comput. *58*, 742–755. https://doi.org/10.1016/j.asoc.2017.05.031.

204. Liu, H., Mi, X., and Li, Y. (2018). Smart deep learning based wind speed prediction model using wavelet packet decomposition, convolutional neural network and convolutional long short term memory network. Energy Convers. Manag. *166*, 120–131. https://doi.org/10.1016/j.enconman.2018.04.021.

205. Mi, X., Liu, H., and Li, Y. (2019). Wind speed prediction model using singular spectrum analysis, empirical mode decomposition and convolutional support vector machine. Energy Convers. Manag. *180*, 196–205. https://doi.org/10.1016/j.enconman.2018.11.006.

206. Ding, M., Zhou, H., Xie, H., Wu, M., Nakanishi, Y., and Yokoyama, R. (2019). A gated recurrent unit neural networks based wind speed error correction model for short-term wind power forecasting. Neurocomputing *365*, 54–61. https://doi.org/10.1016/j.neucom.2019.07.058.

207. Kisvari, A., Lin, Z., and Liu, X. (2021). Wind power forecasting – a data-driven method along with gated recurrent neural network. Renew. Energy *163*, 1895–1909. https://doi.org/10.1016/j.renene.2020.10.119.

208. Yu, C., Li, Y., Bao, Y., Tang, H., and Zhai, G. (2018). A novel framework for wind speed prediction based on recurrent neural networks and support vector machine. Energy Convers. Manag. *178*, 137–145. https://doi.org/10.1016/j.enconman.2018.10.008.

209. Zhang, Z., Qin, H., Liu, Y., Wang, Y., Yao, L., Li, Q., Li, J., and Pei, S. (2019). Long Short-Term

Memory Network based on Neighborhood Gates for processing complex causality in wind speed prediction. Energy Convers. Manag. *192*, 37–51. https://doi.org/10.1016/j.enconman.2019.04.006.

210. Qin, Y., Li, K., Liang, Z., Lee, B., Zhang, F., Gu, Y., Zhang, L., Wu, F., and Rodriguez, D. (2019). Hybrid forecasting model based on long short term memory network and deep learning neural network for wind signal. Appl. Energy *236*, 262–272. https://doi.org/10.1016/j.apenergy.2018.11.063.

211. Lopez, E., Valle, C., Allende, H., and Gil, E. (2017). Efficient training over long short-term memory networks for wind speed forecasting. Progress Pattern Recogn. Image Anal. Comput. VisionAppl. 409–416.

212. Yu, R., Gao, J., Yu, M., Lu, W., Xu, T., Zhao, M., Zhang, J., Zhang, R., and Zhang, Z. (2019). LSTM-EFG for wind power forecasting based on sequential correlation features. Future Generat. Comput. Syst. *93*, 33–42. https://doi.org/10.1016/j.future.2018.09.054.

213. Zhang, Z., Qin, H., Liu, Y., Yao, L., Yu, X., Lu, J., Jiang, Z., and Feng, Z. (2019). Wind speed forecasting based on quantile regression minimal gated memory network and kernel density estimation. Energy Convers. Manag. *196*, 1395–1409. https://doi.org/10.1016/j.enconman.2019.06.024.

214. Duan, J., Zuo, H., Bai, Y., Duan, J., Chang, M., and Chen, B. (2021). Short-term wind speed forecasting using recurrent neural networks with error correction. Energy *217*, 119397. https://doi.org/10.1016/j.energy.2020.119397.

215. Luo, X., Sun, J., Wang, L., Wang, W., Zhao, W., Wu, J., Wang, J.H., and Zhang, Z. (2018). Short-term wind speed forecasting via stacked extreme learning machine with generalized correntropy. IEEE Trans. Ind. Inform. *14*, 4963–4971. https://doi.org/10.1109/TII.2018.2854549.

216. Fei, S.-w., and He, Y. (2015). Wind speed prediction using the hybrid model of wavelet decomposition and artificial bee colony algorithm-based relevance vector machine. Int. J. Electr. Power Energy Syst. *73*, 625–631. https://doi.org/10.1016/j.ijepes.2015.04.019.

217. Jia, G., Li, D., Yao, L., and Zhao, P. (2016). An Improved Artificial Bee Colony-BP Neural Network Algorithm in the Short-Term Wind Speed Prediction (IEEE), pp. 2252–2255.

218. Li, L.-L., Chang, Y.-B., Tseng, M.-L., Liu, J.-Q., and Lim, M.K. (2020). Wind power prediction using a novel model on wavelet decomposition-support vector machines-improved atomic search algorithm. J. Clean. Prod. *270*, 121817. https://doi.org/10.1016/j.jclepro.2020.121817.

219. Dong, Y., Li, J., Liu, Z., Niu, X., and Wang, J. (2022). Ensemble wind speed forecasting system based on optimal model adaptive selection strategy: case study in China. Sustain. Energy Technol. Assess. *53*, 102535. https://doi.org/10.1016/j.seta.2022.102535.

220. Wu, Y.K., Su, P.E., and Hong, J.S. (2016). Stratification-based wind power forecasting in a high-penetration wind power system using a hybrid model. IEEE Trans. Ind. Appl. *52*, 2016–2030. https://doi.org/10.1109/TIA.2016.2524439.

221. Chitsaz, H., Amjady, N., and Zareipour, H. (2015). Wind power forecast using wavelet neural network trained by improved Clonal selection algorithm. Energy Convers. Manag. *89*, 588–598. https://doi.org/10.1016/j.enconman.2014.10.001.

222. Salcedo-Sanz, S., Pastor-Sánchez, A., Prieto, L., Blanco-Aguilera, A., and García-Herrera, R. (2014). Feature selection in wind speed prediction systems based on a hybrid coral reefs optimization – extreme learning machine approach. Energy Convers. Manag. *87*, 10–18. https://doi.org/10.1016/j.enconman.2014.06.041.

223. Salcedo-Sanz, S., Pastor-Sánchez, A., Del Ser, J., Prieto, L., and Geem, Z.W. (2015). A Coral Reefs Optimization algorithm with Harmony Search operators for accurate wind speed prediction. Renew. Energy *75*, 93–101. https://doi.org/10.1016/j.renene.2014.09.027.

224. Zhao, J., Guo, Z., Guo, Y., Lin, W., and Zhu, W. (2021). A self-organizing forecast of day-ahead wind speed: selective ensemble strategy based on numerical weather predictions. Energy *218*, 119509. https://doi.org/10.1016/j.energy.2020.119509.

225. Li, L.-l., Cen, Z.-Y., Tseng, M.-L., Shen, Q., and Ali, M.H. (2021). Improving short-term wind power prediction using hybrid improved cuckoo search arithmetic - support vector regression machine. J. Clean. Prod. *279*, 123739. https://doi.org/10.1016/j.jclepro.2020.123739.

226. Bo, H., Niu, X., and Wang, J. (2019). Wind speed forecasting system based on the variational mode decomposition strategy and immune selection multi-objective dragonfly optimization algorithm. IEEE Access *7*, 178063–178081. https://doi.org/10.1109/ACCESS.2019.2957062.

227. Chen, J., Zeng, G.-Q., Zhou, W., Du, W., and Lu, K.-D. (2018). Wind speed forecasting using nonlinear-learning ensemble of deep learning time series prediction and extremal optimization. Energy Convers. Manag. *165*, 681–695. https://doi.org/10.1016/j.enconman.2018.03.098.

228. Deo, R.C., Ghorbani, M.A., Samadianfard, S., Maraseni, T., Bilgili, M., and Biazar, M. (2018). Multi-layer perceptron hybrid model integrated with the firefly optimizer algorithm for windspeed prediction of target site using a limited set of neighboring reference station data. Renew. Energy *116*, 309–323. https://doi.org/10.1016/j.renene.2017.09.078.

229. Chen, Y., Dong, Z., Wang, Y., Su, J., Han, Z., Zhou, D., Zhang, K., Zhao, Y., and Bao, Y. (2021). Short-term wind speed predicting framework based on EEMD-GA-LSTM method under large scaled wind history. Energy Convers. Manag. *227*, 113559. https://doi.org/10.1016/j.enconman.2020.113559.

230. Fu, W., Wang, K., Li, C., and Tan, J. (2019). Multi-step short-term wind speed forecasting approach based on multi-scale dominant ingredient chaotic analysis, improved hybrid GWO-SCA optimization and ELM. Energy Convers. Manag. *187*, 356–377. https://doi.org/10.1016/j.enconman.2019.02.086.

231. Lu, H., Ma, X., Huang, K., and Azimi, M. (2020). Prediction of offshore wind farm power using a novel two-stage model combining kernel-based nonlinear extension of the Arps decline model with a multi-objective grey wolf optimizer. Renew. Sustain. Energy Rev. *127*, 109856. https://doi.org/10.1016/j.rser.2020.109856.

232. Han, Y., and Tong, X. (2020). Multi-step short-term wind power prediction based on three-level decomposition and improved grey wolf optimization. IEEE Access *8*, 67124–67136. https://doi.org/10.1109/ACCESS.2020.2984851.

233. Wang, C., Zhang, S., Liao, P., and Fu, T. (2022). Wind speed forecasting based on hybrid model with model selection and wind energy conversion. Renew. Energy *196*, 763–781. https://doi.org/10.1016/j.renene.2022.06.143.

234. Fu, W., Zhang, K., Wang, K., Wen, B., Fang, P., and Zou, F. (2021). A hybrid approach for multi-step wind speed forecasting based on two-layer decomposition, improved hybrid DE-HHO optimization and KELM. Renew. Energy *164*, 211–229. https://doi.org/10.1016/j.renene.2020.09.078.

235. Guo, H., Wang, J., Li, Z., and Jin, Y. (2022). A multivariable hybrid prediction system of wind power based on outlier test and innovative multi-objective optimization. Energy *239*, 122333. https://doi.org/10.1016/j.energy.2021.122333.

236. Al-qaness, M.A., Ewees, A.A., Fan, H., Abualigah, L., and Elaziz, M.A. (2022). Boosted ANFIS model using augmented marine predator algorithm with mutation operators for wind power forecasting. Appl. Energy *314*, 118851. https://doi.org/10.1016/j.apenergy.2022.118851.

237. Liu, Z., Jiang, P., Wang, J., and Zhang, L. (2021). Ensemble forecasting system for short-term wind speed forecasting based on optimal sub-model selection and multi-objective version of mayfly optimization algorithm. Expert Syst. Appl. *177*, 114974. https://doi.org/10.1016/j.eswa.2021.114974.

238. Nie, Y., Liang, N., and Wang, J. (2021). Ultra-short-term wind-speed bi-forecasting system via artificial intelligence and a double-forecasting scheme. Appl. Energy *301*, 117452. https://doi.org/10.1016/j.apenergy.2021.117452.

239. Jursa, R., and Rohrig, K. (2008). Short-term wind power forecasting using evolutionary algorithms for the automated specification of artificial intelligence models. Int. J.

Forecast. *24*, 694–709. https://doi.org/10.1016/j.ijforecast.2008.08.007.

240. Pousinho, H.M.I., Mendes, V.M.F., and Catalão, J. (2011). A hybrid PSO–ANFIS approach for short-term wind power prediction in Portugal. Energy Convers. Manag. *52*, 397–402. https://doi.org/10.1016/j.enconman.2010.07.015.

241. Zhang, J. (2014). Wind speed forecasting based on least squares support vector machine and particle swarm optimization. Appl. Mech. Mater. *602–605*, 3251–3255. https://doi.org/10.4028/www.scientific.net/AMM.602-605.3251.

242. He, Z., Chen, Y., Shang, Z., Li, C., Li, L., and Xu, M. (2019). A novel wind speed forecasting model based on moving window and multi-objective particle swarm optimization algorithm. Appl. Math. Model. *76*, 717–740. https://doi.org/10.1016/j.apm.2019.07.001.

243. Cheng, Z., and Wang, J. (2020). A new combined model based on multi-objective salp swarm optimization for wind speed forecasting. Appl. Soft Comput. *92*, 106294. https://doi.org/10.1016/j.asoc.2020.106294.

244. Han, S., Li, J., and Liu, Y. (2011). Tabu search algorithm optimized ANN model for wind power prediction with NWP. Energy Proc. *12*, 733–740. https://doi.org/10.1016/j.egypro.2011.10.099.

245. Samadianfard, S., Hashemi, S., Kargar, K., Izadyar, M., Mostafaeipour, A., Mosavi, A., Nabipour, N., and Shamshirband, S. (2020). Wind speed prediction using a hybrid model of the multi-layer perceptron and whale optimization algorithm. Energy Rep. *6*, 1147–1159. https://doi.org/10.1016/j.egyr.2020.05.001.

246. Cadenas, E., and Rivera, W. (2010). Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA–ANN model. Renew. Energy *35*, 2732–2738. https://doi.org/10.1016/j.renene.2010.04.022.

247. Yu, C., Li, Y., and Zhang, M. (2017). Comparative study on three new hybrid models using Elman Neural Network and Empirical Mode Decomposition based technologies improved by Singular Spectrum Analysis for hour-ahead wind speed forecasting. Energy Convers. Manag. *147*, 75–85. https://doi.org/10.1016/j.enconman.2017.05.008.

248. Zhang, G., and Liu, D. (2020). Causal convolutional gated recurrent unit network with multiple decomposition methods for short-term wind speed forecasting. Energy Convers. Manag. *226*, 113500. https://doi.org/10.1016/j.enconman.2020.113500.

249. Zhang, Y., Liu, K., Qin, L., and An, X. (2016). Deterministic and probabilistic interval prediction for short-term wind power generation based on variational mode decomposition and machine learning methods. Energy Convers. Manag. *112*, 208–219. https://doi.org/10.1016/j.enconman.2016.01.023.

250. Alencar, D.B., Affonso, C.M., Oliveira, R.C.L., and Filho, J.C.R. (2018). Hybrid approach combining SARIMA and neural networks for multi-step ahead wind speed forecasting in Brazil. IEEE Access *6*, 55986–55994. https://doi.org/10.1109/ACCESS.2018.2872720.

251. Fu, W., Wang, K., Tan, J., and Zhang, K. (2020). A composite framework coupling multiple feature selection, compound prediction models and novel hybrid swarm optimizer-based synchronization optimization strategy for multi-step ahead short-term wind speed forecasting. Energy Convers. Manag. *205*, 112461. https://doi.org/10.1016/j.enconman.2019.112461.

252. Gilbert, C., Browell, J., and McMillan, D. (2020). Leveraging turbine-level data for improved probabilistic wind power forecasting. IEEE Trans. Sustain. Energy *11*, 1152–1160.

253. Peng, X., Xu, Q., Wang, H., Lang, J., Li, W., Cai, T., Duan, S., Xie, Y., and Li, C. (2021). A novel efficient DLUBE model constructed by error interval coefficients for clustered wind power prediction. IEEE Access *9*, 61739–61751.

254. Yan, J., Li, K., Bai, E.W., Deng, J., and Foley, A.M. (2016). Hybrid probabilistic wind power forecasting using temporally local Gaussian process. IEEE Trans. Sustain. Energy *7*, 87–95. https://doi.org/10.1109/TSTE.2015.2472963.

255. Wan, C., Lin, J., Wang, J., Song, Y., and Dong, Z.Y. (2017). Direct quantile regression for nonparametric probabilistic forecasting of wind power generation. IEEE Trans. Power Syst. *32*, 2767–2778. https://doi.org/10.1109/TPWRS.2016.2625101.

256. Hu, M., Hu, Z., Yue, J., Zhang, M., and Hu, M. (2017). A novel multi-objective optimal approach for wind power interval prediction. Energies *10*, 419. https://doi.org/10.3390/en10040419.

257. Khosravi, A., Nahavandi, S., and Creighton, D. (2013). Prediction intervals for short-term wind farm power generation forecasts. IEEE Trans. Sustain. Energy *4*, 602–610. https://doi.org/10.1109/TSTE.2012.2232944.

258. Ye, L., Zhou, J., Gupta, H.V., Zhang, H., Zeng, X., and Chen, L. (2016). Efficient estimation of flood forecast prediction intervals via single-and multi-objective versions of the LUBE method. Hydrol. Process. *30*, 2703–2716.

259. Quan, H., Srinivasan, D., and Khosravi, A. (2014). Particle swarm optimization for construction of neural network-based prediction intervals. Neurocomputing *127*, 172–180. https://doi.org/10.1016/j.neucom.2013.08.020.

260. Gendeel, M., Zhang, Y., Qian, X., and Xing, Z. (2021). Deterministic and probabilistic interval prediction for wind farm based on VMD and weighted LS-SVM. Energy Sources A Recovery Util. Environ. Eff. *43*, 800–814.

261. Shrivastava, N.A., Lohia, K., and Panigrahi, B.K. (2016). A multiobjective framework for wind speed prediction interval forecasts. Renew. Energy *87*, 903–910.

262. Sideratos, G., and Hatziargyriou, N.D. (2012). Probabilistic wind power forecasting using radial basis function neural networks. IEEE Trans. Power Syst. *27*, 1788–1796.

263. Zhang, Z., Ye, L., Qin, H., Liu, Y., Wang, C., Yu, X., Yin, X., and Li, J. (2019). Wind speed prediction method using shared weight long short-term memory network and Gaussian process regression. Appl. Energy *247*, 270–284.

264. Taylor, J.W., McSharry, P.E., and Buizza, R. (2009). Wind power density forecasting using ensemble predictions and time series models. IEEE Trans. Energy Convers. *24*, 775–782.

265. Zhang, Y., Wang, J., and Wang, X. (2014). Review on probabilistic forecasting of wind power generation. Renew. Sustain. Energy Rev. *32*, 255–270.

266. Wu, Y.-K., Su, P.-E., and Hong, J.-S. (2016). An Overview of Wind Power Probabilistic Forecasts (IEEE), pp. 429–433.

267. Bludszuweit, H., Domínguez-Navarro, J.A., and Llombart, A. (2008). Statistical analysis of wind power forecast error. IEEE Trans. Power Syst. *23*, 983–991.

268. Yuan, X., Chen, C., Jiang, M., and Yuan, Y. (2019). Prediction interval of wind power using parameter optimized Beta distribution based LSTM model. Appl. Soft Comput. *82*, 105550.

269. Luig, A., Bofinger, S., and Beyer, H.G. (2001). Analysis of Confidence Intervals for the Prediction of Regional Wind Power Output, pp. 725–728.

270. Hodge, B.-M., and Milligan, M. (2011). Wind Power Forecasting Error Distributions over Multiple Timescales (IEEE), pp. 1–8.

271. Kiss, P., and Jánosi, I.M. (2008). Comprehensive empirical analysis of ERA-40 surface wind speed distribution over Europe. Energy Convers. Manag. *49*, 2142–2151.

272. Lange, M. (2005). On the uncertainty of wind power predictions—analysis of the forecast accuracy and statistical distribution of errors. J. Sol. Energy Eng. *127*, 177–184.

273. Aydin, D. (2018). Estimation of the lower and upper quantiles of Gumbel distribution: an application to wind speed data. Appl. Ecol. Env. Res. *16*, 1–15.

274. Ouarda, T., Charron, C., Shin, J.-Y., Marpu, P.R., Al-Mandoos, A.H., Al-Tamimi, M.H., Ghedira, H., and Al Hosary, T. (2015). Probability distributions of wind speed in the UAE. Energy Convers. Manag. *93*, 414–434.

275. Pinson, P. (2012). Very-short-term probabilistic forecasting of wind power with generalized logit—normal distributions. J. Roy. Stat. Soc. C *61*, 555–576.

276. Zhang, Y., Zhao, Y., Pan, G., and Zhang, J. (2020). Wind speed interval prediction based on lorenz disturbance distribution. IEEE Trans. Sustain. Energy *11*, 807–816.

277. Tewari, S., Geyer, C.J., and Mohan, N. (2011). A statistical model for wind power forecast error and its application to the estimation of penalties in liberalized markets. IEEE Trans. Power Syst. *26*, 2031–2039. https://doi.org/10.1109/TPWRS.2011.2141159.

278. Methaprayoon, K., Yingvivatanapong, C., Lee, W.-J., and Liao, J.R. (2007). An integration of ANN wind power estimation into unit commitment considering the forecasting uncertainty. IEEE Trans. Ind. Appl. *43*, 1441–1448.

279. Ali, S., Lee, S.-M., and Jang, C.-M. (2018). Statistical analysis of wind characteristics using Weibull and Rayleigh distributions in Deokjeok-do Island–Incheon, South Korea. Renew. Energy *123*, 652–663.

280. Zhang, Z.-S., Sun, Y.-Z., Gao, D.W., Lin, J., and Cheng, L. (2013). A versatile probability distribution model for wind power forecast errors and its application in economic dispatch. IEEE Trans. Power Syst. *28*, 3114–3125.

281. Celik, A.N. (2004). A statistical analysis of wind power density based on the Weibull and Rayleigh models at the southern region of Turkey. Renew. Energy *29*, 593–604.

282. Chang, T.P. (2011). Performance comparison of six numerical methods in estimating Weibull parameters for wind energy application. Appl. Energy *88*, 272–282. https://doi.org/10.1016/j.apenergy.2010.06.018.

283. Bremnes, J.B. (2004). Probabilistic wind power forecasts using local quantile regression. Wind Energy *7*, 47–54.

284. Wan, C., Xu, Z., Pinson, P., Dong, Z.Y., and Wong, K.P. (2014). Probabilistic forecasting of wind power generation using extreme learning machine. IEEE Trans. Power Syst. *29*, 1033–1044.

285. Ji, J., Sun, Y., Kong, F., and Miao, Q. (2019). A construction approach to prediction intervals based on bootstrap and deep belief network. IEEE Access *7*, 124185–124195.

286. Peng, X., Wang, H., Lang, J., Li, W., Xu, Q., Zhang, Z., Cai, T., Duan, S., Liu, F., and Li, C. (2021). EALSTM-QR: interval wind-power prediction model based on numerical weather prediction and deep learning. Energy *220*, 119692.

287. Khosravi, A., Nahavandi, S., Creighton, D., and Atiya, A.F. (2011). Lower upper bound estimation method for construction of neural network-based prediction intervals. IEEE Trans. Neural Netw. *22*, 337–346.

288. Zhang, G., Wu, Y., Wong, K.P., Xu, Z., Dong, Z.Y., and Iu, H.H.-C. (2015). An advanced approach for construction of optimal wind power prediction intervals. IEEE Trans. Power Syst. *30*, 2706–2715.

289. Ak, R., Vitelli, V., and Zio, E. (2015). An interval-valued neural network approach for uncertainty quantification in short-term wind speed prediction. IEEE Trans. Neural Netw. Learn. Syst. *26*, 2787–2800.

290. Qin, S., Liu, F., Wang, J., and Song, Y. (2015). Interval forecasts of a novelty hybrid model for wind speeds. Energy Rep. *1*, 8–16.

291. Kavousi-Fard, A., Khosravi, A., and Nahavandi, S. (2016). A new fuzzy-based combined prediction interval for wind power forecasting. IEEE Trans. Power Syst. *31*, 18–26.

292. Shen, Y., Lu, X., Yu, X., Zhao, Z., and Wu, D. (2016). Short-term Wind Power Intervals Prediction Based on Generalized Morphological Filter and Artificial Bee Colony Neural Network (IEEE), pp. 8501–8506.

293. Wang, J., Fang, K., Pang, W., and Sun, J. (2017). Wind power interval prediction based on improved PSO and BP neural network. J. Electr. Eng. Technol. *12*, 989–995.

294. Hu, M., Hu, Z., Yue, J., Zhang, M., and Hu, M. (2017). A novel multi-objective optimal approach for wind power interval prediction. Energies *10*, 419.

295. Wu, Y.K., Su, P.E., Wu, T.Y., Hong, J.S., and Hassan, M.Y. (2018). Probabilistic wind-power forecasting using weather ensemble models. IEEE Trans. Ind. Appl. *54*, 5609–5620. https://doi.org/10.1109/TIA.2018.2858183.

296. Li, R., and Jin, Y. (2018). A wind speed interval prediction system based on multi-objective optimization for machine learning method. Appl. Energy *228*, 2207–2220.

297. Jiang, P., Li, R., and Li, H. (2019). Multi-objective algorithm for the design of prediction intervals for wind power forecasting model. Appl. Math. Model. *67*, 101–122.

298. Tsao, H.-H., Leu, Y.-G., and Chou, L.-F. (2021). A center-of-concentrated-based prediction interval for wind power forecasting. Energy *237*, 121467.

299. Zhang, D., Chen, Z., and Zhou, Y. (2022). Wind power interval prediction based on improved whale optimization algorithm and fast learning network. J. Electr. Eng. Technol. *17*, 1785–1802.

300. Shi, Z., Liang, H., and Dinavahi, V. (2018). Direct interval forecast of uncertain wind power based on recurrent neural networks. IEEE Trans. Sustain. Energy *9*, 1177–1187.

301. Zhou, M., Wang, B., Guo, S., and Watada, J. (2021). Multi-objective prediction intervals for wind power forecast based on deep neural networks. Inform. Sci. *550*, 207–220.

302. Saeed, A., Li, C., Danish, M., Rubaiee, S., Tang, G., Gan, Z., and Ahmed, A. (2020). Hybrid bidirectional LSTM model for short-term wind speed interval prediction. IEEE Access *8*, 182283–182294.

303. Li, C., Tang, G., Xue, X., Chen, X., Wang, R., and Zhang, C. (2020). The short-term interval prediction of wind power using the deep learning model with gradient descend optimization. Renew. Energy *155*, 197–211.

304. Li, C., Tang, G., Xue, X., Saeed, A., and Hu, X. (2020). Short-term wind speed interval prediction based on ensemble GRU model. IEEE Trans. Sustain. Energy *11*, 1370–1380.

305. Wang, R., Li, C., Fu, W., and Tang, G. (2020). Deep learning method based on gated recurrent unit and variational mode decomposition for short-term wind power interval prediction. IEEE Trans. Neural Netw. Learn. Syst. *31*, 3814–3827.

306. Xie, Y., Li, C., Tang, G., and Liu, F. (2021). A novel deep interval prediction model with adaptive interval construction strategy and automatic hyperparameter tuning for wind speed forecasting. Energy *216*, 119179.

307. Gan, Z., Li, C., Zhou, J., and Tang, G. (2021). Temporal convolutional networks interval prediction model for wind speed forecasting. Elec. Power Syst. Res. *191*, 106865.

308. Liu, H., Han, H., Sun, Y., Shi, G., Su, M., Liu, Z., Wang, H., and Deng, X. (2022). Short-term wind power interval prediction method using VMD-RFG and Att-GRU. Energy *251*, 123807. https://doi.org/10.1016/j.energy.2022.123807.

309. Dong, Y., Zhang, H., Wang, C., and Zhou, X. (2021). A novel hybrid model based on Bernstein polynomial with mixture of Gaussians for wind power forecasting. Appl. Energy *286*, 116545.

310. Chen, C., and Liu, H. (2021). Dynamic ensemble wind speed prediction model based on hybrid deep reinforcement learning. Adv. Eng. Inform. *48*, 101290.

311. Xiong, D., Fu, W., Wang, K., Fang, P., Chen, T., and Zou, F. (2021). A blended approach incorporating TVFEMD, PSR, NNCT-based multi-model fusion and hierarchy-based merged optimization algorithm for multi-step wind speed prediction. Energy Convers. Manag. *230*, 113680.

312. Barbounis, T.G., Theocharis, J.B., Alexiadis, M.C., and Dokopoulos, P.S. (2006). Long-term wind speed and power forecasting using local recurrent neural network models. IEEE Trans. Energy Convers. *21*, 273–284.

313. Li, D., Yu, X., Liu, S., Dong, X., Zang, H., and Xu, R. (2022). Wind power prediction based on PSO-Kalman. Energy Rep. *8*, 958–968.

314. Liu, Z., and Ware, T. (2022). Capturing spatial influence in wind prediction with a graph convolutional neural network. Front. Environ. Sci. *10*. https://doi.org/10.3389/fenvs.2022.836050.

315. Tuerxun, W., Xu, C., Guo, H., Guo, L., Zeng, N., and Cheng, Z. (2022). An ultra-short-term wind speed prediction model using LSTM based on modified tuna swarm optimization

and successive variational mode decomposition. Energy Sci. Eng. *10*, 3001–3022.

316. Wei, L., Xv, S., and Li, B. (2022). Short-term wind power prediction using an improved grey wolf optimization algorithm with back-propagation neural network. Clean Energy *6*, 288–296.

317. Niu, Z., Yu, Z., Tang, W., Wu, Q., and Reformat, M. (2020). Wind power forecasting using attention-based gated recurrent unit network. Energy *196*, 117081.

318. Huang, Y., Zhang, B., Pang, H., Wang, B., Lee, K.Y., Xie, J., and Jin, Y. (2022). Spatio-temporal wind speed prediction based on Clayton Copula function with deep learning fusion. Renew. Energy *192*, 526–536.

319. Wang, Y., Xiong, W., Liu, Q., Yang, N., Fu, P., Gong, K., and Huang, Y. (2022). Wind power prediction based on A hybrid granular chaotic time series model. Front. Energy Res. *9*. https://doi.org/10.3389/fenrg.2021.823786.